

# Identification of constitutional *MLH1* epimutations and promoter variants in colorectal cancer patients from the Colon Cancer Family Registry

Robyn L. Ward, PhD<sup>1</sup>, Timothy Dobbins, PhD<sup>1,2</sup>, Noralane M. Lindor, PhD<sup>3</sup>, Robert W. Rapkins, PhD<sup>1</sup> and Megan P. Hitchins, PhD<sup>1</sup>

**Purpose:** Constitutional *MLH1* epimutations manifest as promoter methylation and silencing of the affected allele in normal tissues, predisposing to Lynch syndrome-associated cancers. This study investigated their frequency and inheritance.

**Methods:** A total of 416 individuals with a colorectal cancer showing loss of *MLH1* expression and without deleterious germline mutations in *MLH1* were ascertained from the Colon Cancer Family Registry (C-CFR). Constitutive DNA samples were screened for *MLH1* methylation in all 416 subjects and for promoter sequence changes in 357 individuals.

**Results:** Constitutional *MLH1* epimutations were identified in 16 subjects. Of these, seven (1.7%) had mono- or hemi-allelic methylation and eight had low-level methylation (2%). In one subject the epimutation was linked to the c.-27C>A promoter variant. Testing

of 37 relatives from nine probands revealed paternal transmission of low-level methylation segregating with a c.+27G>A variant in one case. Five additional probands had a promoter variant without an *MLH1* epimutation, with three showing diminished promoter activity in functional assays.

**Conclusion:** Although rare, sequence changes in the regulatory region of *MLH1* and aberrant methylation may alone or together predispose to the development of cancer. Screening for these changes is warranted in individuals who have a negative germline sequence screen of *MLH1* and loss of *MLH1* expression in their tumor.

*Genet Med* 2013;15(1):25–35

**Key Words:** colorectal cancer; epimutation; Lynch; methylation; *MLH1*

## INTRODUCTION

Lynch syndrome is a familial cancer syndrome caused by germline heterozygous loss-of-function mutations of the DNA mismatch repair (MMR) genes, most frequently within *MLH1* and *MSH2*.<sup>1</sup> It is characterized by the development of colorectal (CRC), endometrial, and other cancers often at a young age. These tumors are MMR deficient, due to a second hit to the MMR gene, and show microsatellite instability and loss of expression of the relevant MMR protein.<sup>2</sup> Germline sequence mutations in the MMR genes are not found in 25–30% of patients with suspected Lynch syndrome, and a proportion of these mutation-negative individuals have been shown to carry constitutional epimutations of the *MLH1* or *MSH2* genes.<sup>3–8</sup> Constitutional epimutations manifest as promoter methylation and loss of transcription from one genetic allele within normal somatic tissues.<sup>9</sup> They are now well recognized as an alternate mechanism for inactivation of tumor suppressor genes and thus predisposition to cancer.

It is also clear that cancer risk among family members is influenced by whether the disease is caused by an epigenetic aberration or a conventional genetic (DNA sequence) mutation within an MMR gene. In contrast to sequence mutations, constitutional epimutations may demonstrate somatic mosaicism,

defined as heterogeneity in the level of methylation between alleles or tissues. Depending on their underlying cause, epimutations may show Mendelian or non-Mendelian patterns of inheritance.<sup>9</sup>

*MSH2* constitutional epimutations typically show tissue-specific mosaicism, with increased methylation in epithelial tissues such as colonic mucosa, and little or no methylation in other tissues.<sup>6</sup> Because they are secondary to *cis*-acting germline deletions of the upstream *EPCAM* gene,<sup>7</sup> *MSH2* epimutations are inherited in an autosomal dominant manner.

In contrast, constitutional epimutations of *MLH1* are more variable, with at least two distinctive types described to date. The first type, which for the purposes of brevity will be referred to as a “primary” constitutional epimutation, typically shows dense hemiallelic methylation of the *MLH1* promoter throughout all somatic tissues,<sup>10,11</sup> arises spontaneously,<sup>8,12</sup> and is reversed between generations because of erasure in the germline.<sup>13</sup> In the rare instances where intergenerational inheritance of a primary *MLH1* epimutation occurs,<sup>14</sup> it follows a non-Mendelian pattern. In all cases, primary *MLH1* epimutations occur in the absence of any apparent linked sequence change.

The second type, described here as a “secondary” constitutional *MLH1* epimutation, is caused by a *cis*-acting DNA

<sup>1</sup>Adult Cancer Program, Lowy Cancer Research Centre, Prince of Wales Clinical School, Faculty of Medicine at the University of New South Wales, Kensington, Australia;

<sup>2</sup>Sydney School of Public Health, University of Sydney, Sydney, Australia; <sup>3</sup>Department of Medical Genetics, Mayo Clinic College of Medicine, Rochester, Minnesota, USA.

Correspondence: Robyn L. Ward (roblyn@unsw.edu.au)

Submitted 20 January 2012; accepted 21 June 2012; advance online publication 9 August 2012. doi:10.1038/gim.2012.91

sequence alteration, and therefore displays a Mendelian inheritance pattern. In this regard, we recently reported a three-generation Lynch syndrome family with autosomal dominant inheritance of a constitutional *MLH1* epimutation likely caused by the linked c.-27C>A single-nucleotide variant (SNV) within the *MLH1* promoter.<sup>15</sup> The epimutation in this family showed marked somatic mosaicism between tissues and family members. Other examples of secondary epimutations include a family with a mosaic constitutional *MLH1* epimutation that was attributed to a large duplication encompassing the entire *MLH1* gene,<sup>12,16</sup> and an individual with a family history of Lynch syndrome-associated cancers found to have a 6.4kb deletion of *MLH1* from c.-67 to intron 2.<sup>17</sup>

To date, our understanding of the etiology and clinicopathologic consequences of aberrant constitutional *MLH1* methylation have been informed by case studies. To overcome the bias of single-institution family studies, the Colon Cancer Family Registry (C-CFR) was established in 1997 to provide data and biological samples from colorectal cancer patients.<sup>18</sup> Case ascertainment to this resource is from both cancer registries (population-based families) and cancer family clinics (clinic-based families) from the United States, Australia, and Canada. This study uses this resource to determine the frequency of *MLH1* epimutations in individuals without coding mutations in the *MLH1* gene but with an *MLH1*-deficient colorectal cancer. It also assesses the inheritance patterns and clinical phenotypes associated with primary and secondary *MLH1* epimutations, and whether promoter sequence variations are associated with epimutations.

## MATERIALS AND METHODS

### Study cohorts

A dataset (definitions at <https://cfrisc.georgetown.edu/isc/welcome.do>) and constitutive DNA was provided by the C-CFR from 416 individuals ascertained from population- or clinic-based registries through six study centers (Cancer Care, Ontario, Canada; Fred Hutchinson Cancer Research Center, Seattle, WA; Mayo Clinic, Rochester, MN; University of Southern California Consortium, Los Angeles, CA; Universities of Queensland and Melbourne, Australia; and University of Hawaii, Honolulu, HI). Patients were eligible if they had cancer of the colorectum (ICD codes, C18.0–18.9, C19.9, and C20.9), intestinal tract (C26.0), or small intestine (C17.0–17.2 and C17.9). In addition, the cancer must have fulfilled one of two criteria: (i) loss of expression of *MLH1* and either microsatellite instability or unknown microsatellite status; (2) microsatellite instability where immunostaining was not done or not interpretable. In addition to these criteria, the person or their genetic relatives must not have a deleterious mutation in the mismatch repair genes *MLH1*, *MSH2*, *MSH6*, or *PMS2*, or the base excision repair gene *MYH*. Individuals who met criterion 1 or 2, but in whom germline sequence and/or MLPA results did not provide a conclusive normal result were also included. All studies were approved by the Human Investigations Committee and informed consent was obtained from each participant or guardian of each participant.

### Methylation screening analysis, confirmatory assays, and allelic methylation analysis

A total of 416 constitutive DNA samples were screened (Supplementary Figure S1 online) for methylation by the sensitive method of real-time quantitative methylation-specific PCR (qMSP) followed by temperature gradient denaturation analysis of the C-region of the *MLH1* promoter, using primer sequences 5'-TAAATCTCTTCGTCCCTCCCTAAAACG-3' and 5'-CGTTAAGTATTTTTTTCGTTTTGCG-3' at an annealing temperature of 60 °C, as previously applied.<sup>11,19</sup> Methylation levels were calculated against a methylation-independent *MyoD* qMSP run in parallel, and scored as the percentage of methylated reference (PMR) value with reference to the CpGenome Universal Methylated DNA (Chemicon, Merck Millipor, Darmstadt, Germany) control, as previously described.<sup>20</sup> This method has the analytical sensitivity to reliably detect methylation at a PMR level of 0.01. Gel electrophoresis of qMSP products was performed to confirm specific amplification in positive samples that demonstrated the correct denaturation temperature. Furthermore, qMSP products from samples with low-level methylation (PMR 1.5–5%) detected were cloned and sequenced to confirm the presence of methylation on these few alleles.<sup>21</sup> Samples positive in the qMSP screen were tested by CpG pyrosequencing,<sup>10</sup> to verify the qMSP results and to quantify methylation levels more accurately. Following pyrosequencing, “combined bisulphite restriction analysis” and clonal (allelic) bisulfite of “combined bisulphite restriction analysis” products was then undertaken in samples that had methylation levels >10% or where a sequence variant within the promoter or first exon had been co-identified in a methylation-positive case to ascertain allelic methylation patterns.<sup>8,15</sup>

### Sequencing the *MLH1* promoter

Where more than 500 ng of constitutive DNA was provided by the C-CFR (357 of 416 samples), the *MLH1* promoter region was sequenced to show the promoter was intact for methylation studies and to identify any potentially pathogenic promoter variants, or single-nucleotide polymorphisms for use in allelic methylation analyses (Supplementary Materials and Methods online).

Where a promoter sequence alteration or methylation of *MLH1* was identified, we requested a second normal DNA sample from the proband (preferably from a different tissue source), constitutive DNA from relatives, and a copy of the pedigree. At the completion of the laboratory studies, the molecular data were merged with the data set containing the clinicopathologic variables from the C-CFR databases.

### Luciferase promoter reporter assays

To evaluate the effect of the promoter sequence changes identified on transcriptional activity, luciferase promoter reporter assays were performed on each variant by transient transfection of promoter reporter constructs into HCT116 colorectal carcinoma and HEK293 human embryonal kidney cells (Supplementary Materials and Methods online).

## RESULTS

### Characteristics of study cohorts

The 416 study participants (mean age 58.8 years, range 18–90 years), were mostly Caucasian ( $n = 376$ ; 90.4%), female ( $n = 259$ ; 62.3%) and had a right sided CRC ( $n = 347$ ; 83.4%). Individuals were mainly recruited through population-based registries ( $n = 330$ ; 79.3%). In terms of family history, few families met the Amsterdam I ( $n = 7$ ) or II ( $n = 8$ ) criteria and 95 of the 333 individuals on whom data were available had at least one first-degree relative with CRC. Results of germline screening for *MLH1* mutations were unavailable from the C-CFR for 69 individuals, and the characteristics of this group are shown as cohort 2 (Table 1).

### Identification of constitutional *MLH1* methylation and promoter sequence mutations

Twenty-one individuals with either sequence changes or methylation of the *MLH1* promoter were identified from the 416 assessable patients (5.0%, Table 2). The positive cases were grouped into one of five categories based on the level of *MLH1* methylation and any association with promoter sequence changes (Table 2): (i) primary epimutation (qMSP PMR and pyrosequencing >10%) (7/416, 1.7%), (ii) epimutation in association with a promoter DNA sequence variant (1/416, 0.2%), (iii) definite low-level *MLH1* methylation (qMSP PMR range 1.5–5% and confirmed by pyrosequencing) (4/416, 1.0%), (iv) possible low-level methylation (qMSP PMR range >0.01 to <1.5% but unconfirmed by less sensitive pyrosequencing) (4/416, 1.0%), and (v) promoter variant in the absence of *MLH1* methylation (5/357, 1.4%).

### Cases with a primary constitutional *MLH1* epimutation

The clinical, family history, and methylation results for the seven individuals with primary epimutations are shown in Table 2 and Figure 1. Five of these cases had methylation levels approaching 50% (proband A, C, E, F, and G), indicating the majority of copies of the affected allele were methylated. Monoallelic methylation was confirmed in four of these cases heterozygous for the c.-93G>A single-nucleotide polymorphism (A, C, E, and G) (Figure 1; Table 2). Two (B and D) had methylation levels below 20%, indicating mosaicism (Table 2). Proband B was heterozygous for the c.-269C>G single-nucleotide polymorphism and monoallelic methylation was demonstrated (Figure 1). In uninformative proband D, dense hypermethylation of a proportion of alleles was confirmed. In contrast to the study population overall, a disproportionate number of individuals (three) were ascertained from the study center in Australia and through clinic-based (rather than population-based) recruitment. Given our past work on epimutations, we obtained permission to cross-match the Australian patients identified in the current study against the identity of our previously reported individuals. In doing so, we identified that persons F and G in Table 2 had been previously reported by our group.<sup>8,14</sup> Thus five of seven cases in the current study are newly identified.

### An epimutation associated with the c.-27C>A and c.85G>T haplotype

One proband had *MLH1* methylation in association with a c.-27C>A promoter variant (H; Table 2, Supplementary Figure S2 online). This haplotype has been previously linked to constitutional *MLH1* epimutation in another unrelated Australian family.<sup>15</sup> A comparison of the two extended pedigrees of both families did not reveal an obvious connection, suggesting these are two distinct families bearing the same variant haplotype.

### Individuals with mosaic or low-level constitutional methylation

Four probands (N–Q, Table 2) were identified with low but confirmed methylation of the *MLH1* promoter (Figure 2). In one proband methylation occurred in association with a variant sequence change (c.+27G>A), which is discussed below.

A final group of four potential positive probands (R–U) were identified with only trace levels of methylation detected by qMSP (Table 2). This was not confirmed by other methods due to limitations of analytical sensitivity, and the cause of the positive findings is unknown. This group is not discussed further.

### Transmission of epimutations in family members

DNA from family members was obtained from nine positive probands (six with primary epimutations, two with confirmed low-level methylation and one with unconfirmed trace-level of methylation, see Table 2, and indicated by filled diamonds in Figures 1 and 2). The methylation status in relatives of Proband F (two individuals),<sup>14</sup> and Proband G (six relatives),<sup>8</sup> have been previously reported as negative. No methylation was detected in 14 relatives from the other four probands with primary epimutations (Probands B–E, Table 2 and Figure 1). Of the probands with confirmed low-level methylation, only N and Q had relatives available for testing (Figure 2). No methylation was found in the family of Q. Of note, however, the son of proband N, who had died at 22 years from CRC, also carried the variant c.+27 G> A allele and was also positive for a similarly low level of methylation (5% by CpG pyrosequencing) in his constitutional DNA. The other five relatives tested in this family did not harbor the variant allele or have *MLH1* methylation, yet two of them had CRC (microsatellite status unknown, ages 49 and 61 years). The two relatives from proband R, who carried trace levels of methylation, were also negative for methylation.

### *MLH1* promoter sequence changes without constitutional methylation

Five probands (I–M, Table 2) from the population-based registry in Seattle had a sequence variant within the promoter or 5'UTR of *MLH1* without any methylation (Table 2 and Figure 3). These comprised two SNVs, c.-11C>T and c.-42C>T and two novel microdeletions, delc.-411-413GAG and delc.-432-435AAAG. Because RNA was unavailable to directly assess whether the promoter variants caused a reduction in allelic expression in the subjects, luciferase promoter reporter assays were instead used to assess their effect on promoter

**Table 1** Summary of characteristics of study participants

		Cohort 1 (N = 347)		Cohort 2 (N = 69)	
		N	%	N	%
Recruitment site	Seattle	116	33.4%	50	72.5%
	Ontario	71	20.5%	16	23.2%
	Australia	77	22.2%	3	4.3%
	Mayo	66	19.0%	0	0%
	Hawaii	14	4.0%	0	0%
	LA	3	0.9%	0	0%
Registry	Population-based	267	76.9%	63	91.3%
	Clinic-based	79	22.8%	6	8.7%
	Missing	1	0.3%	0	0%
Race	Caucasian	308	89.5%	68	92.1%
	Japanese	9	2.6%	0	0%
	Latino	4	1.2%	0	0%
	Polynesian	3	0.9%	0	0%
	Korean/South Asian	5	1.4%	1	0%
	Black Af Am/Carrib	4	1.2%	0	0%
	Unknown	14	3.2%	0	0%
Sex	Female	217	62.5%	42	60.9%
	Male	129	37.2%	27	39.1%
	Missing	1	0.3%	0	0%
Age group	<30	7	2.0%	2	2.9%
	30–39	29	8.4%	6	8.7%
	40–49	51	14.7%	12	17.4%
	50–59	71	20.5%	5	7.2%
	60–69	107	30.8%	23	33.3%
	70+	81	23.3%	21	30.4%
	Missing	1	0.3%	0	0%
Primary site	Right-sided	288	83.0%	59	85.5%
	Left	48	13.8%	9	13.0%
	Rectum	9	2.6%	1	1.4%
	Intestine NOS	1	0.3%	0	0%
	Missing	1	0.3%	5	6.6%
MSI and/or loss of MLH1 IHC	Both	257	74.1%	41	59.4%
	One	90	25.9%	28	40.6%
	Missing	0	0%	0	0%
Synch or metachro-nous Lynch tumor any age	No	279	80.4%	59	85.5%
	Yes	39	11.2%	10	14.5%
	Missing	29	8.4%	0	0%
Proband's FDR history of colorectal cancer	0 affected FDRs	182	52.4%	48	69.6%
	1 affected FDR	57	16.4%	10	14.5%
	2 affected FDRs	22	6.3%	1	1.4%
	≥3 affected FDRs	4	1.2%	1	1.4%
	Missing	80	23.1%	3	4.3%

Cohort 1 includes individuals with an *MLH1* mismatch repair, deficient CRC, and no deleterious germline sequence mutations in this gene. Cohort 2 includes individuals on whom data were missing on germline *MLH1* testing.

Af Am, African American; Carrib, Caribbean; CRC, colorectal cancer; FDR, first-degree relative; IHC, immunohistochemistry; LA, Los Angeles; MSI, microsatellite instability; NOS, not otherwise specified.

**Table 2** Clinicopathological characteristics of patients with a constitutional epimutation and/or promoter variant of *MLH1* categorized by the level of *MLH1* promoter methylation and its association with promoter sequence changes

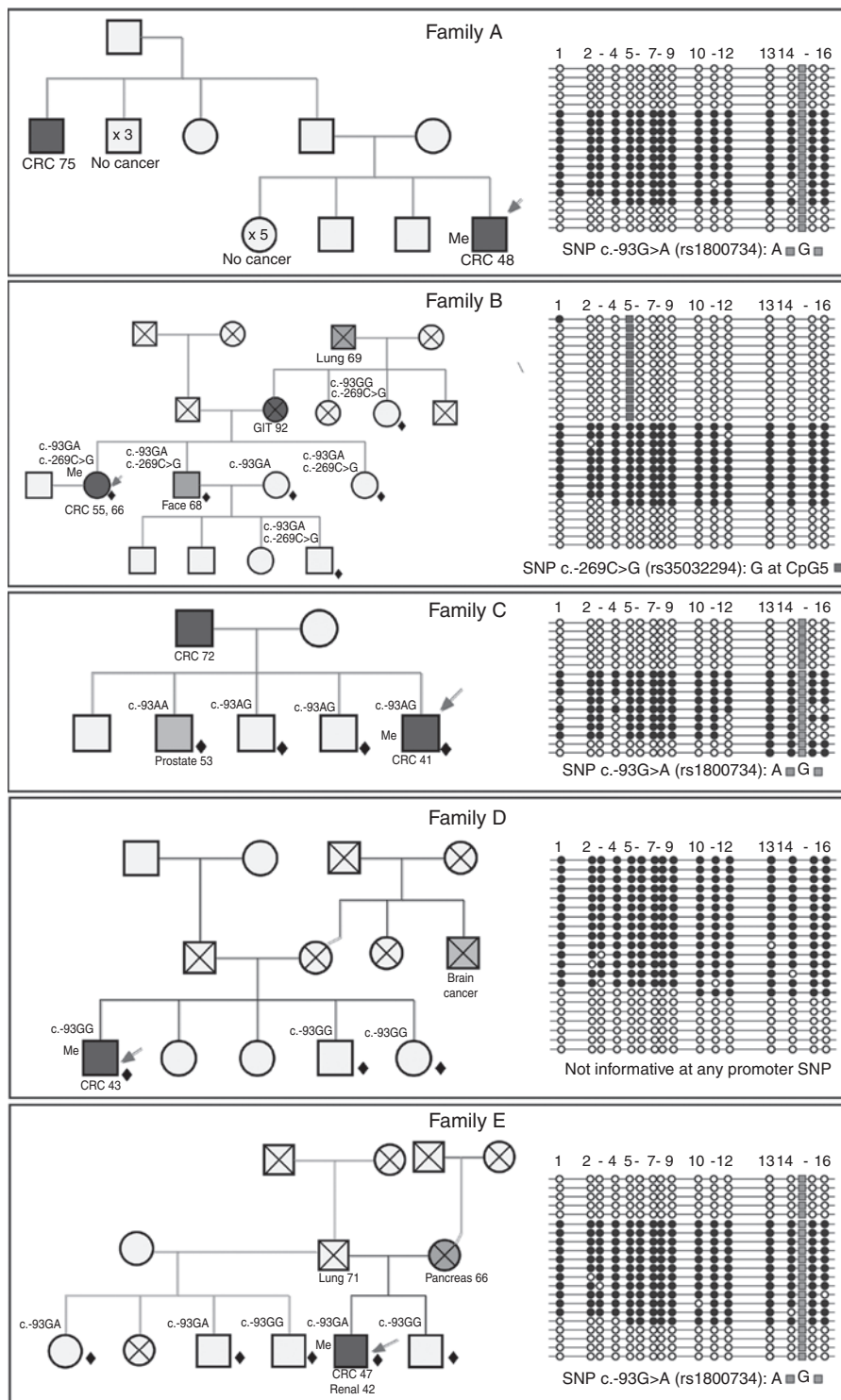
Center	Reg	Cohort	Sex	Race	Number of tumors	Age of onset	Tumor sites	MMR-deficient MSI/IHC	Methylation level	c.-93 SNP	Promoter variant	Number of relatives for testing	Number of relatives with methylation	
Primary constitutional epimutation														
A	Ontario	Pop	1	M	Filipino	1	48	R/S	One	36%	GA <sup>Me</sup>	None	Nil	N/A
F	Aust	Clinic	1	M	Caucasian	1	38	TC	Both	49%	GG <sup>NI</sup>	None	2	0
G	Aust	Clinic	1	F	Caucasian	4	45, 59, 60, 68	Uterus, AC, Rec, CT	Both	50%	GA <sup>Me</sup>	None	6	0
B	Aust	Clinic	1	F	Caucasian	1	55, 66	Colon, TC	One	17%	GG <sup>NI</sup>	None, SNP, c.-269C <sup>Me</sup> >G	5	0
C	Hawaii	Pop	1	M	Japanese	1	41	R/S	Both	45%	G <sup>Me</sup> A	None	3	0
D	Seattle	Pop	1	M	Caucasian	1	43	Cecum	Both	11–15%	GG <sup>NI</sup>	None	2	0
E	Seattle	Pop	1	M	Caucasian	2	42, 47	Kidney, AC	One	46%	G <sup>Me</sup> A	None	4	0
Secondary constitutional epimutation occurring in association with a promoter variant														
H	Aust	Clinic	1	F	Caucasian	1	43	AC	One	32%	GG <sup>NI</sup>	c.-27-C>A <sup>Me</sup> and c.85G>T <sup>Me</sup>	Nil	N/A
Promoter variant, no methylation														
I	Seattle	Pop	1	M	Caucasian	1	66	TC	Both	0%	AA	c.-11C>T	Nil	N/A
J	Seattle	Pop	1	F	Caucasian	1	53	Cec	Both	0%	GA	c.-42C>T	Nil	N/A
K	Seattle	Pop	2	F	Caucasian	3	64, 66, 66	Cec, 2x uterus	Both	0%	GA	Del-411-413GAG	Nil	N/A
L	Seattle	Pop	1	F	African Am	1	55	AC	Both	0%	GG	Del -432-435AAAG	Nil	N/A
M	Seattle	Pop	2	M	Caucasian	1	34	Cec	One	0%	GA	c.-11C>T	Nil	N/A
Low-level methylation certain														
N	Ontario	Pop	1	M	Caucasian	3	50, 52, 54	DC, 2x cec	Both	2–4%	GA	c.+27G>A <sup>Me</sup>	7	1
O	Ontario	Pop	1	F	Caucasian	1	45	SC	One	4%	GA	None	Nil	N/A
P	Aust	Clinic	1	F	Caucasian	1	69	AC	One	2%	G <sup>Me</sup> A	None	Nil	N/A
Q	Aust	Pop	1	M	Caucasian	2	39, 44	AC, skin	Both	3%	GG	None	6	0
Low-level methylation uncertain														
R	Ontario	Pop	1	M	Caucasian	1	40	AC	One	1.2%	GA	None	2	0
S	Seattle	Pop	1	F	Caucasian	1	63	Cecum	Both	1%	GG	None	Nil	N/A
T	Seattle	Pop	1	F	Caucasian	1	64	AC	Both	1%	GG	None	Nil	N/A
U	Seattle	Pop	1	M	Caucasian	1	42	TC	Both	1%	GG	None	Nil	N/A

*MLH1* constitutional epimutations manifest as promoter methylation and loss of transcription from one genetic allele within normal somatic tissues. Epimutations may be primary, which manifest as hemiallelic *MLH1* promoter methylation in all somatic tissues, or secondary to a *cis*-acting DNA sequence change. Individuals with low-level *MLH1* methylation were also identified in this study. Eligible is yes if the person met the pre-specified criteria for inclusion in the study and no if they did not. Colorectal cancer (CRC)–AC, ascending colon; cec, cecum, CT, connective tissue; DC, descending colon; R/S, rectosigmoid; TC, transverse colon; SC, sigmoid. MMR-deficient status is based on MSI and IHC results only for colon cancers, if one CRC is MSI high and IHC shows loss of expression of *MLH1* the person is labeled both, if either MSI or IHC is abnormal the person label is one. Person R, no methylation was detected in a second DNA sample derived from normal colonic mucosa. Persons K and M were part of cohort 2, as results of germline sequencing or MLPA for the mismatch repair genes were not provided.

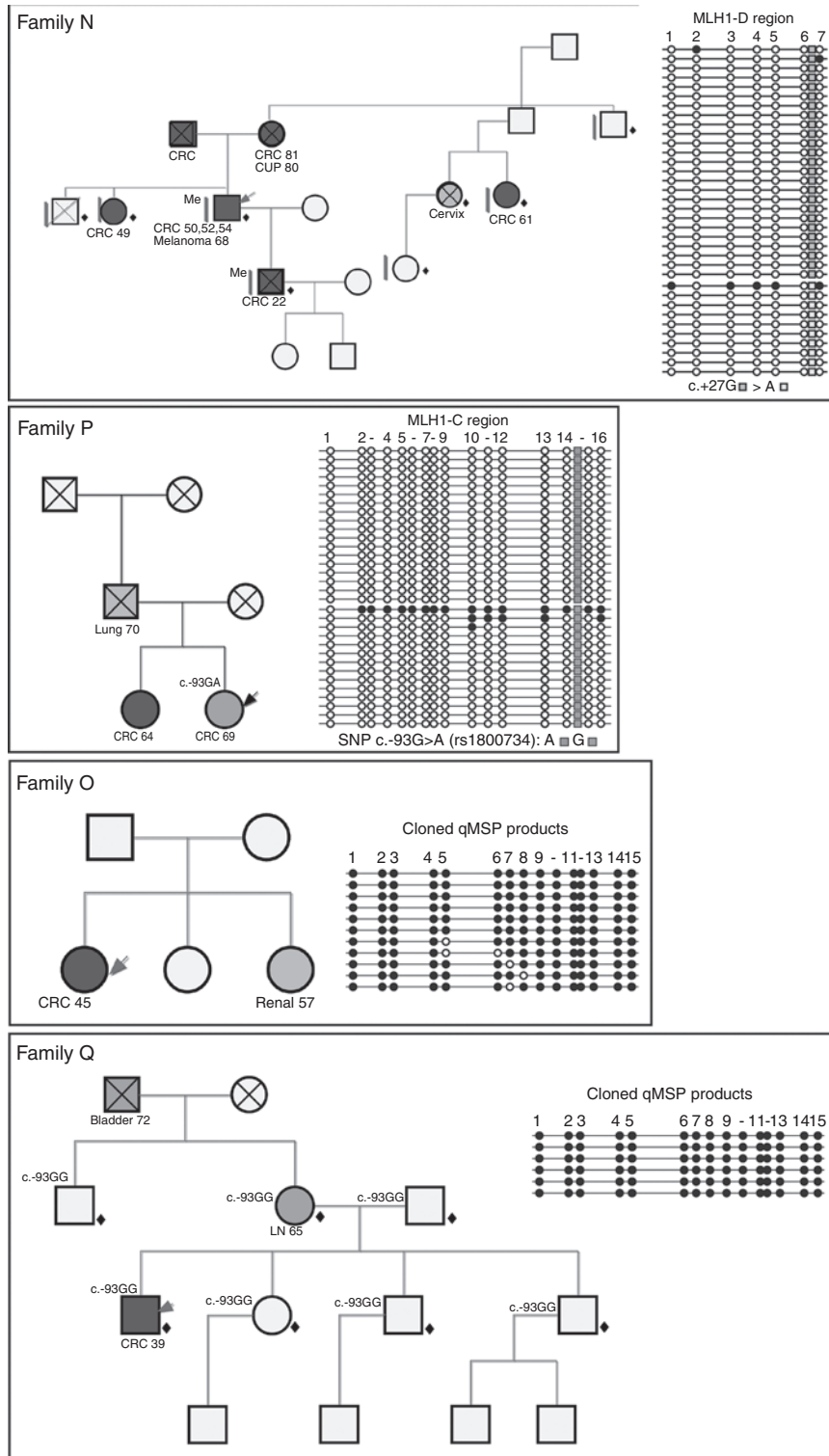
Am, American; Aust, Australia; F, female; IHC, immunohistochemistry; M, male; <sup>Me</sup>, methylated allele; MMR, mismatch repair; MSI, microsatellite instability; N/A, not available; <sup>NI</sup>, not informative; Pop, population; Reg, registry; SNP, single-nucleotide polymorphism.

activity (Figure 3b). The c.-11C>T and c.-42C>T SNVs showed a major reduction in promoter activity as compared with the wild-type promoter sequence, whereas the delc.-411-413GAG microdeletion showed a statistically significant but lesser

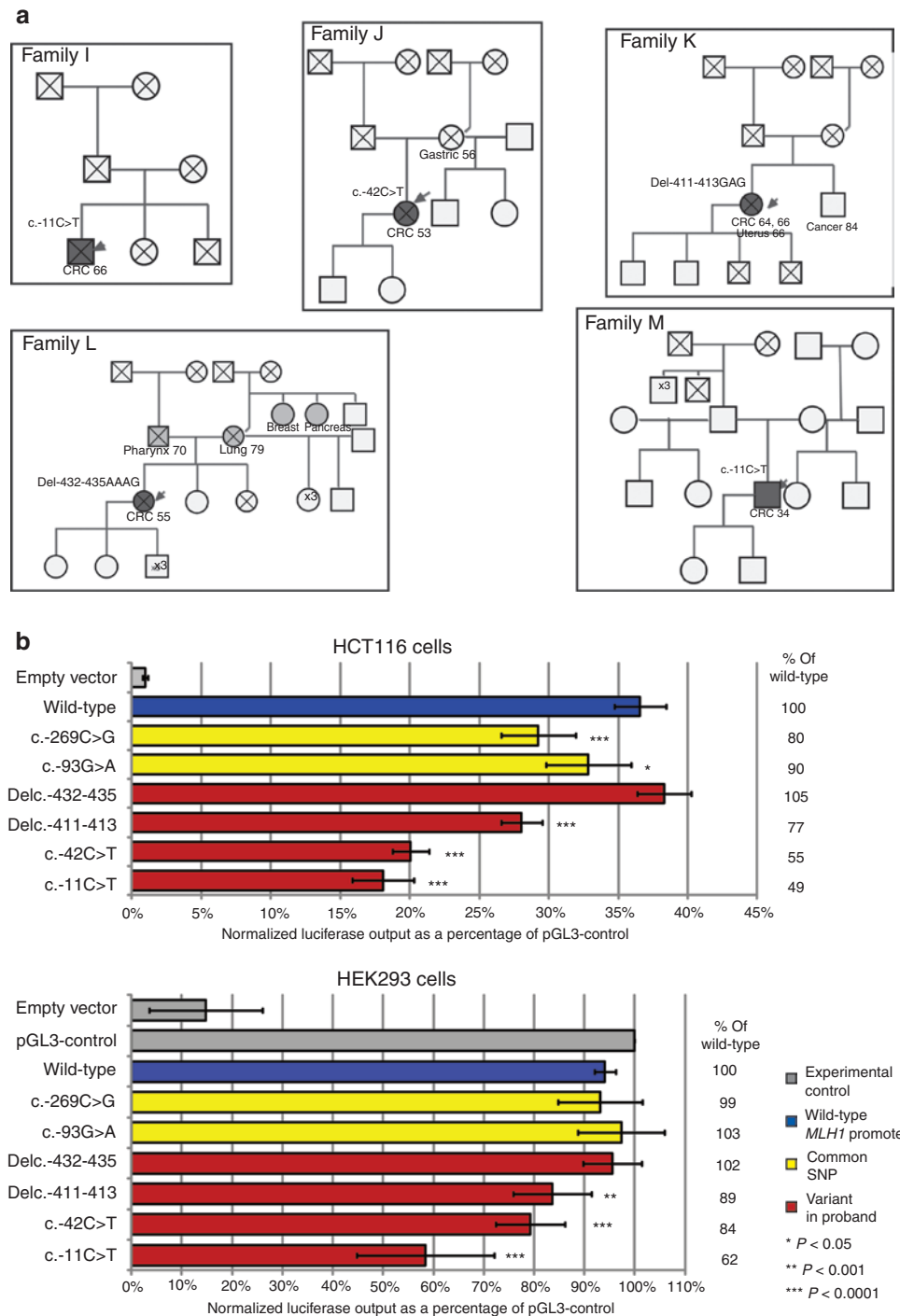
degree of transcriptional reduction. These findings were consistent in transient transfections of both HCT116 and HEK293 cell lines. The delc.-432-435AAAG had no detrimental effect on promoter activity in these assays.



**Figure 1** Pedigrees and allelic bisulfite sequencing of individuals with primary *MLH1* epimutations (family A–E). Pedigree: Family labels correspond with Table 2. The proband is shown by an arrow, squares are males, circles are females, and crosses indicate deceased individuals. The dark shading represents individuals with CRC, the gray shading is for other cancers, and age at diagnosis is shown. The small diamonds denote relatives on whom DNA was available for testing. Genotype at the c.-93G>A SNP is shown adjacent to the person symbol. Me = *MLH1* epimutation. All individuals with primary epimutations had young-onset CRC and an unremarkable family history of CRC. Allelic bisulfite sequencing (right panel) confirmed monoallelic methylation in three individuals who were heterozygous for the c.-93G>A SNP and one who was heterozygous for the c.-269C>G SNP. Hemiallelic methylation was observed in another individual who was not informative for either SNP. Each horizontal line represents a single DNA strand and circles represent individual CpG sites (filled = methylated; open = unmethylated). CRC, colorectal cancer; SNP, single-nucleotide polymorphism.



**Figure 2 Pedigrees and bisulfite sequencing of individuals with confirmed low-level *MLH1* methylation (families N–Q).** The family labels correspond with Table 2. The proband is shown by an arrow, squares are males, circles are females, and crosses indicate deceased individuals. The dark shading represents individuals with CRC; the gray shading is for other cancers. The small diamonds denote relatives on whom DNA was available for testing. The genotype at the  $c.-93G>A$  SNP is shown adjacent to the person symbol. Me = *MLH1* epimutation. <sup>Me</sup> shows the association between sequence change ( $c.+27A$ ) and methylation in family N. Bisulfite sequencing on the right, shows in two index cases from families N and P, methylation of a small number of alleles detected in combined bisulphite restriction analysis fragments in which both methylated and unmethylated alleles are sequenced equally, indicating extensive epigenetic mosaicism. In the other two (proband from families O and Q), the presence of low-level methylation was confirmed by cloning and sequencing the amplified qMSP products, which solely amplify methylated templates, confirming the presence of methylation in a small number of alleles. CRC, colorectal cancer; SNP, single-nucleotide polymorphism.



**Figure 3 Pedigrees of individuals with sequence changes in the *MLH1* promoter without constitutional methylation (families I–M) and their effect on promoter activity.** (a) The family labels correspond with Table 2. The proband is shown by an arrow, squares are males, circles are females, and crosses indicate deceased individuals. The dark shading represents individuals with CRC, the gray shading is for other cancers. The sequence variant for each family is shown. Apart from one case (aged 34 years), the probands were over 50 years of age, all had right-sided CRC, and one person had multiple cancers including uterine cancer. (b) The promoter activity of a luciferase reporter construct containing the entire wild-type (blue) *MLH1* promoter and 5'-UTR (from c.-513 to c.-1 with respect to the ATG start codon), and constructs containing the promoter sequence variants c.-11C>T, c.-42C>T, delc.-411-413GAG, and delc.-432-435AAAG (red) and also the SNPs c.-269C>G and c.-93G>A (yellow) in transient transfections of HCT116 and HEK293 cells are shown. The c.-11C>T and c.-42C>T SNVs each resulted in a major reduction in promoter activity as compared with the wild-type promoter sequence. The delc.-411-413GAG microdeletion showed a lesser degree of transcriptional reduction (77–89% of the wild type), although this was statistically significant, whereas the delc.-432-435AAAG appeared to have no detrimental effect on promoter activity in these assays. The c.-269C>G and c.-93G>A SNPs also demonstrated marginally diminished promoter activity specifically in HCT116 cells. Values are shown as the mean  $\pm$  SD of the luciferase output normalized to a cotransfected *Renilla* control, and expressed as a percentage of the positive control vector, for multiple independent experiments performed in triplicate. CRC, colorectal cancer; SNP, single-nucleotide polymorphism; SNV, single-nucleotide variant.



## DISCUSSION

Through screening a large multinational cohort of individuals with MMR-deficient CRC, this study has revealed the frequency and spectrum of changes associated with the phenomenon of constitutional *MLH1* epimutation and identified potentially pathogenic promoter sequence variants. We have also demonstrated the complex relationship between sequence variations in the regulatory region of *MLH1* and methylation of the promoter.

Since the first description of primary constitutional *MLH1* epimutations, opinion has been divided as to the circumstances, if any, which justify screening individuals for epimutations. The case for screening relates in part to the likelihood of detecting a positive case, yet the reported variation in the frequency of primary *MLH1* epimutations ranges from 0.7<sup>4</sup> to 13.3%.<sup>22</sup> These differences reflect the selection criteria used in various studies, including *BRAF* tumor status, family history, and age. In the current study, seven individuals (<2.0%) with an *MLH1*-deficient CRC and no germline sequence mutations displayed all the characteristic features of primary *MLH1* epimutations. Specifically, their DNA displayed dense mono- or hemi-allelic methylation of the *MLH1* promoter, or dense methylation of a proportion of alleles in a mosaic case uninformative for any promoter single-nucleotide polymorphism, in the absence of any promoter sequence change. All were identified because of suspicion of having Lynch syndrome, having developed a MMR-deficient CRC. The five new cases identified herein had an unremarkable family history, and we found no *MLH1* methylation in any of their assessable genetic relatives. Although lack of biological samples from these five new cases meant we could not formally confirm loss of function of the wild-type allele in tumor, or of the methylated allele in normal tissue, our previous findings in similar cases suggest the epimutation as the likely cause of their cancer predisposition.<sup>9</sup>

The most compelling evidence of a link between methylation and sequence variants within the promoter was our finding of an individual with an extensive maternal family history of cancer and allele-specific methylation of the variant c.-27C>A and c.85G>T haplotype. We have previously reported autosomal dominant inheritance of a mosaic constitutional *MLH1* epimutation linked to this very haplotype in another Caucasian Australian family.<sup>15</sup> This new proband thus represents the second case of constitutional methylation in association with the c.-27C>A and c.85G>T haplotype, providing further evidence that this genetic haplotype confers cancer susceptibility via a mechanism of epigenetic suppression. Unfortunately, no samples were available from relatives of this new proband to establish segregation of the haplotype with methylation or disease status. A third Lynch syndrome family in which this haplotype was first reported was not tested for methylation.<sup>23</sup> The independent findings of the two linked (c.-27C>A and c.85G>T) substitutions in three seemingly unrelated Lynch-like families raises the possibility of a founder disease-susceptibility haplotype. We propose that methylation studies on a number of *MLH1* promoter SNVs in suspected Lynch syndrome cases<sup>24-26</sup> may help resolve their pathogenicity.

The individuals with confirmed low-level methylation in peripheral blood without a genetic change within the *MLH1* promoter had dense methylation present on only a small proportion of alleles (epigenetic mosaicism). This is consistent with occasional cells having sufficient methylation to cause partial loss of allelic expression, and perhaps a proportionate change in risk of neoplastic transformation. Alternatively, low-level methylation may flag the presence of a cryptic change-causing gene dysregulation. Consistent with this observation is the previous finding of complete transcriptional inactivation in an individual with low-level (~10%) constitutional *MLH1* methylation.<sup>10</sup> The finding of methylation segregating with a c.+27G>A variant in the cancer-affected father and son of family N suggests a genetic basis to their low-level methylation. However, this SNV does not cause an amino acid change or the creation of a new splice site, and does not segregate with disease in this pedigree. Together, these observations argue against a major role for this SNV in cancer causation.

Sequence analysis of the entire *MLH1* promoter revealed an additional four sequence variants in five index cases, none of which were associated with constitutional methylation. The luciferase assays demonstrated that the novel c.-11C>T SNV causes a profound constitutive downregulation of transcription, although one of the two novel microdeletions caused a minor reduction in transcriptional activity. Consistent with previous reports, the c.-42C>T SNV was associated with significantly reduced transcriptional output.<sup>15,25</sup> Collectively, these findings suggest that some *MLH1* promoter variants cause transcriptional dysregulation through a primarily DNA alteration-based mechanism; presumably by creating or destroying individual DNA-nucleoprotein binding motifs and thereby altering interactions with particular transcription factors. It is interesting to note that with the exception of one carrier of the c.-11C>T SNV, each of the other probands developed cancer at an older age of onset and did not have a remarkable family history of cancer, suggesting that the mechanism of inactivation of *MLH1* may influence phenotype. Review of the *MLH1* promoter sequences from 304 healthy Australians did not identify any of these sequence changes among them.<sup>15</sup>

This study had notable limitations. To find sufficient CRC cases with loss of expression of *MLH1* not due to germline *MLH1* gene mutations, we included all available cases from the C-CFR. Because of the complex ascertainment schemes of the six centers and the changes in ascertainment over three recruitment phases, we cannot define an exact population frequency for *MLH1* epimutations for the whole population of all CRC patients. The C-CFR cases were over-selected for young onset cases and multiplex families; we did not have constitutional DNA on all relatives with reported Lynch-related cancers, nor did we have tumor phenotype on all cases or relatives.

Although *MLH1* constitutional epimutations are uncommon, their identification has clinical significance for those affected. When related to underlying promoter sequence

changes, there appears to be a family phenotype that is similar to Lynch syndrome, although work is needed to more precisely identify the associated cancer risks. Relatives who do not carry this mutation can be reassured that their cancer risks may not be different from the general population. Those who do carry the mutation warrant very close follow-up and frequent communication as knowledge of this field matures. For those with *MLH1* hemiallelic hypermethylation but no identified underlying DNA sequence change, the clinical and familial risks remain poorly defined. It may be optimal to evaluate relatives for constitutional *MLH1* hypermethylation and to follow any who manifest this phenotype under a research protocol, as if they may have risks akin to those reported in Lynch syndrome. It is not clear if those who do not show *MLH1* methylation in blood are at average risk, whether they may develop methylation as they age, or whether the methylation may be present in other tissues. Collaboration with other centers to generate larger numbers will be necessary to more fully understand the clinical ramifications of *MLH1* epimutations.

Despite their rarity, our findings support consideration of screening for *MLH1* constitutional epimutations and promoter sequence changes in individuals with CRC who have a negative germline sequence screen of the *MLH1* coding region after loss of expression of *MLH1* has been discovered in a tumor. At a more fundamental level, this study shows that assessment of cancer predisposition must now take into account constitutional changes in the epigenome as well as the genome.

#### SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

#### ACKNOWLEDGMENTS

M.P.H. is supported by a CDAll from the Australian National Health and Medical Research Council (NHMRC) and Cancer Institute NSW. This work was funded by the NHMRC and Cancer Council NSW. The Colon Cancer Family Registry is supported by the National Cancer Institute, National Institutes of Health, grant number: RFA #CA-95-011; Australian Colorectal Cancer Family Registry, UO1 CA097735; USC Familial Colorectal Neoplasia Collaborative Group, grant number: UO1 CA074799; Mayo Clinic Cooperative Family Registry for Colon Cancer Studies, grant number: UO1 CA074800; Ontario Registry for Studies of Familial Colorectal Cancer, grant number: UO1 CA074783; Seattle Colorectal Cancer Family Registry, grant number: UO1 CA074794; University of Hawaii Colorectal Cancer Family Registry, grant number: UO1 CA074806; University of California, Irvine Informatics Center, grant number: UO1 CA078296. Acknowledgment of contributors who did not meet the requirements for authorship: We thank the Colon CFR study coordinators and data managers who helped prepare the data set for these analyses (Terrilea Burnett, Helen Chen, Charisma Sanders, Darshana Daftary, Douglas Snazel, Allyson Templeton, Christine Davies, and Paul Mosquin), as well as the participants in the Colon CFR who have generously donated their time for this project.

The content of this article does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the CFRs, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government or the CFR.

#### DISCLOSURE

The authors declare no conflict of interest.

#### REFERENCES

- Peltomäki P, Vasen H. Mutations associated with HNPCC predisposition – Update of ICG-HNPCC/INSIGHT mutation database. *Dis Markers* 2004;20:269–276.
- Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, Boland CR. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin Genet* 2009;76:1–18.
- 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;62:3925–3928.
- Miyakura Y, Sugano K, Akasu T, et al. Extensive but hemiallelic methylation of the h*MLH1* promoter region in early-onset sporadic colon cancers with microsatellite instability. *Clin Gastroenterol Hepatol* 2004;2:147–156.
- Suter CM, Martin DI, Ward RL. Germline epimutation of *MLH1* in individuals with multiple cancers. *Nat Genet* 2004;36:497–501.
- Chan TL, Yuen ST, Kong CK, et al. Heritable germline epimutation of *MSH2* in a family with hereditary nonpolyposis colorectal cancer. *Nat Genet* 2006;38:1178–1183.
- Ligtenberg MJ, Kuiper RP, Chan TL, et al. Heritable somatic methylation and inactivation of *MSH2* in families with Lynch syndrome due to deletion of the 3' exons of *TACSTD1*. *Nat Genet* 2009;41:112–117.
- Hitchins M, Williams R, Cheong K, et al. *MLH1* germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. *Gastroenterology* 2005;129:1392–1399.
- Hitchins MP, Ward RL. Constitutional (germline) *MLH1* epimutation as an aetiological mechanism for hereditary non-polyposis colorectal cancer. *J Med Genet* 2009;46:793–802.
- Goel A, Nguyen TP, Leung HC, et al. 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;128:869–878.
- Hitchins MP, Owens SE, Kwok CT, Godsmark G, Algar UF, Ramesar RS. Identification of new cases of early-onset colorectal cancer with an *MLH1* epimutation in an ethnically diverse South African cohort. *Clin Genet* 2011;80:428–434.
- Morak M, Schackert HK, Rahner N, et al. 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;16:804–811.
- Hitchins MP, Ward RL. Erasure of *MLH1* methylation in spermatozoa-implications for epigenetic inheritance. *Nat Genet* 2007;39:1289.
- Hitchins MP, Wong JJ, Suthers G, et al. Inheritance of a cancer-associated *MLH1* germ-line epimutation. *N Engl J Med* 2007;356:697–705.
- Hitchins MP, Rapkins RW, Kwok CT, et al. 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;20:200–213.
- Morak M, Koehler U, Schackert HK, et al. Biallelic *MLH1* SNP cDNA expression or constitutional promoter methylation can hide genomic rearrangements causing Lynch syndrome. *J Med Genet* 2011;48:513–519.
- Gylling A, Ridanpää M, Vierimaa O, et al. Large genomic rearrangements and germline epimutations in Lynch syndrome. *Int J Cancer* 2009;124:2333–2340.
- Newcomb PA, Baron J, Cotterchio M, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. *Cancer Epidemiol Biomarkers Prev* 2007;16:2331–2343.
- Kwok CT, Ward RL, Hawkins NJ, Hitchins MP. Detection of allelic imbalance in *MLH1* expression by pyrosequencing serves as a tool for the identification of germline defects in Lynch syndrome. *Fam Cancer* 2010;9:345–356.
- Eads CA, Danenberg KD, Kawakami K, et al. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res* 2000;28:E32.

21. Wong JJ, Hawkins NJ, Ward RL, Hitchins MP. Methylation of the 3p22 region encompassing *MLH1* is representative of the CpG island methylator phenotype in colorectal cancer. *Mod Pathol* 2011;24:396–411.
22. Crépin M, Dieu MC, Lejeune S, et al. Evidence of constitutional *MLH1* epimutation associated to transgenerational inheritance of cancer susceptibility. *Hum Mutat* 2012;33:180–188.
23. Raevaara TE, Korhonen MK, Lohi H, et al. Functional significance and clinical phenotype of nontruncating mismatch repair variants of *MLH1*. *Gastroenterology* 2005;129:537–549.
24. Müller-Koch Y, Kopp R, Lohse P, et al. 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;6:473–482.
25. Green RC, Green AG, Simms M, Pater A, Robb JD, Green JS. Germline *hMLH1* promoter mutation in a Newfoundland HNPCC kindred. *Clin Genet* 2003;64:220–227.
26. Isidro G, Matos S, Gonçalves V, et al. Novel *MLH1* mutations and a novel *MSH2* polymorphism identified by SSCP and DHPLC in Portuguese HNPCC families. *Hum Mutat* 2003;22:419–420.



**This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>**