Cancer Cell
Article

Dominantly Inherited Constitutional Epigenetic Silencing of *MLH1* in a Cancer-Affected Family Is Linked to a Single Nucleotide Variant within the 5′UTR

Megan P. Hitchins,[1] Robert W. Rapkins,[1] Chau-To Kwok,[1] Sameer Srivastava,[1] Justin J.L. Wong,[1] Levon M. Khachigian,[2] Patsie Polly,[3] Jack Goldblatt,[4] and Robyn L. Ward[1,*]
[1]Adult Cancer Program, Lowy Cancer Research Centre, Prince of Wales Clinical School
[2]Centre for Vascular Research, Lowy Cancer Research Centre
[3]Inflammation and Infection Research Centre and Department of Pathology, School of Medical Sciences
Faculty of Medicine at the University of New South Wales, NSW 2052, Australia
[4]Genetic Services of Western Australia and School of Paediatrics and Child Health, University of Western Australia, Perth, WA 6008, Australia
*Correspondence: robyn@unsw.edu.au
DOI 10.1016/j.ccr.2011.07.003

SUMMARY

Constitutional epimutations of tumor suppressor genes manifest as promoter methylation and transcriptional silencing of a single allele in normal somatic tissues, thereby predisposing to cancer. Constitutional *MLH1* epimutations occur in individuals with young-onset cancer and demonstrate non-Mendelian inheritance through their reversal in the germline. We report a cancer-affected family showing dominant transmission of soma-wide highly mosaic *MLH1* methylation and transcriptional repression linked to a particular genetic haplotype. The epimutation was erased in spermatozoa but reinstated in the somatic cells of the next generation. The affected haplotype harbored two single nucleotide substitutions in tandem; c.-27C > A located near the transcription initiation site and c.85G > T. The c.-27C > A variant significantly reduced transcriptional activity in reporter assays and is the probable cause of this epimutation.

INTRODUCTION

Altered DNA methylation is frequently observed in cancer cells and includes both hypomethylation associated with activation of oncogenes (Feinberg and Vogelstein, 1983), and hypermethylation of the CpG island promoters of tumor suppressor genes associated with transcriptional silencing (Greger et al., 1989; Herman et al., 1998). More recently, constitutional epimutations of DNA mismatch repair and tumor suppressor genes, including *MLH1*, *MSH2*, *DAPK1*, and *KILLIN*, have been identified in patients who do not carry a sequence mutation within the cognate gene, yet have early-onset cancer consistent with a familial cancer syndrome (Bennett et al., 2010; Chan et al., 2006; Gazzoli et al., 2002; Raval et al., 2007). These so-called constitutional epimutations, characterized by monoallelic promoter methylation and transcriptional silencing of the affected allele within normal somatic tissues, represent another mechanism for cancer predisposition.

Lynch syndrome is an autosomal dominant cancer susceptibility syndrome characterized by the development of colorectal, endometrial and other cancers demonstrating microsatellite instability (MSI) as a consequence of defective mismatch repair, usually at a young age (Lynch and Lynch, 2004). The syndrome is typically caused by heterozygous germline mutations within the coding region of the mismatch repair genes (Peltomäki and Vasen, 2004). A proportion of cases with suspected Lynch syndrome, in whom no germline mismatch repair sequence mutation is found, carry a constitutional epimutation of *MLH1*

Significance

Constitutional epimutations of cancer-related genes represent an alternative mechanism for cancer predisposition. Some are caused by underlying *cis*-acting genetic alterations proximate to the affected gene and display autosomal dominant inheritance. However, the mechanistic basis for *MLH1* epimutations that are reversible between generations remains undefined. This study of a cancer-affected family with dominant inheritance of a mosaic *MLH1* epimutation through three generations provides evidence of a *cis*-genetic basis for constitutional epigenetic silencing of *MLH1*. The identification of a 5′UTR germline c.-27C > A variant raises the interesting possibility that promoter variants of unknown pathogenic significance within *MLH1* and other genes may confer cancer susceptibility through their association with epigenetic modifications. This finding indicates a close interaction between genotype and epigenotype in a cancer-associated gene.

(Gazzoli et al., 2002; Hitchins et al., 2005; Miyakura et al., 2004; Suter et al., 2004) or *MSH2* (Chan et al., 2006; Ligtenberg et al., 2009). Individuals with an *MSH2* epimutation exhibit tissue-specific *MSH2* promoter methylation, most apparent in epithelial tissues. In these cases, somatic methylation is caused by a *cis*-acting germline deletion encompassing the transcription termination signal of the upstream *EPCAM* gene, which results in elongation of transcription from *EPCAM* into *MSH2* due to failed transcription termination (Ligtenberg et al., 2009). Predictably, families with *MSH2* epimutations display autosomal dominant inheritance. In contrast, constitutional *MLH1* epimutations are reported in individuals with a significant personal but not familial history of cancer. Clear evidence of intergenerational transmission of *MLH1* epimutation has been reported in just two families (Hitchins et al., 2007; Morak et al., 2008), one of which showed non-Mendelian inheritance. In this family, the affected mother transmitted her *MLH1* epimutation to just one of three sons, each of whom had inherited the identical maternal allele, whereas the allele reverted to the normal functional state in her other two sons (Hitchins et al., 2007). Moreover, *MLH1* epimutations are erased in the spermatozoa of male carriers, as evidenced by complete demethylation and transcriptional reactivation of the somatically-affected allele (Hitchins and Ward, 2007; Hitchins et al., 2007).

The mechanistic basis of constitutional *MLH1* epimutation is unknown. Certainly no primary genetic defect has been identified in the vicinity of the affected locus in carriers. Families with constitutional *MLH1* and other epimutations offer an opportunity to understand how epigenetic aberrations arise in humans. Understanding the pattern of inheritance and etiology of constitutional epimutations will inform genetic counseling and clinical management. The goal of this study was to investigate the mechanistic basis of *MLH1* epimutations and to define the pattern of inheritance associated with them.

RESULTS

Haplotype-Specific Soma-Wide Mosaic *MLH1* Epimutation in a Proband with Young-Onset Colorectal Cancer

A previous population-based study of subjects from Western Australia who had developed early-onset (<60 years) MSI colorectal cancer identified 17 cases with a MLH1-deficient tumor but no pathogenic germline mismatch repair mutation (Schofield et al., 2009). To determine if these individuals carried a constitutional *MLH1* epimutation, we screened their peripheral blood lymphocyte (PBL) DNA for the presence of *MLH1* promoter methylation by quantitative CpG methylation pyrosequencing (Figure 1A). This led to the identification of one proband, a Caucasian female diagnosed with a right-sided colorectal cancer at age 41 years, who had variable levels of *MLH1* methylation throughout her normal somatic tissues (Figure 1B). This was confirmed by clonal bisulphite sequencing across the critical region of the *MLH1* promoter (see Figure S1 available online). Sequencing of the entire CpG island encompassing the *MLH1* promoter (2 kb upstream to 1 kb downstream of the first codon) in the proband's genomic DNA identified two heterozygous single nucleotide substitutions within exon 1; c.-27C > A in the 5′UTR and missense variant c.85G > T (rs63750656). The two genetic alleles were separated by cloning, and resequencing revealed a variant haplotype on which both the c.-27A and c.85T variants were linked in tandem (Figure 1C). These two sites were incorporated into assays designed to distinguish the epigenetic modifications and transcriptional activity associated with each genetic allele (Figure 2A). Allelic bisulphite sequencing demonstrated that the methylation occurred specifically on the variant haplotype and involved CpG sites flanking both nucleotide substitutions (c.-27A and c.85T) (Figure 2B). Moreover, bisulphite sequencing revealed a mixture of fully, partially and unmethylated forms of the variant haplotype, indicative of epigenetic mosaicism (Figure 2B). Chromatin immunoprecipitation followed by allele-specific real-time PCR (ChIP-AS) in cultured lymphoblastoid cells (LCLs) from the proband showed a coalescence of repressive histone modifications including H3-K9 and H3-K27 trimethylation, and depletion of the H3-K9 acetylation and H3-K4 trimethylation motifs permissive to transcription, on the variant haplotype (Figure 2C). Quantitative allelic expression analyses showed transcriptional repression of the variant haplotype in different tissues, ranging from complete to partial loss of expression, consistent with mosaic epigenetic silencing (Figure 2D). This variable allelic expression imbalance was confirmed at the benign c.655A > G SNP within *MLH1* exon 8, at which the proband was also informative (Figure S2A). An acquired loss-of-heterozygosity (LOH) of the wild-type allele was detected in the tumor by allele quantification, accounting for its loss of MLH1 and MSI (Figure 2E; Figure S2B). In summary, the proband bore a mosaic soma-wide *MLH1* epimutation with variable transcriptional suppression on a variant genetic haplotype, marked by c.-27C > A and c.85G > T substitutions, as the probable cause of her cancer susceptibility.

Dominant Intergenerational Transmission of the *MLH1* Epimutation Linked to the Variant Haplotype

The proband reported a family history of Lynch syndrome-associated cancers. Specifically her brother had a colorectal cancer at 46 years of age that also failed to express MLH1 and was MSI, while her mother developed endometrial and metachronous colorectal cancers at a young age. Twelve members from three generations of the family, designated family 16, consented to join our study (Figure 3A). Each relative was tested for *MLH1* methylation by CpG pyrosequencing and genotyped at multiple polymorphic markers across a 4.3 Mb region encompassing the *MLH1* locus. Strikingly, the affected mother (I1), affected brother (II1), and five currently asymptomatic relatives (II2, II3, II4, III1, III2) spanning all three generations also had soma-wide *MLH1* methylation. The methylation segregated faithfully with the variant genetic haplotype (marked by the c.-27C > A and c.85G > T substitutions) and followed an autosomal dominant inheritance pattern (Figure 3A). Furthermore, allelic methylation and expression studies demonstrated at a molecular level that the promoter methylation and transcriptional repression were linked to the variant haplotype in each of these family members (Figures 3 and 4; Figures S3A–S3E). In contrast, no methylation was detected in relatives who did not carry this variant haplotype (III3, III4, III5), and they each demonstrated balanced biallelic expression of *MLH1* at the common c.655A > G SNP (Figure 3; Figures S3F–S3H). The dominant inheritance of the *MLH1* epimutation on the variant

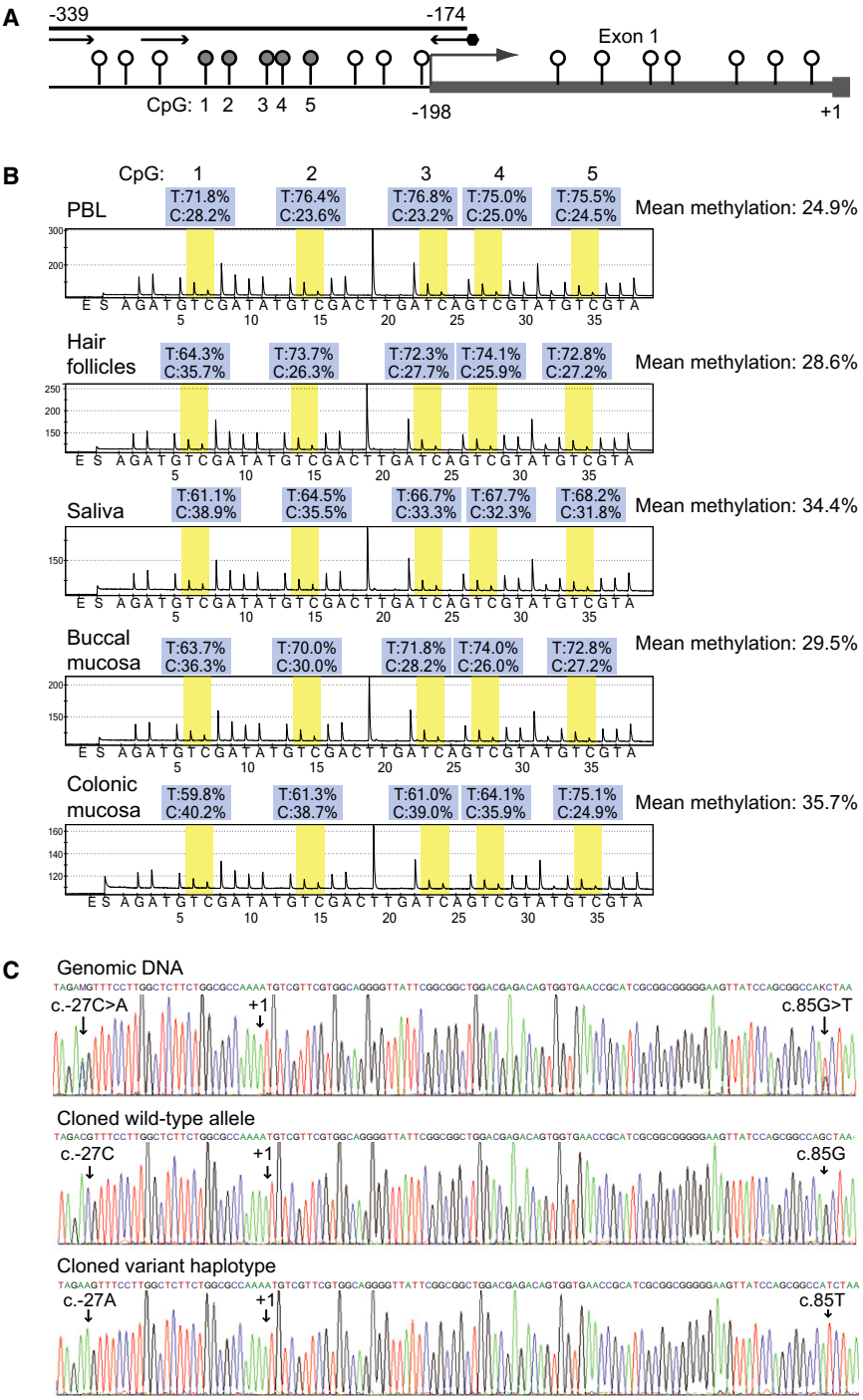


Figure 1. Soma-Wide Methylation of the *MLH1* Promoter in the Female Proband

(A) Map of CpG methylation pyrosequencing assay used to measure *MLH1* promoter methylation levels. Gray rectangle; exon with transcription start site according to reference sequence (GenBank accession NM_000249) indicated by an arrow. +1; A of ATG start codon. Circles: CpG dinucleotides with the five sites assayed numbered and filled in gray. Black bar: region amplified by outer primers (horizontal arrows). Hexagon: biotinylated primer. An inner primer was used for pyrosequencing of the biotinylated strand.

(B) Representative pyrograms show the level of *MLH1* promoter methylation detected in various normal somatic tissues. Five CpGs interrogated are shaded yellow and the percentage methylation (in blue box) at each site is calculated as the C:T ratio of peak heights (representing methylated:unmethylated cytosine).

(C) Sequence electropherogram from genomic DNA showing heterozygosity for the c.-27C > A and c.85G > T variants (top). Beneath, plasmid inserts containing a cloned amplicon show a wild-type allele, and a variant haplotype bearing both c.-27A and c.85T variants in *cis*.

See also Figure S1.

haplotype in family 16 implicates an underlying *cis*-acting genetic defect that predisposes this particular haplotype to epigenetic modification and silencing.

Somatic Allelic Mosaicism

Epigenetic mosaicism was observed as inter-tissue and inter-individual differences in the extent of *MLH1* promoter methylation and of transcriptional silencing of the variant haplotype in family 16 (Figures 3B and 3C). Comparison of overall methylation levels from four different somatic tissues (PBL, buccal mucosa, saliva, and hair follicles), from each of seven carriers of the variant haplotype (including three cancer-affected and four unaffected family members) failed to show a correlation between the degree of methylation and any particular tissue source or disease status. However, in the two affected siblings with colonic mucosa available (proband II5 and brother II1), methylation levels were highest in this tissue (Figure 3B). The inter-individual variability in overall methylation levels were best illustrated by two asymptomatic relatives, a 22-year-old female (III2) who had the highest methylation levels (range, 27.3%–38.5%) and her 49-year-old uncle (II3) who had the lowest levels (range, 18.7%–26.0%). Allelic bisulphite sequencing revealed a heterogeneous pattern of hypermethylated, partially methylated, and unmethylated copies of the variant *MLH1* haplotype among the somatic tissues of the epimutation carriers (Figures 2 and 4; Figures S3A–S3F). Further illustrating this epigenetic mosaicism, the comparative level of expression of the variant haplotype in those tissues ranged from 0%–60% of the wild-type allele (Figure 3C; Figure S3). Interestingly, the degree of epigenetic silencing also demonstrated a temporal difference within the same individual. In two PBL samples collected from the affected brother (II1) at time intervals 2 months

apart, the levels of transcripts derived from the variant haplotype differed substantially (3.4% and 26.4% of total *MLH1* transcripts), yet the methylation levels were similar in both samples (26.0% and 27.1%) (Figures 3B,3C and 4; Figure S4). Although transcripts derived from the variant *MLH1* haplotype are presumably generated from those chromosomes that have escaped promoter methylation, the discordance between the methylation and transcription levels in these two PBL samples suggest that some unmethylated copies may also be silenced. Thus, histone modifications or other nuclear factors may also induce transcriptional silencing of the variant haplotype in the absence of CpG methylation.

Erasure of the Epigenetic Aberration in the Male Germline

Allelic methylation analyses and haplotyping demonstrated that the *MLH1* epimutation was transmitted faithfully with the variant genetic haplotype from the affected brother (II1) to his son (III1) (Figure 4; Figure S4). To determine if the epimutation was conveyed through the germline to the next generation with its epigenetic modifications intact via "gametic epigenetic inheritance," motile spermatozoa from the affected brother (II1) were isolated by the swim-up procedure. Methylation analyses using various techniques showed the spermatozoa cells were unmethylated at the *MLH1* promoter (Figures 5A–5C). The negligible amount of *MLH1* methylation detected by the most sensitive of these assays could be attributed to contamination of the sample with residual somatically-derived genomic DNA, because this was matched by traces of methylation of the imprinted *SNRPN* control gene (Figure 5A). Furthermore, the levels of expression from both *MLH1* alleles were equivalent in spermatozoa, indicating restoration of transcriptional activity from the naked variant haplotype to normal levels (Figure 5D). This indicates that the altered epigenetic state associated with the somatically-affected variant haplotype was completely erased in the male germline, and that the epigenetic aberration was reestablished once more in the soma of the next generation. Furthermore, this showed the variant haplotype has the capacity for normal transcriptional activity in the absence of repressive epigenetic modifications.

Transcriptional Activity of the Variant Haplotype Is Dependent on the Epigenetic Status of the CpG Island Overlapping *MLH1* and *EPM2AIP1*

The *MLH1* and *EPM2AIP1* genes share the same bi-directional CpG island promoter, are transcribed head-to-head from opposite strands (Figure 6A), and are epigenetically coregulated (Goel et al., 2011; Lin et al., 2007). To determine if the epimutation was confined to this CpG island, or included flanking genes *TRANK1* and *LRRFIP2* (Figure S5A), the methylation status and allelic activity of these two genes was investigated. Neither gene was methylated in family 16 (Figure S5B). Several carriers of the variant haplotype were heterozygous for common SNPs within the 3′UTR of *TRANK1* and *LRRFIP2* (Figure 3A). Using these SNPs for allelic expression studies, we showed that both genes were biallelically transcribed in LCLs from these individuals (Figure S5C), indicating the epimutation was localized to the *MLH1-EPM2AIP1* CpG island.

Although the *MLH1* reference mRNA begins at nucleotide c.-198 (Figure 2A), distinct mRNA isoforms, namely 1a and 1b, have been isolated from a colorectal carcinoma cell line (Lin et al., 2007). To identify the major *MLH1* transcripts and their initiation sites in normal somatic cells, 5′RACE was performed on PBL mRNA from a healthy individual. We also identified two prevalent *MLH1* isoforms that utilize alternative exons 1a or 1b (Figure 6A). Interestingly, the dominant transcription initiation site of the 1a isoform was located at c.-29, just two bases upstream of the c.-27C > A variant. Because the 1a transcript encompasses both the c.-27C > A and c.85G > T variants identified in family 16 (Figure 6A), we performed allelic representation analyses on the *MLH1*-1a and 1b isoforms individually, as well as on *EPM2AIP1*, to determine if somatic transcriptional silencing of the variant haplotype was confined to the *MLH1*-1a isoform, or if all transcripts derived from this CpG island were equally affected. Isoform-specific amplification was performed between *MLH1* exons 1a or 1b and exon 9 to encompass the c.655A > G SNP within exon 8 in LCLs from members of family 16 who are heterozygous for this SNP (proband II5, II3 and III2). Allele quantification was then performed at the c.655A > G site. Both *MLH1* isoforms demonstrated a similar extent of allelic expression imbalance (Figure 6B). Allelic expression analyses of *EPM2AIP1* in family members who were informative for an expressible 3′UTR SNP also revealed a significant allelic expression imbalance in carriers of the variant haplotype (II5, III2), whereas balanced *EPM2AIP1* expression was observed in a noncarrier (III3) (Figure 6C). The equivalent reduction in allelic expression observed for each of the three major transcripts, irrespective of the precise location of their initiation sites, suggests the transcriptional activity of the variant haplotype correlates with the epigenetic state of the shared *EPM2AIP1-MLH1* CpG island.

To further investigate the relationship between epigenetic modifications and expression of the variant haplotype in somatic cells, LCLs from family 16 were cultured with a combination of the demethylating agent 5-aza-2′-deoxycytidine (Aza) and the histone deacetylase inhibitor trichostatin A (TSA). A significant upregulation of the variant *MLH1* haplotype was observed in the LCLs from carriers following combinatorial drug treatment, concomitant with a reduction in CpG methylation levels (Figure 6D) and a partial reversal of histone modifications (Figure S5D). In contrast, no change in the allelic expression ratio was detected at the c.655A > G SNP in another family member (III3) who was unmethylated at *MLH1* and did not inherit the variant haplotype (Figure 6D). Similarly, a partial upregulation of the coaffected *EMP2AIP1* allele was observed in LCLs from carriers of the variant haplotype following treatment with AZA and TSA (Figure 6C). These findings confirm that the transcriptional activity of the variant haplotype is indeed epigenetically regulated. This suggests that a fundamental genetic defect located on the variant haplotype in family 16 induces epigenetic remodeling in somatic cells, which in turn governs the transcriptional activity of the *MLH1-EPM2AIP1* locus. However, this does not rule out the involvement of additional nuclear interactions that may also mediate transcriptional suppression of unmethylated copies of the variant haplotype.

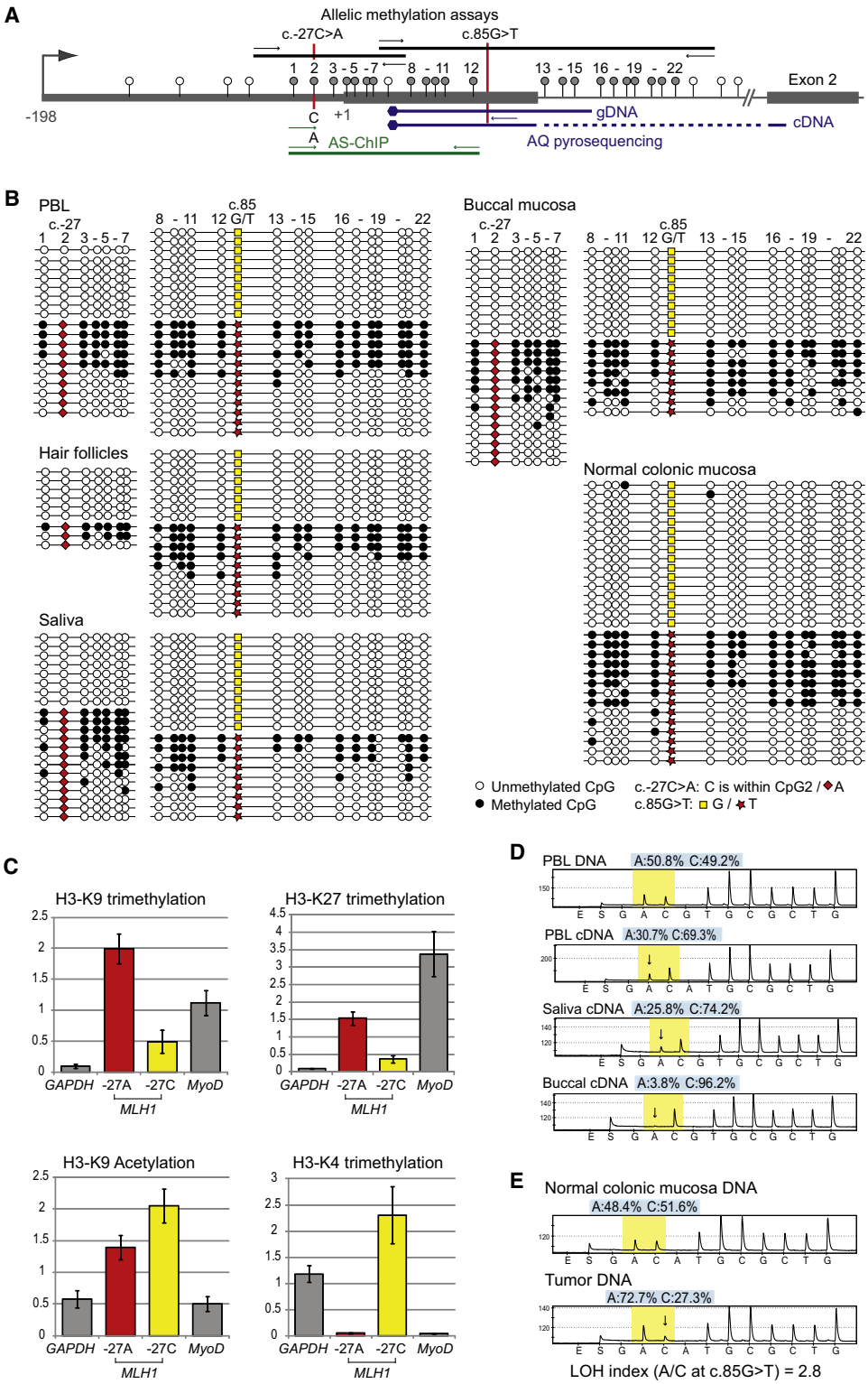
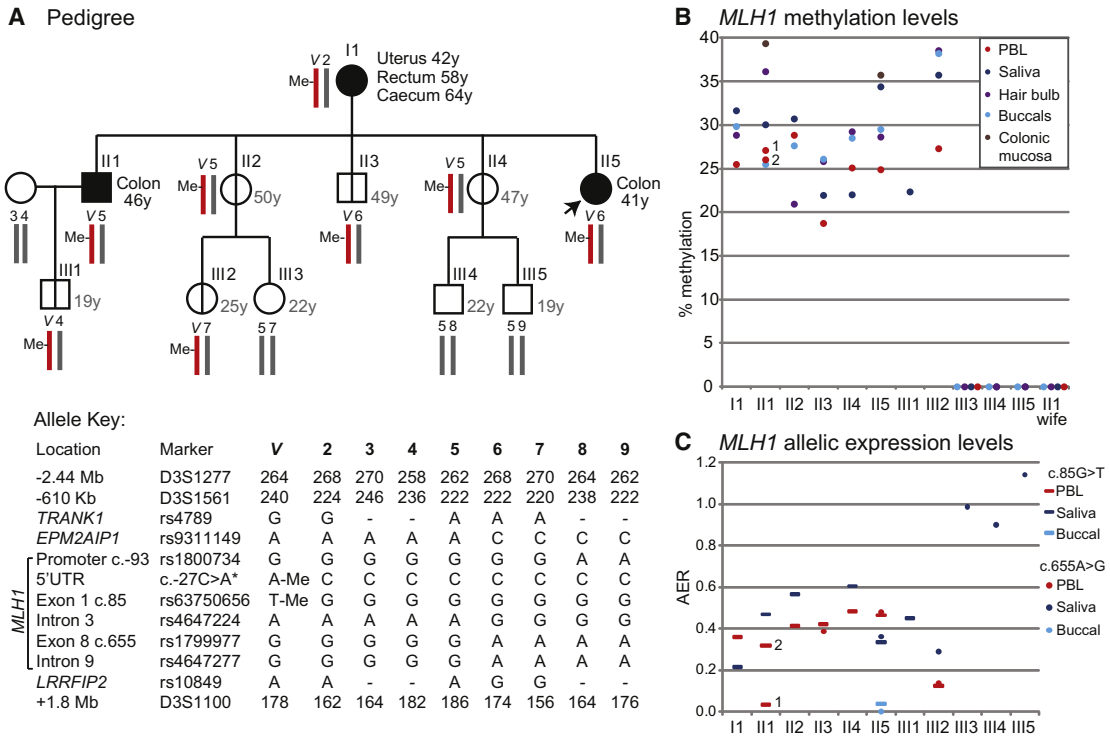


Figure 2. Soma-Wide Epigenetic Repression of a Single Allele of *MLH1* in the Proband

(A) Map of *MLH1* assays used to assess allelic methylation status (black bars) across the c.-27C > A and c.85G > T sites, allele quantification (AQ) of cDNA or genomic DNA templates (blue bars) at c.85G > T, and AS-ChIP (allele-specific real-time PCR following chromatin immunoprecipitation) to profile histone modifications at the c.-27C > A site (green). Exons in gray show 5′UTR (narrow box) and coding sequence (wide boxes), with transcription initiation site indicated by a gray arrow. Lollipops; CpG sites, with those analyzed for methylation by clonal bisulphite sequencing filled in gray. Horizontal arrows show primer positions.



Location	Marker	*V*	2	3	4	5	6	7	8	9
-2.44 Mb	D3S1277	264	268	270	258	262	268	270	264	262
-610 Kb	D3S1561	240	224	246	236	222	222	220	238	222
TRANK1	rs4789	G	G	-	-	A	A	A	-	-
EPM2AIP1	rs9311149	A	A	A	A	A	C	C	C	C
MLH1 Promoter c.-93	rs1800734	G	G	G	G	G	G	G	A	A
MLH1 5′UTR	c.-27C>A*	A-Me	C	C	C	C	C	C	C	C
MLH1 Exon 1 c.85	rs63750656	T-Me	G	G	G	G	G	G	G	G
MLH1 Intron 3	rs4647224	A	A	A	A	A	G	G	G	G
MLH1 Exon 8 c.655	rs1799977	G	G	G	G	G	A	A	A	A
MLH1 Intron 9	rs4647277	G	G	G	G	G	A	A	A	A
LRRFIP2	rs10849	A	A	-	-	A	G	G	-	-
+1.8 Mb	D3S1100	178	162	164	182	186	174	156	164	176

Figure 3. Pedigree of Family 16 Showing Constitutional *MLH1* Methylation Segregates with the Variant Genetic Haplotype

(A) Pedigree. Age of cancer diagnosis (black font) and current age of asymptomatic individuals (gray font) is shown. Circles, female; squares, male; filled, affected methylation carrier; vertical line at center, asymptomatic methylation carrier. Chromosome 3p haplotypes, generated by compiling the genotypes from multiple SNP and microsatellite markers located within or flanking *MLH1*, are represented as vertical lines and numbered according to the key beneath. Soma-wide *MLH1* methylation (Me) segregated faithfully with the variant haplotype, *V* (in red), marked by the c.-27A and c.85T variants. Other unmethylated genetic alleles are shown in gray.

(B) Levels of *MLH1* methylation as measured in duplicate by CpG pyrosequencing in somatic tissues of family members.

(C) Relative levels of allelic expression of *MLH1* in somatic tissues. Allelic expression ratios (AER) obtained at two expressible *MLH1* SNPs in informative family members by AQ pyrosequencing of cDNA normalized to genomic DNA from the same tissue. AER of 1.0 indicates equal levels of expression from both alleles. AER < < 1.0 indicates reduced expression from the variant haplotype.

See also Figure S3.

The c.-27C > A or c.85G > T Variants Are Implicated as the Causative Genetic Defect

To search for a germline structural sequence alteration of the variant haplotype in family 16, the Affymetrix Genome-wide Human SNP 6.0 array was employed for genotyping and copy number variant (CNV) analyses of five family members from all three generations (I1, II2, II4, II5, III2). Three algorithms (QuantiSNP, PennCNV, COKGEN) were used for CNV estimation against a reference baseline generated from 300 healthy Caucasian controls from the Wellcome Trust Case Control Consortium (Colella et al., 2007). No evidence for any CNV within 30Mb of *MLH1* was found in family 16 (data not shown). Furthermore, the finding that neighboring genes *TRANK1* and *LRRFIP2* were structurally and functionally intact argues against a cryptic genetic defect within either gene that could influence the epigenetic state of the intervening *MLH1-EPM2AIP1* locus. This suggests the *cis*-genetic cause of the epimutation is located within the *MLH1-EPM2AIP1* promoter, prompting further

For AQ pyrosequencing, the biotinylated sense strand is indicated by a hexagon. Blue arrow shows internal pyrosequencing primer. For AS-ChIP, two forward primers ending in C or A were used with a common reverse primer to specifically amplify the c.-27C or c.-27A alleles, respectively.

(B) Allelic methylation patterns in normal somatic tissues. Clonal bisulphite sequencing shows methylation specifically affects the (left) A allele at c.-27C > A and the (right) T allele at c.85G > T. Each horizontal line represents a single DNA strand and circles individual CpG sites.

(C) Allele-specific histone modifications in LCLs by AS-ChIP with antibodies targeting specific histone tail moieties at the c.-27C > A site. *GAPDH* and *MyoD* provide controls for transcriptionally active and repressed genes, respectively. Mean relative enrichment of gene or allele-specific DNA performed in triplicate ±1 standard deviation (SD) is shown.

(D) Allelic imbalance in *MLH1* expression. Representative pyrograms showing AQ at the c.85G > T site on the complementary strand. Relative peak heights of the C-G and A-T alleles in the yellow shaded areas are provided above in blue. Variable transcriptional repression of the T allele was detected in cDNAs from normal somatic tissues (indicated by downward arrows).

(E) Loss of heterozygosity (LOH) of the wild-type c.85G allele in colorectal tumor DNA is detected compared to normal tissue by AQ pyrosequencing.

See also Figure S2.

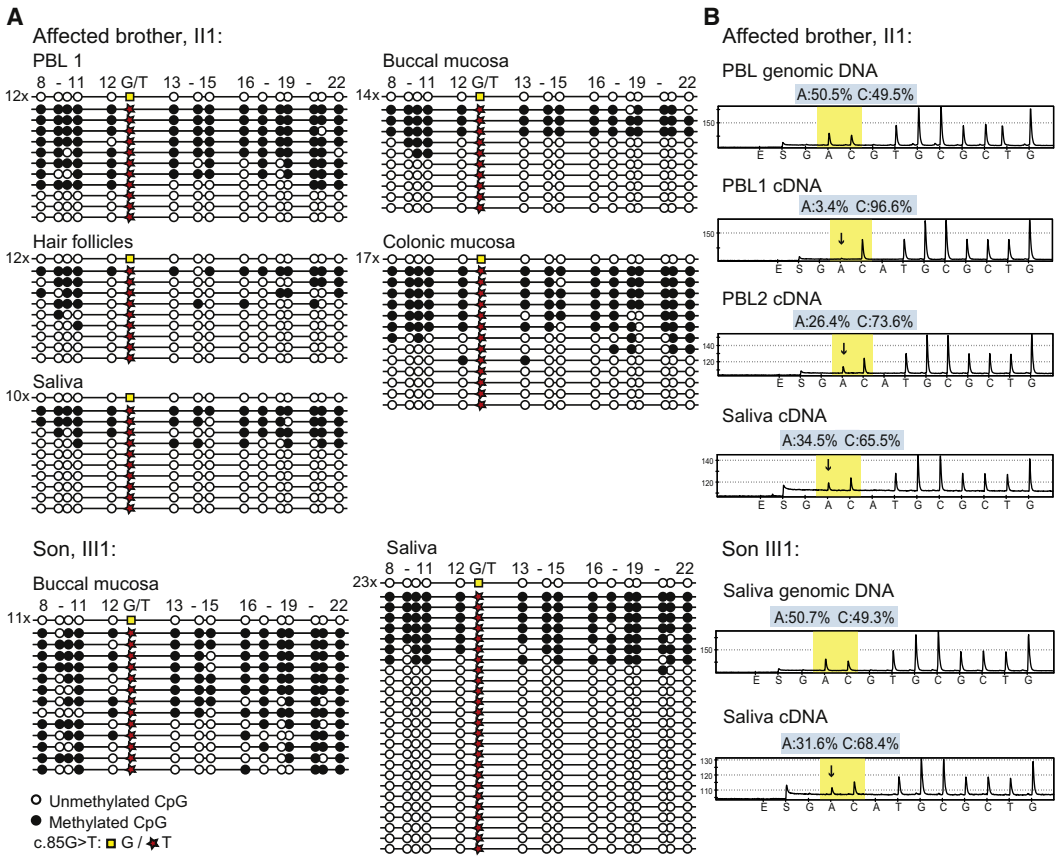


Figure 4. Dominant Transmission of Somatic Epigenetic Silencing of *MLH1* Linked to the Variant Haplotype from the Affected Brother II1 to his Son III1

(A) Allelic *MLH1* methylation patterns in somatic tissues from the affected brother II1 and his son III1 by clonal bisulphite sequencing, show promoter methylation specifically affects the T allele at the c.85G > T SNP in father and son. One representative unmethylated G allele is shown, with the number of replicates of the same pattern given on the left.

(B) Allelic imbalance in *MLH1* expression. Illustrative examples of pyrograms from AQ pyrosequencing at the c.85G > T site on the complementary strand in genomic DNA and cDNA templates are shown. PBL1 and 2 refer to two separate PBL samples collected from II1 2 months apart. cDNA synthesis and pyrosequencing were performed in the same batch, hence the difference in allelic expression imbalance observed between these two samples cannot be attributed to technical vagary.

See also Figure S4.

investigation of the germline c.-27C > A and c.85G > T variants. Neither variant was found in our screen of 303 healthy controls.

The c.-27C > A Variant Alone Diminishes Transcriptional Activity in Functional Assays

Luciferase reporter assays were performed to determine if the variant haplotype comprising the c.-27C > A and c.85G > T variants in tandem reduce transcriptional activity in somatic cells using constructs containing the wild-type or naked variant haplotype inserted into the promoterless pGL3-Basic luciferase reporter vector. A 2.5-fold reduction in normalized luciferase activity was observed from the variant haplotype compared to the wild-type allele in transiently transfected HEK293 cells (Figure 7A). To determine if either the c.-27C > A or c.85G > T substitution in isolation downregulates transcriptional activity, or if the two variants have an additive effect when located in *cis*, the two variants were isolated in separate constructs and assayed individually. The isolated c.-27C > A variant diminished promoter activity to a similar extent as the combined variant haplotype, whereas the c.85G > T variant produced transcriptional output equivalent to the wild-type sequence (Figure 7A). Additional promoter reporter assays were performed whereby the constructs were stably integrated into the genome of undifferentiated NCCIT embryonal carcinoma cells. NCCIT transfectants containing the combined variant haplotype and the c.-27C > A substitution alone showed over 8-fold reduction in luciferase output compared to the wild-type and c.85G > T sequences (Figure 7A). This substantial decrease in transcriptional activity occurred independently of overt epigenetic change, because no differential methylation or histone modification was observed between the four transgenic *MLH1* constructs (data not shown).

To confirm the c.-27C > A variant solely downregulates *MLH1* promoter activity in the context of the full *MLH1* promoter sequence, additional assays were performed using constructs containing the entire promoter and 5′UTR. A transcriptional reduction of nearly 3-fold in NCCIT cells (Figure 7A) and

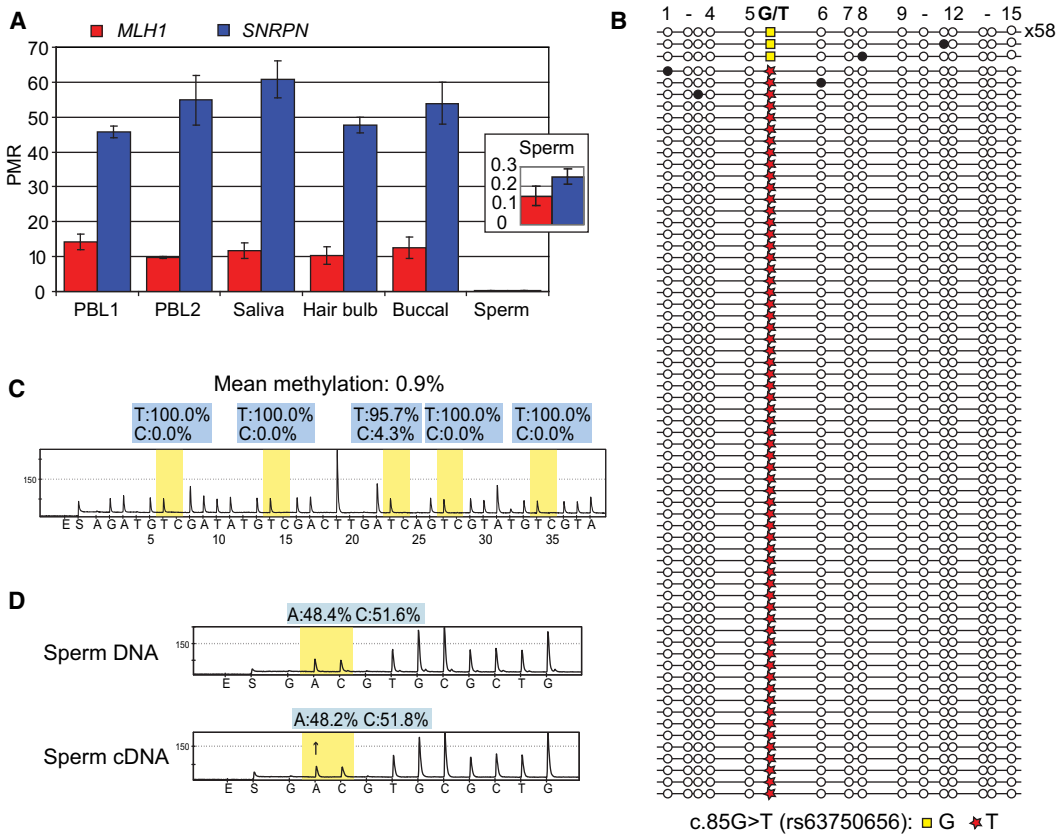


Figure 5. Erasure of *MLH1* Epigenetic Aberration from the Variant Haplotype in the Spermatozoa of Affected Brother II1

(A) Histogram showing levels of hypermethylated *MLH1* and *SNRPN* alleles in somatic tissues and motile spermatozoa from II1, as measured by quantitative real-time methylation-specific PCR. The mean percentage of methylation reference (PMR) from reactions performed in triplicate is plotted ± SD. The imprinted *SNRPN* is methylated specifically on the maternal allele in somatic cells and detected at ~50%, but is unmethylated in spermatozoa due to gametic epigenetic reprogramming. *SNRPN* methylation thus serves as a control for contamination of spermatozoa with somatically-derived DNA. Inset, expanded scale showing negligible levels of *MLH1* methylation in spermatozoa are matched by *SNRPN*, attributable to marginal somatic contamination of the sample.

(B) Clonal bisulphite sequencing of spermatozoa at the c.85G > T site shows erasure of methylation from each T allele.

(C) CpG pyrosequencing confirms erasure of methylation from the *MLH1* promoter. The mean methylation score across five CpG sites interrogated is given above the pyrogram.

(D) AQ pyrosequencing at the c.85G > T site shows an equal level of expression from both *MLH1* alleles in spermatozoa cDNAs.

1.5-fold in HEK293 cells (Figure 7B) was observed for the c.-27C > A variant compared to the wild-type. These functional assays provide supportive evidence that the c.-27C > A substitution singularly predisposes the variant haplotype in family 16 to transcriptional repression, whereas c.85G > T likely represents a rare SNP with no detrimental effect on gene function, but which merely lies in linkage disequilibrium with the true pathogenic change.

Investigation of Heat-Shock Factor as a Nuclear Intermediary in Transcriptional Silencing of the c.-27C > A Variant

An in silico search of databases containing known transcription factor binding motifs predicted the c.-27C > A variant created an inverted dyad repeat (AGAAGTTTCC) of the consensus nGAAn logo for the heat-shock transcription factor (HSF) family (Figure S6A). To investigate whether HSFs are recruited to the c.-27C > A variant and cause transcriptional downregulation, various functional analyses were performed. Electrophoretic mobility shift assays (EMSAs) conducted with nuclear extracts from NCCIT and heat-stressed HEK293 cells showed HSF2 bound with clear preference to the c.-27C > A variant in vitro (Figure S6B), whereas HSF1 bound both alleles in HEK293 extracts (Figure S6C). However, no evidence was found to support a role for HSF in mediating transcriptional silencing of the c.-27C > A variant allele in luciferase reporter assays conducted in HSF-activated host cells (Figure S6D), or in LCLs from family 16 that were demethylated then subjected to heat-stress (Figures S6E–S6F). Because HSFs form trimers on activation by stress stimuli and bind heat-shock elements that typically contain a triad of the pentanucleotide logo (Sakurai and Enoki, 2010), it is possible that the dyad motif created by the c.-27C > A variant fails to confer sufficient binding affinity for activated forms of HSF. Interestingly, however, the EMSAs also showed specific binding of other nucleoproteins unrelated to HSF to the wild-type sequence in the extracts of both cell types (Figures S6B–S6C), suggesting the c.-27C > A variant may result in a loss of interaction with as yet undefined nuclear factors.

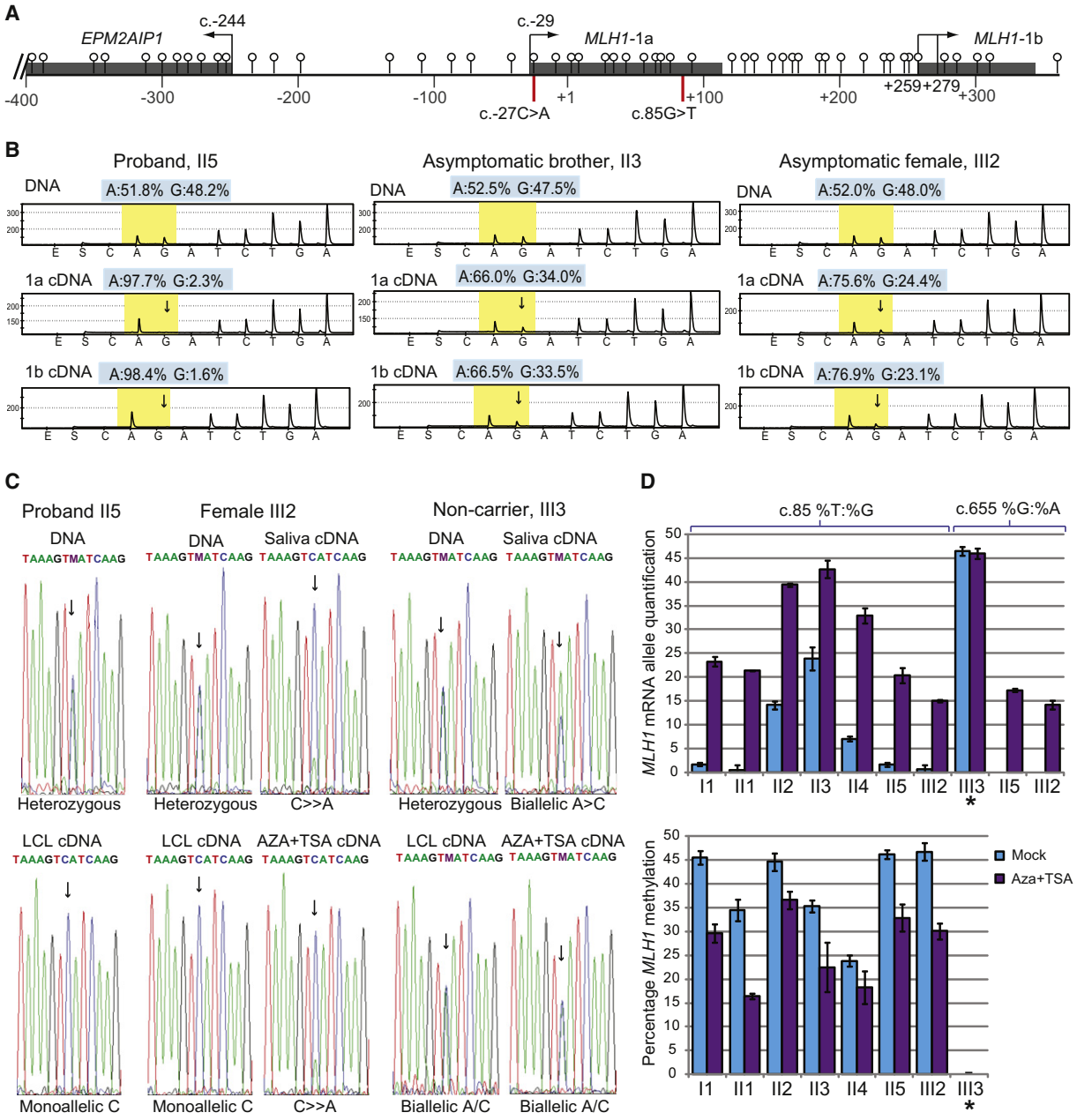


Figure 6. Transcriptional Regulation of the Variant Haplotype in Somatic Cells Is Dependent on the Epigenetic Status of the Bidirectional *EPM2AIP1-MLH1* Promoter

(A) Schematic of the CpG island encompassing the transcription initiation sites of *EPM2AIP1* and two *MLH1* isoforms, 1a and 1b, as identified by 5′RACE in PBL. Nucleotide numbering is with reference to the translation start site (+1) within *MLH1* exon 1a. Gray rectangles show exons, transcription start sites and orientation are indicated by arrows. Lollipops, CpG sites. Relative positions of the single nucleotide variants identified within exon 1a in family 16 are shown. The *MLH1*-1a isoform is transcribed from c.-29, the 1b isoform has two transcription start sites 20 bp apart. Exons 1a and 1b represent alternative first exons and splice directly onto exon 2. The *EPM2AIP1* transcription start site is according to GenBank accession NM_014805.

(B) Concomitant allelic silencing of both *MLH1*-1a and 1b isoforms in carriers of the variant haplotype. Pyrograms from AQ pyrosequencing at the c.655A > G SNP site (shaded yellow) within *MLH1* exon 8 in LCLs from heterozygous carriers of the variant haplotype show a similar reduction of the G allele in both isoforms 1a and 1b.

(C) Representative sequence electropherograms across the A > C SNP (rs9311149) of *EPM2AIP* is shown for informative individuals; proband II5 and III2 (carriers of the variant haplotype) and III3 (noncarrier). Loss or significant reduction of the A allele is shown in cDNAs from II5 and III2, indicating *EPM2AIP1* is subject to concomitant transcriptional silencing with *MLH1* in family members methylated at the paired locus. Allelic silencing is partially reprieved in LCLs treated with 5-aza-2′-deoxycytidine (Aza) and trichostatin A (TSA). In noncarrier III3, both *EPM2AIP1* alleles are represented at similar levels in all cDNAs.

(D) Partial upregulation of the c.85G > T allele concomitant with reduced *MLH1* promoter methylation in LCLs treated with Aza and TSA. Above, histogram showing the mean (±SD) level of expression of the variant c.85T:G allele or c.655G:A in all carriers of the variant haplotype, and noncarrier III3 (*), by AQ pyrosequencing in treatments performed in triplicate. Below, level of *MLH1* promoter methylation in the same cell treatments measured by CpG pyrosequencing. See also Figure S5.

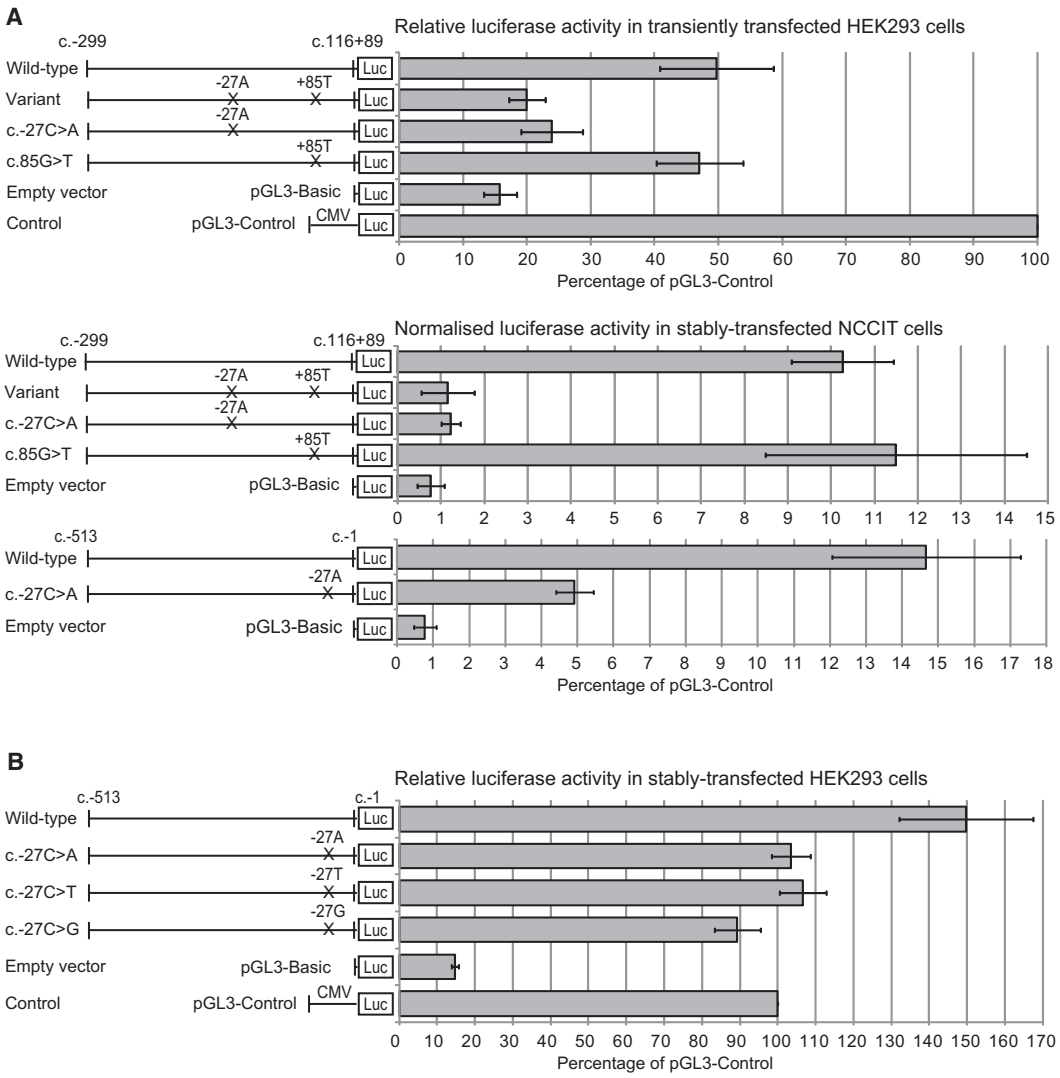


Figure 7. The variant Haplotype and c.-27C > A Substitution Alone Significantly Reduce *MLH1* Promoter Activity in Luciferase Reporter Assays

(A) Left: map of *MLH1* fragments inserted upstream of the *Firefly* luciferase reporter (Luc) in pGL3-Basic. A set of four *MLH1* constructs (c.-299 to c.116+80) containing either the wild-type sequence, variant haplotype (c.-27A and c.85T in *cis*), c.-27A variant alone, or c.85T alone were created to address the functional relevance of the two variants together and separately. A second pair of *MLH1* constructs (c.-513 to −1) containing the wild-type or c.-27A tested the functional relevance of the c.-27C > A variant in the context of the full *MLH1* promoter. Variants are marked with an x. Right: *Firefly* luciferase activity normalized to *Renilla* luciferase in HEK293 cells cotransfected with the *Renilla*-expressing pRL-TK vector. For NCCIT stable transfectants, luciferase activity was normalized to cell number. Histograms show the mean (±SD) normalized luciferase output from transfections performed in triplicate, expressed as a percentage of the pGL3-Control vector.

(B) A further set of constructs containing the core *MLH1* promoter with the wild-type c.-27C site mutated to all three other nucleotides was used to test the functional importance of this nucleotide site. Histogram shows the mean (±SD) relative *Firefly* luciferase normalized to *Renilla* luciferase from stable transfections performed in triplicate, expressed as a percentage of the normalized pGL3-Control vector.

See also Figure S6.

Retention of the Cytosine at Position c.-27 Is Crucial for Optimal *MLH1* Transcription

To determine whether loss of the wild-type cytosine at nucleotide c.-27, or the nature of the substitution in family 16 is of greater functional significance, luciferase promoter reporter assays were performed with constructs containing the full *MLH1* promoter in which the c.-27C was substituted for each of the other nucleotides. The c.-27C > G and c.-27C > T substitutions diminished promoter activity to a similar extent as the c.-27C > A (Figure 7B). This suggests that retention of a cytosine at the c.-27 position is required for optimal *MLH1* transcriptional performance in somatic cells.

DISCUSSION

We show Mendelian inheritance of a constitutional *MLH1* epimutation in a cancer-affected family and demonstrate that the epimutation is likely to be induced by a c.-27C > A single

nucleotide variant within the 5′UTR, located just two bases from the dominant transcription start site. Because the dominant inheritance pattern of the epimutation implicates a causative *cis*-acting genetic defect, we investigated the possibility of large structural alterations across chromosome 3p, as well as point mutations within the *MLH1* CpG island. We found only germline c.-27C > A and c.85G > T substitutions within exon 1. Neither change was found in healthy controls. However, this variant haplotype has previously been reported to segregate with colorectal cancer in another family, although epigenetic and expression studies were not undertaken (Raevaara et al., 2005). The conservative p.A29S amino acid change encoded by the c.85G > T variant has been shown to function normally in terms of mismatch repair efficiency, subcellular localization, protein stability and heterodimerization, arguing against a role for this variant in altered protein activity (Raevaara et al., 2005; Takahashi et al., 2007).

Our finding that the epigenetic aberration was localized to the CpG island promoter shared by *MLH1* and *EPM2AIP1*, and did not involve neighboring genes, suggested a proximate *cis*-genetic cause, further implicating a pathogenic role for one or both of the single nucleotide variants. Analysis of the variants separately and together in luciferase promoter reporter assays led us to conclude that the c.-27C > A variant was the sole cause of transcriptional repression. We also showed that mutagenesis of the c.-27C site reduced transcriptional activity, irrespective of the nature of the nucleotide substitution. This indicates that the wild-type cytosine needs to be retained at this site for efficient transcription in somatic cells. This nucleotide is located two bases from the major transcription initiation site of the *MLH1*-1a isoform, thus may reside within an important, but as yet undefined regulatory element. These findings, together with the preferential binding of undefined nucleoproteins to the c.-27C allele in vitro, suggest that in somatic cells the c.-27C > A variant may result in loss of binding of a nuclear factor necessary for maximal transcription and/or protection from chromatin modification. The link between loss of *MLH1* regulatory elements and induction of epigenetic change is supported by the previous finding of a germline *cis*-deletion encompassing c.-68 to intron 2 in an individual with Lynch syndrome who also displayed constitutional *MLH1* methylation (Gylling et al., 2009).

Autosomal dominant sequence-dependent epimutations have also been described in diseases caused by genes other than *MLH1*. In a familial case of α-thalassemia, an interstitial deletion downstream of *HBA2* caused an epimutation of this gene (Tufarelli et al., 2003). *MSH2* epimutations are caused by a similar mechanism (Ligtenberg et al., 2009). These examples pertain to epimutations caused by large structural alterations in the vicinity of the affected gene. To our knowledge the only other instance of a single nucleotide variant inducing a constitutional epimutation has been in familial chronic lymphocytic leukemia. In this study, heritable *DAPK1* methylation was associated with a single nucleotide variant within a regulatory element over 6 kb upstream, which recruited the HOXB7 repressor (Raval et al., 2007). In neoplasia, the correlation between particular genotypes at germline promoter SNPs and the presence of promoter methylation also suggest an interplay between sequence variation within functional elements and the epigenetic apparatus, as exemplified by *VHL* (Banks et al., 2006), *MGMT* (Hawkins et al., 2009; Ogino et al., 2007), and *GSTP1* (Rønneberg et al., 2008).

Our observations regarding the allelic patterns of *MLH1* methylation and expression in a wide range of tissues from members of family 16 allows us to speculate on the mechanism by which the epimutation is established on the variant haplotype and transmitted through the germline (Figure 8). The erasure of methylation from the variant haplotype and restoration of normal transcriptional activity in the spermatozoa of the cancer-affected male first indicates complete but transient reversion of the epigenetic error in the germline. The coincidence of allelic reactivation with erasure of the repressive epigenetic components indicates the variant haplotype is capable of normal transcription in its epigenetically naked state. Therefore the underlying genetic defect does not directly abrogate transcription per se. The reinstatement of the epimutation on the variant haplotype in somatic cells of successive generations indicates somatic resetting of the epigenetic error with each new generation. However, significant epigenetic heterogeneity was observed throughout the soma of family members, suggesting the underlying genetic defect confers a propensity for the accrual of repressive chromatin modifications in a somatic context, which in turn stabilize transcriptional silencing of the variant haplotype. It appears the variant haplotype may exist in three different methylation and transcriptional states in somatic tissues (labeled Type I, II, and III on Figure 8). Some variant alleles were methylated and transcriptionally silent, some were unmethylated and active, yet others appeared to be unmethylated and transcriptionally silent. The latter finding, taken together with the observations in promoter reporter assays, suggests the involvement of an intermediary nuclear factor influencing transcriptional regulation of the variant allele in somatic but not germ cells. This somatic factor may take the form of a transcriptional repressor, epigenetic modifier or noncoding RNA, which is specifically recruited by the variant haplotype and induces secondary epigenetic modification. An alternative mechanism would be the loss of binding of a chromatin insulator that would normally mask the *MLH1* promoter from epigenetic modification (Figure 8).

Irrespective of the precise mechanism of epigenetic suppression, it is probable that the variant genetic haplotype of *MLH1* identified in family 16 confers cancer susceptibility. Our study should prompt a reevaluation of the pathogenicity of several germline variants identified within the *MLH1* 5′UTR and promoter in Lynch syndrome cases, such as at the neighboring c.-28 nucleotide (Isidro et al., 2003; Müller-Koch et al., 2001).These variants may similarly confer cancer-susceptibility by mediating allelic epigenetic inactivation. Finally our study lends weight to recent reports correlating allele-specific methylation in a small number of nonimprinted genes with particular SNP genotypes (Kerkel et al., 2008; Zhang et al., 2009). Although evidence for genotype-epigenotype interactions is currently limited, our study suggests that germline single nucleotide variants in gene regulatory elements may play a generalized role in inducing epigenetic silencing of tumor suppressor and other disease-associated genes.

EXPERIMENTAL PROCEDURES

Subjects

This study was approved by the Human Research Ethics Committee of St. Vincent's Hospital Sydney and subjects provided written consent. Clinical specimens were handled as described in Supplemental Experimental Procedures.

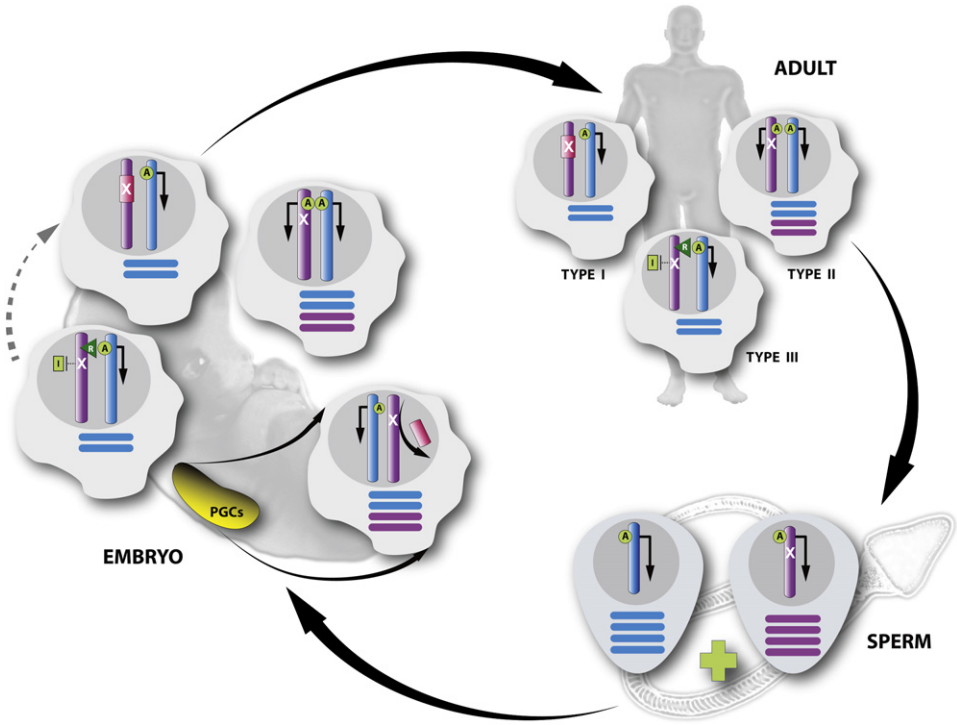


Figure 8. Model of the Potential Mechanism by Which the Variant *MLH1* Haplotype Induces a Soma-Wide Epimutation and Its Mode of Intergenerational Transmission

Wild-type (blue) and variant (purple) *MLH1* alleles and their color-coded mRNA transcripts are shown within somatic or germ cells. The c.-27C > A variant is indicated by a white x and promoter methylation and repressive histone marks are a pink rectangle. Transcriptional activity is indicated by black arrows. In green, transcription activators (A) are represented by a circle, transcription repressors or chromatin modifiers (R) by a triangle, and chromatin insulators (I) by a square. Top right: the variant *MLH1* haplotype is shown in three different epigenetic states in adult somatic cells. In some cells (type I) the variant haplotype is methylated and transcriptionally silent, in other cells (type II) the haplotype is unmethylated and accessible to transcription activators. For type III, it is inferred that the allele is unmethylated but transcriptionally repressed by a putative intermediary nuclear factor. This may take the form of a repressive complex that is recruited to the allele and mediates chromatin modulation, or loss of a protective insulator that normally prevents chromatin modification.

Bottom right: in mature spermatozoa the variant haplotype was completely unmethylated and this coincided with restoration of transcription, indicating erasure of the epimutation in gametes. This shows the variant haplotype is inherently capable of normal transcriptional activity in its naked form. Left: the epimutation is reestablished with each new generation. This is predicted to occur contemporaneously with de novo methylation of other genomic loci in somatic cells of the gastrulating embryo during germ cell lineage commitment. The variant haplotype confers a susceptibility to aberrant de novo methylation through altered interaction with an intermediary nuclear factor(s). Mosaicism may be attributable to competition for occupancy between transcription activators and chromatin modifiers. The three epigenetic states established on the variant haplotype are maintained through mitosis into adulthood, with the onset of methylation permanently stabilizing epigenetic silencing of the affected allele. Erasure of the epimutation from the variant haplotype most likely occurs in the primordial germ cells (PGC) as they migrate through the genital ridge of the developing embryo to become gamete precursors. This process most likely occurs by active demethylation, coinciding with epigenetic reprogramming of imprinted genes. Picture of spermatozoan kindly provided by BIODIDAC.

Methylation Analyses

Genomic DNA (1 μg) was treated with sodium bisulphite using the EZ Methylation Gold Kit (Zymo Research) and 100 ng used for PCR. *MLH1* methylation levels were determined by CpG pyrosequencing, as previously described (Goel et al., 2011). For allelic methylation analyses, PCR was performed with primers flanking the c.-27C > A or the c.85G > T variants that were unbiased with respect to methylation status (Supplemental Experimental Procedures). Amplicons were cloned in the pGEMTEasy vector (Promega), plasmids isolated from single bacterial colonies, and clonal bisulphite sequencing of inserts was performed using vector primers. Quantitative real-time methylation specific PCR was performed as previously described (Hitchins et al., 2007; Kwok et al., 2010).

Allelic Representation Analyses

Allelic expression and LOH analyses were performed by AQ pyrosequencing at the c.85G > T site from the complementary (C > A) strand. First-strand cDNAs and genomic DNA templates were PCR-amplified using the same biotin-labeled forward primer, but different reverse primers; a primer spanning the junction of exons 1 and 2 for cDNAs and an intronic primer for genomic DNA. Pyrosequencing of the denatured biotinylated strand from both cDNA and DNA templates was performed in the antisense direction using the same pyrosequencing primer and nucleotide dispensation order GACATGCGCTG to interrogate sequence A/CTGGCCGCTGG. Primer sequences are given in Supplemental Experimental Procedures. The PyroMark AQ software was used to determine the relative percentage of each allele based on peak height. To normalize for any systematic bias in the AQ assay, allelic expression ratios (AER) were normalized to genomic DNA, calculated as $(cDNA_{\%A}/cDNA_{\%C})/(DNA_{\%A}/DNA_{\%C})$. LOH in tumor DNA (T) was sought compared to normal DNA (N) using the LOH index calculated as $(T_{\%A}/T_{\%C})/(N_{\%A}/N_{\%C})$ and defined as an LOH index <0.6 or >1.7. AQ pyrosequencing at the c.655A > G SNP was performed as previously described (Kwok et al., 2010), with AER and LOH values calculated as the ratio of G:A. Allelic expression analysis of *EPM2AIP1* was performed as previously described (Goel et al., 2011).

Allele-Specific Chromatin Immunoprecipitation

LCLs (6×10^6) were fixed in 0.1% formaldehyde and subjected to ChIP using the Millipore Magna ChIP A Chromatin Immunoprecipitation Kit (Millipore).

Rabbit polyclonal antibodies; anti-H3K9 acetylation, anti-H3K9me3, anti-H3K27me3, anti-H3K4me3, and IgG control (Millipore) were used. Semiquantitative real-time PCR of *MLH1* was performed using two forward primers that differed only at the final 3′ base to specifically amplify the c.-27A or the c.-27C allele in separate reactions, in combination with the same reverse primer. Relative enrichment of each histone modification was calculated as $2^{\Delta Ct\ (input\text{-}eluate)}$ for each locus, with background $2^{\Delta Ct\ (input\text{-}IgG\ negative\ control)}$ subtracted.

Treatment with Epigenetic Reversal Drugs

LCLs (2×10^5/ml) were cultured in media containing 8 μM 5-Aza-2′-deoxycytidine for 72 hr with daily drug replenishment. Cells were then treated with medium containing 300 nM trichostatin A for a further 24 hr. Mock-treatments were performed with drug diluent. Cells were cultured for a further 72 hr without drug prior to harvesting.

Luciferase Promoter Reporter Assays

Promoter reporter constructs containing *MLH1* sequences inserted into the pGL3-Basic vector (Promega) upstream of the *Firefly* luciferase reporter were transfected into HEK293 and NCCIT cells (Supplemental Experimental Procedures). Briefly, transfections were performed in triplicate in parallel with the pGL3-Control luciferase reporter under the SV40 promoter, and empty pGL3-Basic vector. HEK293 cells were cotransfected with pRL-TK *Renilla* luciferase reporter to normalize for transfection efficiency. NCCIT stable transfectants were normalized by cell count and plasmid integration rates were measured by real-time PCR of host genomic DNA.

ACCESSION NUMBERS

The Affymetrix Genome-wide Human SNP 6.0 array data is deposited in Gene Expression Omnibus accession number GSE30348.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.ccr.2011.07.003.

ACKNOWLEDGMENTS

M.P.H. is supported by a CDAII from the Australian NHMRC and Cancer Institute NSW. S.S. was supported by a postdoctoral Endeavour Travel Award from the Australian Federal Government. This work was funded by the NHMRC and Cancer Council NSW. Thanks to Deborah Packham for technical assistance, Genevieve Bennett for sample collection, and Drs. Peter Molloy and Konsta Duesing for copy number analyses. The authors have no conflicts of interest to declare.

Received: February 23, 2011
Revised: May 16, 2011
Accepted: July 5, 2011
Published: August 15, 2011

REFERENCES

Banks, R.E., Tirukonda, P., Taylor, C., Hornigold, N., Astuti, D., Cohen, D., Maher, E.R., Stanley, A.J., Harnden, P., Joyce, A., et al. (2006). Genetic and epigenetic analysis of von Hippel-Lindau (VHL) gene alterations and relationship with clinical variables in sporadic renal cancer. Cancer Res. *66*, 2000–2011.

Bennett, K.L., Mester, J., and Eng, C. (2010). Germline epigenetic regulation of KILLIN in Cowden and Cowden-like syndrome. JAMA *304*, 2724–2731.

Chan, T.L., Yuen, S.T., Kong, C.K., Chan, Y.W., Chan, A.S., Ng, W.F., Tsui, W.Y., Lo, M.W., Tam, W.Y., Li, V.S., and Leung, S.Y. (2006). Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat. Genet. *38*, 1178–1183.

Colella, S., Yau, C., Taylor, J.M., Mirza, G., Butler, H., Clouston, P., Bassett, A.S., Seller, A., Holmes, C.C., and Ragoussis, J. (2007). QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res. *35*, 2013–2025.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation of ras oncogenes in primary human cancers. Biochem. Biophys. Res. Commun. *111*, 47–54.

Gazzoli, I., Loda, M., Garber, J., Syngal, S., and Kolodner, R.D. (2002). 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. *62*, 3925–3928.

Goel, A., Nguyen, T.P., Leung, H.C., Nagasaka, T., Rhees, J., Hotchkiss, E., Arnold, M., Banerji, P., Koi, M., Kwok, C.T., et al. (2011). 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 *128*, 869–878.

Greger, V., Passarge, E., Höpping, W., Messmer, E., and Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum. Genet. *83*, 155–158.

Gylling, A., Ridanpää, M., Vierimaa, O., Aittomäki, K., Avela, K., Kääriäinen, H., Laivuori, H., Pöyhönen, M., Sallinen, S.L., Wallgren-Pettersson, C., et al. (2009). Large genomic rearrangements and germline epimutations in Lynch syndrome. Int. J. Cancer *124*, 2333–2340.

Hawkins, N.J., Lee, J.H., Wong, J.J., Kwok, C.T., Ward, R.L., and Hitchins, M.P. (2009). MGMT methylation is associated primarily with the germline C>T SNP (rs16906252) in colorectal cancer and normal colonic mucosa. Mod. Pathol. *22*, 1588–1599.

Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K., Hamilton, S.R., Kinzler, K.W., et al. (1998). Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc. Natl. Acad. Sci. USA *95*, 6870–6875.

Hitchins, M.P., and Ward, R.L. (2007). Erasure of MLH1 methylation in spermatozoa-implications for epigenetic inheritance. Nat. Genet. *39*, 1289.

Hitchins, M., Williams, R., Cheong, K., Halani, N., Lin, V.A., Packham, D., Ku, S., Buckle, A., Hawkins, N., Burn, J., et al. (2005). MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology *129*, 1392–1399.

Hitchins, M.P., Wong, J.J., Suthers, G., Suter, C.M., Martin, D.I., Hawkins, N.J., and Ward, R.L. (2007). Inheritance of a cancer-associated MLH1 germline epimutation. N. Engl. J. Med. *356*, 697–705.

Isidro, G., Matos, S., Gonçalves, V., Cavaleiro, C., Antunes, O., Marinho, C., Soares, J., and Boavida, M.G. (2003). Novel MLH1 mutations and a novel MSH2 polymorphism identified by SSCP and DHPLC in Portuguese HNPCC families. Hum. Mutat. *22*, 419–420.

Kerkel, K., Spadola, A., Yuan, E., Kosek, J., Jiang, L., Hod, E., Li, K., Murty, V.V., Schupf, N., Vilain, E., et al. (2008). Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. Nat. Genet. *40*, 904–908.

Kwok, C.T., Ward, R.L., Hawkins, N.J., and Hitchins, M.P. (2010). Detection of allelic imbalance in MLH1 expression by pyrosequencing serves as a tool for the identification of germline defects in Lynch syndrome. Fam. Cancer *9*, 345–356.

Ligtenberg, M.J., Kuiper, R.P., Chan, T.L., Goossens, M., Hebeda, K.M., Voorendt, M., Lee, T.Y., Bodmer, D., Hoenselaar, E., Hendriks-Cornelissen, S.J., et al. (2009). Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3′ exons of TACSTD1. Nat. Genet. *41*, 112–117.

Lin, J.C., Jeong, S., Liang, G., Takai, D., Fatemi, M., Tsai, Y.C., Egger, G., Gal-Yam, E.N., and Jones, P.A. (2007). Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. Cancer Cell *12*, 432–444.

Lynch, H.T., and Lynch, J.F. (2004). Lynch syndrome: history and current status. Dis. Markers *20*, 181–198.

Miyakura, Y., Sugano, K., Akasu, T., Yoshida, T., Maekawa, M., Saitoh, S., Sasaki, H., Nomizu, T., Konishi, F., Fujita, S., et al. (2004). Extensive but hemiallelic methylation of the hMLH1 promoter region in early-onset sporadic

colon cancers with microsatellite instability. Clin. Gastroenterol. Hepatol. *2*, 147–156.

Morak, M., Schackert, H.K., Rahner, N., Betz, B., Ebert, M., Walldorf, C., Royer-Pokora, B., Schulmann, K., von Knebel-Doeberitz, M., Dietmaier, W., et al. (2008). 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. *16*, 804–811.

Müller-Koch, Y., Kopp, R., Lohse, P., Baretton, G., Stoetzer, A., Aust, D., Daum, J., Kerker, B., Gross, M., Dietmeier, W., and Holinski-Feder, E. (2001). 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. *6*, 473–482.

Ogino, S., Hazra, A., Tranah, G.J., Kirkner, G.J., Kawasaki, T., Nosho, K., Ohnishi, M., Suemoto, Y., Meyerhardt, J.A., Hunter, D.J., and Fuchs, C.S. (2007). MGMT germline polymorphism is associated with somatic MGMT promoter methylation and gene silencing in colorectal cancer. Carcinogenesis *28*, 1985–1990.

Peltomäki, P., and Vasen, H. (2004). Mutations associated with HNPCC predisposition — Update of ICG-HNPCC/INSiGHT mutation database. Dis. Markers *20*, 269–276.

Raevaara, T.E., Korhonen, M.K., Lohi, H., Hampel, H., Lynch, E., Lönnqvist, K.E., Holinski-Feder, E., Sutter, C., McKinnon, W., Duraisamy, S., et al. (2005). Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology *129*, 537–549.

Raval, A., Tanner, S.M., Byrd, J.C., Angerman, E.B., Perko, J.D., Chen, S.S., Hackanson, B., Grever, M.R., Lucas, D.M., Matkovic, J.J., et al. (2007). Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. Cell *129*, 879–890.

Rønneberg, J.A., Tost, J., Solvang, H.K., Alnaes, G.I., Johansen, F.E., Brendeford, E.M., Yakhini, Z., Gut, I.G., Lønning, P.E., Børresen-Dale, A.L., et al. (2008). GSTP1 promoter haplotypes affect DNA methylation levels and promoter activity in breast carcinomas. Cancer Res. *68*, 5562–5571.

Sakurai, H., and Enoki, Y. (2010). Novel aspects of heat shock factors: DNA recognition, chromatin modulation and gene expression. FEBS J. *277*, 4140–4149.

Schofield, L., Watson, N., Grieu, F., Li, W.Q., Zeps, N., Harvey, J., Stewart, C., Abdo, M., Goldblatt, J., and Iacopetta, B. (2009). Population-based detection of Lynch syndrome in young colorectal cancer patients using microsatellite instability as the initial test. Int. J. Cancer *124*, 1097–1102.

Suter, C.M., Martin, D.I., and Ward, R.L. (2004). Germline epimutation of MLH1 in individuals with multiple cancers. Nat. Genet. *36*, 497–501.

Takahashi, M., Shimodaira, H., Andreutti-Zaugg, C., Iggo, R., Kolodner, R.D., and Ishioka, C. (2007). Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. Cancer Res. *67*, 4595–4604.

Tufarelli, C., Stanley, J.A., Garrick, D., Sharpe, J.A., Ayyub, H., Wood, W.G., and Higgs, D.R. (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. Nat. Genet. *34*, 157–165.

Zhang, Y., Rohde, C., Tierling, S., Jurkowski, T.P., Bock, C., Santacruz, D., Ragozin, S., Reinhardt, R., Groth, M., Walter, J., and Jeltsch, A. (2009). DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. PLoS Genet. *5*, e1000438.