Primary constitutional MLH1 epimutations: a focal epigenetic event

Estela Dámaso1, Adela Castillejo2, María del Mar Arias3, Julia Canet-Hermida1, Matilde Navarro1, Jesús del Valle1, Olga Campos1, Anna Fernández1, Fátima Marín1, Daniela Turchetti4, Juan de Dios García-Díaz5, Conxi Lázaro1, Maurizio Genuardi6, Daniel Rueda7, Ángel Alonso8, Jose Luis Soto9, Megan Hitchins9, Marta Pineda1 and Gabriel Capellá1

BACKGROUND: Constitutional MLH1 epimutations are characterised by monoallelic methylation of the MLH1 promoter throughout normal tissues, accompanied by allele-specific silencing. The mechanism underlying primary MLH1 epimutations is currently unknown. The aim of this study was to perform an in-depth characterisation of constitutional MLH1 epimutations targeting the aberrantly methylated region around MLH1 and other genomic loci.

METHODS: Twelve MLH1 epimutation carriers, 61 Lynch syndrome patients, and 41 healthy controls, were analysed by Infinium 450 K array. Targeted molecular techniques were used to characterise the MLH1 epimutation carriers and their inheritance pattern.

RESULTS: No nucleotide or structural variants were identified in-cis on the epimutated allele in 10 carriers, in which inter-generational methylation erosion was demonstrated in two, suggesting primary type of epimutation. CNVs outside the MLH1 locus were found in two cases. EPM2AIP1-MLH1 CpG island was identified as the sole differentially methylated region in MLH1 epimutation carriers compared to controls.

CONCLUSION: Primary constitutional MLH1 epimutations arise as a focal epigenetic event at the EPM2AIP1-MLH1 CpG island in the absence of cis-acting genetic variants. Further molecular characterisation is needed to elucidate the mechanistic basis of MLH1 epimutations and their heritability/irreversibility.

INTRODUCTION
Lynch syndrome (LS) is characterised by an increased risk for colorectal cancer (CRC) and cancers of the endometrium, ovary, stomach, small intestine, hepatobiliary tract, urinary tract, brain, and skin1. LS is caused by a germline genetic variant within a mismatch repair (MMR) gene, MLH1 (MutL homolog 1), MSH2 (MutS Homolog 2), MSH6 (MutS homolog 6) or PMS2 (PMS1 Homolog 2), or a terminal deletion of EPCAM (Epithelial Cell Adhesion Molecule gene) with consequent epigenetic inactivation of MSH2. In a small proportion of LS patients the cancer predisposition is caused by a constitutional epimutation of MLH1, in which one allele of the CpG island promoter is aberrantly hypermethylated throughout normal tissues with associated loss-of-expression from this allele2.

Two types of constitutional MLH1 epimutation have been defined: secondary, which are linked in-cis to a genetic alteration and follow an autosomal dominant pattern of inheritance5–10, and primary, which occur in the absence of any apparent linked sequence change, typically arise de novo6,11–16, and demonstrate null10,11,13,17 or non-Mendelian inheritance9,11,18. To date, 75 index cases with a constitutional MLH1 epimutation have been reported19–20, accounting for 2–3% of mutation negative cases with suspected LS whose tumors are MLH1-deficient17. Most of these (66/75) have been considered primary9,9,20. The available evidence from these cases suggests that constitutional MLH1 epimutations cause a severe LS phenotype, including a young age of cancer onset and multiple primary tumors.

Previous studies on constitutional MLH1 epimutation have focused on the role of this molecular defect in cancer causation by confirming the presence of MLH1 promoter methylation and corresponding transcriptional silencing within normal tissues, refining the selection criteria for patients warranting screening for it, and the inheritance patterns within families. It has been proposed that primary MLH1 epimutations arise in the germline or early stages of embryonic development, since they are monoallelic and soma-
wide, but frequently exhibit mosaic methylation and expression loss\(^1\). However, no comprehensive studies have been undertaken to elucidate the mechanism(s) underlying primary MLH1 epimutations, such as whether they are localized within the MLH1 locus due to a focal defect, or if additional genes are concomitantly affected due to widespread epigenetic perturbation. The main aim of our study was to perform an in-depth characterisation of the methylation profile in cases with a confirmed constitutional MLH1 epimutation to define the extent of the aberrantly methylated region around MLH1, as well as other loci on a genome-wide scale. The results from this study contribute to the understanding of primary constitutional MLH1 epimutations by showing that the methylation error occurs in a localized manner.

**PATIENTS AND METHODS**

**Patients and samples**

Twelve Caucasian MLH1 epimutation carriers were recruited from the west Mediterranean area, among which eight were previously reported\(^12\,23\). The four new cases were referred from Complejo Hospitalario de Navarra and Doce de Octubre University Hospital due to the identification of MLH1 methylation in blood by MS-MLPA. The original numerical code for each proband was maintained from previous publications; however, a prefix corresponding to the referral center (below) was added (Table S1). Clinico-pathological data collected included age at cancer diagnosis, tumor location, immunohistochemistry of MMR proteins in tumors, and MLH1 methylation status in tumor and blood DNA (Table S1). Forty-one healthy controls (selected to appropriately match the patients by age, race, and geographic location), 61 LS patients with confirmed genetic mutations (21 MLH1, 28 MSH2, 9 MSH6, 3 PMS2), and four CRC patients whose tumors demonstrated somatic MLH1-methylation, were included as controls (Table S2). Written informed consent was obtained from all individuals, and the ethics committee of the respective hospitals approved the study.

Samples of peripheral blood leukocytes (PBL) DNA were collected from biobanks at the Catalan Institute of Oncology (ICO), Eche University Hospital (HGUDE), Valencia Biobank Network (IBSP-CV), Complejo Hospitalario de Navarra (CHN), Doce de Octubre University Hospital (H120), and Policlínico S.Orsola-Malpighi (M). Lymphoblastoid cell lines were established from the 8 lymphocytes of MLH1 epimutation carriers by standard EBV-transformation using supernatant from the B95.5 cell lines, as described in Supplementary Methods. Colorectal tumors and distal normal mucosa were collected as formalin-fixed paraffin embedded (FFPE) tissues when available (Table S2).

For each FFPE specimen, 10–20 × 10-μm sections were cut from a single-representative block per case, using macrodissection with a scalpel if needed to enrich for tumor cells. After deparaffinization using Deparaffinization Solution (Qiagen, Hilden, Germany), DNA was isolated using the QIAamp DNA FFPE Tissue Kit (Qiagen) according to the manufacturer’s instructions.

**Methylation analyses**

The initial detection or confirmation of the presence of MLH1 methylation in each proband and consenting family members was performed by methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) on genomic DNA from PBL. For all other methylation analyses, DNA was subject to sodium bisulphite modification using the EZ DNA Methylation Kit (Zymo Research), according to the manufacturer’s instructions.

**MS-MLPA.** The presence of methylation at the MLH1 promoter was detected and confirmed using 100 ng of PBL DNA using the SALSA MLPA ME011 Mismatch Repair genes probemix (MRc-Holland), according to the manufacturer’s instructions. DNA from the RKO CRC cell line was used as the MLH1 methylation-positive control. The amplification products were run on an ABI Prism 3130 DNA sequencer and analysed using GeneMapper v4.0 (Applied Biosystems).

**Clonal bisulfite sequencing.** Clonal bisulfite sequencing of fragments of the MLH1 promoter was used to determine the allelic methylation profile in bisulfite modified PBL DNA from MLH1 epimutation carriers who were found to be heterozygous for a promoter variant. Primers were used to amplify specific promoter regions encompassing each promoter variant (Table S3), and PCR products were cloned in E. coli cells using the TOPO TA cloning kit (Invitrogen) to separate the ampiclons into individual alleles. The inserts from at least 24 individual colonies were sequenced using plasmid vector primers. Additional methodological details are provided in Supplementary Methods. The methylation status at individual CpG site and the allele of each heterozygous variant within each promoter fragment was determined using SeqMan software (DNASTAR) or Mutation Surveyor (SoftGenetics).

**Methylation array analysis.** PBL DNA samples (1000 ng) from patients and controls, and 500–1000 ng of available FFPE colorectal cancer DNA and distal normal colon tissue DNA, were randomized within 96-well plates. For internal quality control, in vitro methylated and unmethylated DNAs were included in each batch.

For quality control purposes, PBL DNA integrity was evaluated by agarose gel electrophoresis and spectrophotometry. DNAs from FFPE samples were analysed by qPCR using the Infinium FFPE QC kit (Illumina), and samples showing a delta-Ct value below five were restored using the Infinium HD FFPE Restore kit (Illumina), following the manufacturer’s instructions.

Quantity of 1000 ng PBL and 500 ng FFPE DNAs were bisulfite converted using the EZ DNA Methylation Kit (Zymo Research), according to the manufacturer’s instructions. To determine the efficiency of the bisulfite conversion, a predetermined genomic region was evaluated by Sanger sequencing in the methylated and unmethylated controls and one sample from each batch.

Genome-wide methylation profiling was performed using the Infinium Human Methylation 450 K Beadchip (Illumina), which interrogates the methylation status of 485,764 CpG sites across the genome. After hybridization, sample scanning was performed using the HiScan platform (Illumina), which has a laser scanner with two colors (332 nm/660 nm). The relative intensity of each dye was analysed using GenomeStudio software (Methylation Module). For each analysed CpG site, a β-value was obtained depending on the fluorescence intensity. β measures took values between 0 (unmethylated) and 1 (fully methylated). Analysis for batch effects was performed using RnBeads software (Max-Planck-Institut Informatik). Group comparisons and statistical analyses (based on differentially methylated CpG sites, CpG islands, promoters, genes and tiling) were also performed using RnBeads software. CpG methylation was visualized by Integrative Genome Viewer (Broad Institute). GRCh37/hg19 was used as the reference genome.

The data discussed in this publication are accessible through GEO Series accession number GSE107353 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107353).

**Analysis of Alu elements.** The sequence encompassing the MLH1-EPM2AIP1 CpG island identified as the differentially methylated region (DMR) in carriers of an MLH1 epimutation was screened for repetitive elements by the RepeatMasker Web Server (Institute for Systems Biology) using “Cross_match” as the search engine. No repetitive elements were found in this DMR. Downstream of the DMR, Alu elements were found in MLH1 intron 1, as previously described\(^24\). The methylation status of these Alu elements was tested in PBL DNA from three MLH1 epimutation carriers, one healthy control and the RKO cell line by bisulfite sequencing of the PCR fragments designated p24 (8 CpGs from Alu 2) and p25 (1...
CpG from Alu 3), as previously described\(^2^4\). Amplification of the region p10 (control region, including 6 CpGs, unmethylated in normal tissues and methylated in cancer tissues) was performed using methylation-specific primers. Primers and conditions are summarized in Table S3.

Mutational analyses

Screen of the EPM2AIP1-MLH1 CpG island for point mutations. The EPM2AIP1-MLH1 CpG island is defined by its CpG content as the region Chr3:37,034,229–37,035,355 by the UCSC genome browser (CpG:93). Screening for promoter variants was performed by amplification of two overlapping PCR fragments followed by direct sequencing. Primers and PCR conditions are provided in Table S3. Sequencing was performed using the BigDye Terminator v.3.1 Sequencing Kit (Applied Biosystems). Mutation Surveyor (SoftGenetics) was used for sequence visualization.

Detection of structural variants around the MLH1 locus. A screen for the presence of structural alterations was performed on PBL DNA from MLH1 epimutation carriers using a custom high-definition CGH array designed with eArray Software (Agilent Technologies) and manufactured by Agilent’s SurePrint oligo technology. This comprised 15,000 probes encompassing the MLH1 locus (region Chr3:36,450,000–37,900,000 within cytoband 3p22.2), with an average probe spacing of 100 bp intervals. Bioinformatics analysis was performed in R using the 2.15.12 Bioconductor statistical packages. Results were visualized on CytoGenomics Software (Agilent Technologies).

In silico functional evaluation of the genetic variants identified PROMO 3.0 software\(^2^5,2^6\) was used to predict any changes in transcription factor binding between wild-type alleles and promoter variants. Only human transcription factors were considered and 5% was selected as maximum matrix dissimilarity rate.

Structural variants were compared with those previously reported in the Database of Genomic Variants\(^2^7\), using chromosomal coordinates.

MLH1 allelic expression analyses

For allelic expression analyses at the MLH1 single-nucleotide polymorphism (SNP) rs1799977 (c.655A>G), the relative levels of MLH1 protein were considered and 5% was selected as maximum matrix dissimilarity rate.

For allelic expression analyses at the MLH1 single-nucleotide polymorphism (SNP) rs1799977 (c.655A>G), the relative levels of MLH1 gene (Chr3:36,416,879–36,416,879) (Figure S4B). This region appears to be prone to copy number variations in the control region (DGV database), as shown in Table S4.

Determination of the methylation-associated allele (MAA) Clonal bisulfite sequencing was performed on PBL DNA of probands harboring a heterozygous variant within the MLH1 promoter region to determine the MAA. Dense monoallelic MLH1 methylation was linked to the G allele at SNP rs1800734 (c.—93G>A) in cases H120-A and HGUE_2 and to the A allele in patients CHN_1 and HGUE_5 (Table 1). In case HGUE_5, who was heterozygous for both the novel heterozygous variant c.—234_–236del and the c.—93G>A SNP, methylation was linked to the non-deleted and c.—93A alleles, and the c.—234_–236del and c.—93G>G alleles were unmethylated (Figure S5). This finding indicated the c.—234_–236del variant was not causally linked to the epimutation in HGUE_5. In patient ICO_34, methylation was linked to the G allele at SNP rs34566456 (c.—93G>C) with ~38% of G alleles methylated, indicating mosaicism (Table 1), consistent with our prior findings using other methods\(^2^1\).

Inheritance pattern of the epimutant alleles and designation of MLH1 epimutation as likely primary or secondary To investigate the inheritance pattern of the genetic allele harboring the epimutation, the MLH1 promoter methylation status and haplotypes within a 12 Mb around MLH1 were determined in available PBL DNA from the first-degree relatives of epimutation carriers (Table S5). MS-MLPA analysis showed no evidence of MLH1 methylation in any of the relatives tested from any family,
Table 1. Results of molecular characterisation of constitutional MLH1 epimutation carriers in our series

| Case ID | % MLH1 promoter methylation in blood (MS-MLPA) | Germline variants | Functional impact of epimutation | Inheritance pattern | Other features |
|---------|-----------------------------------------------|-------------------|--------------------------------|---------------------|---------------|
| ICO_1   | 56                                            | None              | Not identified                  | MLH1 c.655A and EPM2AIP1 c.*2570G | Blood lymphocytes | MLH1 = 0 and EPM2AIP1 = 0.02 | MLH1 c.−93G>A | 100% of c.−93A clones | One daughter (Unmethylated) | Inter-generational erasure |
| ICO_34  | 25                                            | None              | Not identified                  | EPM2AIP1 c.*2570G   | Blood lymphocytes | 0.48              | c.−593G>C | MLH1 c.−593G | 38% of c.−593G clones (3 of 8 clones) | Two sisters (both unmethylated) |
| M60     | 56                                            | None              | Not identified                  | MLH1 c.655A         | Blood lymphocytes | 0.3               | WT          | No promoter variants | NA | NA |
| HGUE_1  | 48                                            | None              | One deletion (20 Kb) outside DM CpG island | No heterozygous coding variants | NA | NA | WT | No promoter variants | NA | NA |
| HGUE_2  | 64                                            | None              | Two deletions (15 and 19 Kb) outside DM CpG island | No heterozygous coding variants | NA | NA | c.−93G>A | MLH1 c.−93G | 94% of c.−93G clones (16 of 17 clones) | No transmission of MAA |
| HGUE_3  | 70                                            | None              | Not identified                  | No heterozygous coding variants | NA | NA | WT | No promoter variants | NA | NA |
| HGUE_4  | 49                                            | None              | Not identified                  | No heterozygous coding variants | NA | NA | WT | No promoter variants | NA | NA |
| HGUE_5  | 49                                            | c.−234_−236 delCTT in trans with methylation | Not identified                  | No heterozygous coding variants | NA | NA | c.−234_−236 delCTT and c.−93G>A | MLH1 c.−93A (allele without deletion) | 80% of c.−93A clones (8 of 10 clones) | NA |
| CHN_1   | 47                                            | None              | Not identified                  | MLH1 c.655A         | Blood lymphocytes | 0 ± 0.00          | MLH1 c.−93A | 100% or c.−93A clones (9 clones) | One sister and proband’s son and daughter (all of them unmethylated) | Inter-generational erasure |
| CHN_2   | 36                                            | None              | Not identified                  | MLH1 c.655A         | EBV cell line (44% methylation) | 0.53 ± 0.04        | WT          | No promoter variants | NA | NA |
| H12O_A  | 40                                            | None              | Not identified                  | MLH1 c.655G         | EBV cell line (9% methylation) | 0.79 ± 0.03        | c.−93G>A | MLH1 c.−93G | 100% of c.−93G clones (13 clones) | NA |
| H12O_B  | 47                                            | None              | Not identified                  | No heterozygous coding variants | NA | NA | WT | No promoter variants | NA | NA |

New results obtained in the current study are highlighted in bold. Abbreviations: DM = differentially methylated; MAA = methylation-associated allele; NA = not available.
thus no case of transmission of an epimutation was observed among the family members tested (Figs. 1, 2). Haplotype analyses revealed that the genetic allele bearing the epimutation (in the proband) was shared in unmethylated form by two sisters of case ICO_34, suggesting a de novo occurrence of the epimutation in proband ICO_34; one sister and two offspring of case CHN_1, suggesting de novo occurrence of the epimutation in proband CHN_1 and indicating inter-generational erasure of the epimutation in her offspring (Fig. 1). The inter-generational erasure newly observed in case CHN_1 herein, and previously observed in case ICO_113, allowed us to definitively classify the epimutation in these two cases as primary. In six additional cases, the absence of methylation in the relatives tested, and/or the lack of a family history of cancer, also suggested their epimutation was of the primary type.

In patient HGUE_2, although clonal bisulphite sequencing showed monoallelic methylation of the c.-93A allele, we were unable to determine whether this was linked in-cis to either or both of the 15 Kb and 19 Kb deletions flanking MLH1 that were also identified in this patient. The MAA was not transmitted to either of his offspring, hence this family was non-informative for whether the type of epimutation carried by this proband was primary or secondary. Similarly, in case HGUE_1, we were unable to determine whether the epimutation was linked in-cis to the 20 Kb deletion identical upstream of MLH1, hence whether the deletion may have been causally related.

Global epigenetic characterisation of MLH1 constitutional epimutations

In order to evaluate whether aberrant methylation was restricted to MLH1 or extended to other loci across the genome in MLH1 epimutation carriers, methylation array analysis was performed using the Infinium 450 k array in PBL DNA. This revealed that the MLH1-EP2MAIP1 promoter CpG island was the sole differentially methylated region (DMR) in MLH1 epimutation carriers when compared to healthy controls (false discovery rate (FDR) adjusted p-value = 2.45E–42) (Fig. 3a). The same result was obtained when MLH1 epimutation carriers were compared to LS cases (FDR adjusted p-value = 3.45E–61) (Fig. 3b), and the subset of LS cases with germline MLH1 mutation (FDR adjusted p-value = 6.41E–21) (Fig. 3c). Accordingly, the same analysis identified MLH1 and EPM2AIP1 as the unique differentially methylated promoters in MLH1 epimutation carriers compared to healthy controls (EPM2AIP1 FDR adjusted p-value = 2.21E–43, MLH1 FDR adjusted p-value = 2.19E–35), LS patients (EPM2AIP1 FDR adjusted p-value = 5.07E–61, MLH1 FDR adjusted p-value = 9.23E–51), and MLH1-mutated LS cases (EPM2AIP1 FDR adjusted p-value = 8.41E–21, MLH1 FDR adjusted p-value = 1.51E–16) (Fig. 3).

A deeper analysis at the MLH1 promoter region showed that the methylation levels in PBL from constitutional epimutation carriers spread across the entire CpG island, MLH1 exon 1 and the first part of intron 1 (Fig. 4). The CpG island shores were also affected by hypermethylation, represented by probes at positions chr3: 37,033,791 and chr3: 37,035,400 at the boundaries of the 1.6 Kb DMR. Of note, no other common DMR were found in the genome-wide methylation analysis among the carriers of a MLH1 constitutional epimutation.

The DMR identified in the PBL of MLH1 epimutation carriers was also present in subject-matched normal colonic mucosa and colorectal cancer (CRC) (when available for analysis). Although methylation levels across the DMR were marginally higher in CRC tissue than normal tissues, this did not reach statistical
significance, suggesting that methylation is not a common second hit in these tumors (Fig. 5a). Similarly, no significant differences were found in the levels of methylation among the tissues tested from mutation-positive LS cases (Fig. 5b).

In normal colonic mucosa tissue, a lack of MLH1 methylation was similarly observed in the LS control group and the sporadic cases with somatic MLH1 methylation in their tumors (Figure S3A). By contrast, in tumor tissue, the sporadic MLH1-methylated CRC cases showed a more extensive DMR spanning ≈50 Kb (from chr3:36,985,516 to chr3:37,035,399), within which four CpG sites located at chr3:36,986,513; 36,986,532; 36,986,555, and 36,986,642 were differentially methylated only in the sporadic CRC (FDR adjusted p-value <0.01) (Figure S3B and S3C).

Methylation state of Alu elements
In order to evaluate whether demethylation of MLH1 intron 1 Alu sequences is associated with constitutional epimutation, as has been suggested for sporadic MSI tumors24, bisulfite sequencing analysis was performed in PBL DNA from 3 MLH1 epimutation carriers (cases ICO_1, ICO_34, and CHN_1), one healthy control and the RKO cell line. As previously described, partial demethylation was observed in Alu 3 of RKO cells (Figure S4). By contrast, the intron 1 Alu CpGs analysed were found to be fully methylated in carriers of a constitutional MLH1 epimutation and a healthy control (Figure S4), arguing against demethylation of these nearby Alu sequences as a causal mechanism for constitutional epimutation.

DISCUSSION
A comprehensive genetic and epigenetic analysis has provided a deeper and broader molecular characterisation of 12 cases with constitutional MLH1 epimutation than previously undertaken. The series, which includes 4 novel cases, accounts for 16% of all index
cases of constitutional MLH1 epimutation reported to date (Table S6). Our findings suggest that the epimutations in at least 10 of these cases are primary based on the following: (i) the absence of rare in-cis genetic aberrations within the MLH1 CpG island, (ii) the lack of familial cancer history and the presence of the MAA in an unmethylated state in siblings and/or parents that point to a de novo occurrence of the epimutation in the probands (Table S1 and S5); and (iii) the demonstration in two cases of the inter-generational erasure of the epimutation in offspring who inherited the MAA. Interestingly, few studies of MLH1 epimutation that have involved family members from more than one generation have ever been undertaken, but among these, both inter-generational inheritance and erasure of the epimutation between the proband and offspring have been described.

Fig. 3 Scatterplot of the normalized mean $\beta$-values obtained using the Infinium 450 k Human Methylation array to identify differentially methylated CpG islands and promoters in blood DNA from constitutional MLH1 epimutation carriers against healthy controls a and Lynch syndrome patients with a germline sequence mutation b. The transparency corresponds to point density. One % of the points in the sparsest populated plot regions are drawn explicitly. The colored points represent differentially methylated CpG islands (depicted as arrow #1) and promoters (arrows #2 and #3). FDR adjusted $p$-values of differentially methylated regions are summarized at the bottom c.
Healthy individuals, Mutation-positive Lynch syndrome patients, Mean alterations in-cis epimutations have been reported with concomitant genetic deletions occurred on the MAA. Noteworthy, secondary sample availability. Also, we could not determine whether the presence of these deletions by other methods due to the lack of MLH1 single-nucleotide variant, and a large duplication encompassing MLH1 autosomal dominant inheritance of the absence of family history and methylation in blood from degree relatives, led us to propose to classify them as (Chr3:37,033,791–37,035,399) in PBL DNA demonstrated that a region of 1.6 Kb DMR unique to the group of MLH1 epimutation carriers. Above, CpG islands (CI) are represented as black rectangles and their shores are represented in grey. Location of Deng’s promoter regions (DR) is indicated as white rectangles A–D. Genes (G) containing the displayed CpG sites are represented as grey rectangles. Ensembl GRCh37 was taken as reference for gene coordinates (Table S5). The two cases of inter-generational erasure, we describe herein provide further evidence of the reversibility of this defect between generations (Table S5, Fig. 1).

In two cases, HGUE_1 and HGUE_2, deletions external to MLH1 were identified by aCGH. However, we were unable to confirm the presence of these deletions by other methods due to the lack of sample availability. Also, we could not determine whether the deletions occurred on the MAA. Noteworthy, secondary MLH1 epimutations have been reported with concomitant genetic alterations in-cis, including promoter deletions, a promoter single-nucleotide variant, and a large duplication encompassing MLH1 and neighboring genes. These cases have displayed autosomal dominant inheritance of the MLH1 epimutation with faithful segregation of the methylation with the genetic alteration. Therefore, although we cannot completely rule out that deletions are causative of the epimutation in cases HGUE_1 and HGUE_2, the absence of family history and methylation in blood from first-degree relatives, led us to propose to classify them as “suspected primary” (Table S6).

Transcriptional silencing usually correlates with methylation levels (Table S6). In case H12O-A, differences in the values of methylation and transcriptional silencing could be found depending upon the sample tested. The lymphoblastoid cell line displayed less methylation than that observed in blood, in line with a previous observation. These results highlight the limitations of using lymphoblastoid cell lines for epigenetic-based analyses, although this may be a necessity when primary tissue samples from patients are limited.

The array-based genome-wide methylation analysis performed in PBL DNA demonstrated that a region of 1.6 Kb (Chr3:37,033,791–37,035,400)—encompassing the shared MLH1-EPM2AIP1 CpG island and its corresponding shores—was the sole DMR unique to the group of MLH1 epimutation carriers, as compared to mutation-positive LS patients or healthy controls. For cases HGUE_1 and HGUE_2, classified as suspected primary epimutations the same 1.6 Kb DMR was identified. No differences in the length of the DMR were observed in the normal colon mucosa or CRC tissue among the MLH1 epimutation carriers. The absence of any other DMR in the MLH1 epimutation carrier group suggests that the constitutional epimutation is caused by a shared focal event, since it spans the very same region in all patients tested. Of note, our preliminary analysis does not suggest a genetic cause. Also, we have provided evidence against MLH1 intron 1 Alu sequence demethylation as a causal mechanism for constitutional MLH1 epimutation.

Interestingly, the extent of the DMR observed (consistently in all three tissue types analysed) in MLH1 epimutation carrier was narrower than that observed in sporadic MLH1-methylated CRC. Although our cases of sporadic MLH1-methylated CRC were few (n = 4), the extent of their methylation (50 Kb) is consistent with tumors with available MSI and methylation data from TCGA project (n = 43) (Figure S5).

While the causal mechanism of primary constitutional MLH1 epimutations remains unknown, it may be speculated that it acts in the very early stages of embryonic development. MLH1 de novo methylation has been induced in vitro in human embryonic stem cells suggesting that pluripotency is a mandatory requirement for establishing the epimutation. The detection of non-methylated clones on the MAA in distinct patients, including those with methylation levels close to 50% (Table S5), suggests that if an epimutation is established in the germline, some alleles may lose the methylation during the demethylation phase that occurs between zygote formation and the preimplantation embryogenesis. The methylated allele could behave as a “leaky” allele when self-propagating the MLH1 epimutation within individual cells.

Our results show that aberrant methylation of the MLH1 promoter is targeted, as opposed to disseminated, since we found no other common DMRs among the MLH1 epimutation...
Fig. 5  Representation of differentially methylated region in blood, normal colonic mucosa, and colorectal cancer tissue in constitutional MLH1 epimutation patients a and mutation-positive Lynch syndrome patients b. $\beta$-values obtained after global methylation analysis are displayed against CpG sites’ coordinate location. The location of the CpG sites is not drawn to scale. Above, EMP2AIP1-MLH1 CpG island (CI) is represented as black rectangle and their shores are represented in grey. Location of Deng’s promoter regions (DR) are indicated as white rectangles A–D. Genes (G) including displayed CpG sites are represented as grey rectangles. Ensembl GRCh37 was taken as reference for gene coordinates.

Conflict of interest: The authors declare that they have no conflict of interest.

REFERENCES
1. Lynch, H. T., Snyder, C. L., Shaw, T. G., Heinig, C. D. & Hitchins, M. P. Milestones of Lynch syndrome: 1895-2015. Nat. Rev. Cancer 15, 181–194 (2015).
2. Hitchins, M. P. Constitutional epimutation as a mechanism for cancer causality and heritability? Nat. Rev. Cancer 15, 625–634 (2015).
3. Hitchins Megan, P. et al. Dominantly inherited constitutional epigenetic silicancer-affected family is linked to a single nucleotide variant within truncing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the SUITR. Cancer Cell 20, 200–213 (2011).
4. Kwok, C.-T. et al. The MLH1 c.-27C>A and c.85G>T variants are linked to dominantly inherited constitutional epigenetic silicancer-affected family is linked to a single nucleotide variant within truncing of MLH1 promoter regions (DR) are indicated as white rectangles A–D. Genes (G) including displayed CpG sites are represented as grey rectangles. Ensembl GRCh37 was taken as reference for gene coordinates.

ACKNOWLEDGEMENTS
We are indebted to the patients and their families. We thank all members of the involved Genetic Counseling Cancer Units and Valencian Biobank Network. We thank Lauro Sumoy and Gabriel Rech for their support with methylation array data analysis, Miguel Ángel Peinado and Mireia Jordà for a critical revision of the manuscript. This work was funded by the Spanish Ministry of Economy and Competitiveness and cofunded by FEDER funds -a way to build Europe- (grants SAF2012-33636 and SAF2015-68016-R), Instituto de Salud Carlos III (grant PI16/00563), CIBERONC, RTICC Network (RD12/0036/0031 and RD12/0036/0008), the Spanish Association Against Cancer (080253), the Government of Catalonia (grant 2014SGR336) and Fundación Mutua Madrileña (grant AP114252013). This article is based upon work from COST Action BM1206, supported by COST (European Cooperation in Science and Technology, www.cost.eu). E.D. was supported by a grant from the Spanish Ministry of Economy and Competitiveness.

ADDITIONAL INFORMATION
Supplementary information is available for this paper at https://doi.org/10.1038/s41416-018-0019-8.
MLH1 epimutation is a focal epigenetic event

E Dámás et al.

10. Gylling, A. et al. Large genomic rearrangements and germline epimutations in Lynch syndrome. Int. J. Cancer 124, 2333–2340 (2009).

11. Hitchins, M. P. et al. Inheritance of a cancer-associated MLH1 germ-line epimutation. N Engl. J. Med. 356, 697–705 (2007).

12. Niessen, R. C. et al. Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. Genes Chromosomes Cancer 48, 737–744 (2009).

13. Pineda, M. et al. MLH1 methylation screening is effective in identifying epimutation carriers. Eur. J. Hum. Genet. 20, 1256–1264 (2012).

14. Valle, L. et al. MLH1 germline epimutations in selected patients with early-onset non-polyposis colorectal cancer. Clin. Genet. 71, 232–237 (2007).

15. Goel, A. 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 128, 869–878 (2011).

16. Hitchins, M. et al. MLH1 germline epimutations as a factor in hereditary non-polyposis colorectal cancer. Gastroenterology 129, 1392–1399 (2005).

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

18. Sloane, M. A., Nunez, A. C. & Packham, D. et al. Mosaic epigenetic inheritance as a cause of early-onset colorectal cancer. JAMA Oncol. 1, 953–957 (2015).

19. Hitchins, M. P. Finding the needle in a haystack: identification of cases of Lynch syndrome with MLH1 epimutation. Fam. Cancer 15, 413–422 (2016).

20. Takeda, T. et al. Methylation analysis of DNA mismatch repair genes using DNA derived from the peripheral blood of patients with endometrial cancer: epimutation in endometrial carcinogenesis. Genes 7, 86 (2016).

21. Peltohäuser, P. Update on Lynch syndrome genomics. Fam. Cancer 15, 385–393 (2016).

22. Castillo, A. et al. Prevalence of MLH1 constitutional epimutations as a cause of Lynch syndrome in unselected versus selected consecutive series of patients with colorectal cancer. J. Med. Genet. 52, 498–502 (2015).

23. Crucianelli, F. et al. MLH1 constitutional and somatic methylation in patients with MLH1 negative tumors fulfilling the revised Bethesda criteria. Epigenetics 9, 1431–1438 (2014).

24. Wang, X. et al. Spreading of Alu methylation to the promoter of the MLH1 gene in gastrointestinal cancer. PLoS ONE 6, e25913 (2011).

25. Messegue, X. et al. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics 18, 333–334 (2002).

26. Farré, D. et al. Identification of patterns in biological sequences at the ALGEN server: PROMO and MALGEN. Nucleic Acids Res. 31, 3651–3653 (2003).

27. MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. 42(D1), D986–D992 (2014).

28. Borrias, E. et al. MLH1 founder mutations with moderate penetrance in Spanish Lynch syndrome families. Cancer Res. 70, 7379–7391 (2010).

29. Crépin, M. et al. Evidence of constitutional MLH1 epimutation associated to transgenerational inheritance of cancer susceptibility. Hum. Mutat. 33, 180–188 (2012).

30. Kwok, C.-T., Ward, R. L., Hawkins, N. J. & Hitchins, M. P. 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 (2010).

31. Takahashi, Y. et al. Integration of CpG-free DNA induces de novo methylation of CpG islands in pluripotent stem cells. Science 356, 503–508 (2017).