



HAL
open science

Methylation Tolerance-Based Functional Assay to Assess Variants of Unknown Significance in the MLH1 and MSH2 Genes and Identify Patients With Lynch Syndrome

Delphine Bouvet, Sahra Bodo, Annie Munier, Erell Guillerm, Romane Bertrand, Chrystelle Colas, Alex Duval, Florence Coulet, Martine Muleris

► To cite this version:

Delphine Bouvet, Sahra Bodo, Annie Munier, Erell Guillerm, Romane Bertrand, et al.. Methylation Tolerance-Based Functional Assay to Assess Variants of Unknown Significance in the MLH1 and MSH2 Genes and Identify Patients With Lynch Syndrome. *Gastroenterology*, 2019, 157 (2), pp.421-431. 10.1053/j.gastro.2019.03.071 . hal-02332061

HAL Id: hal-02332061

<https://hal.science/hal-02332061>

Submitted on 31 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Methylation Tolerance-Based Functional Assay to Assess Variants of Unknown Significance in the *MLH1* and *MSH2* Genes and Identify Patients With Lynch Syndrome

Short Title: Characterization of *MLH1* and *MSH2* variants

**BOUVET Delphine^{1,2}, BODO Sahra^{1,2}, MUNIER Annie³, GUILLERM Erell^{1,2,4},
BERTRAND Romane^{1,2}, COLAS Chrystelle^{1,2,5}, DUVAL Alex^{2,6}, COULET Florence^{1,2,4},
MULERIS Martine^{1,2}**

1. Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, CRSA, F-75012 Paris, France

2. SIRIC CURAMUS, Equipe labellisée par la Ligue Nationale contre le Cancer

3. Sorbonne Université, Inserm, Centre de recherche Saint-Antoine, UMS30-LUMIC, Plateforme de Cytométrie en Flux CISA, site Saint-Antoine, F-75012 Paris, France

4. Genetics Department, AP-HP, Hôpital Universitaire Pitié-Salpêtrière, Paris, France

5. Institut Curie, Paris Sciences Lettres Research University, Department of Genetics, Paris, France

6. Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, CRSA, AP-HP, Biochimie, biologie moléculaire, F-75012 Paris, France

Grant support: This work was supported by grants from SIRIC CURAMUS and the Ligue

Nationale Contre le Cancer.

Abbreviations: ACMG: American College of Medical Genetics and Genomics; CRC: colorectal carcinoma; CMMRD: constitutional mismatch repair deficiency; InSiGHT: International Society for Gastrointestinal Hereditary Tumours; LS: Lynch syndrome; MMR: mismatch repair; MNNG: N-Methyl-N'-Nitro-N-Nitrosoguanidine; MT: methylation tolerance; VUS: variant of unknown significance.

Correspondence should be addressed to:

Martine Muleris, Centre de Recherche Saint-Antoine (INSERM U938), Batiment Kourilsky,
2ème étage, 34 rue Crozatier, 75012 Paris France

martine.muleris@inserm.fr

tel: +33 149 286 667

Disclosures: The authors have declared that no conflicts of interest exist.

Author Contributions: study concept and design: MM, DB, AD; acquisition of data: DB; analysis and interpretation of data: DB, MM; drafting of the manuscript: MM, DB, AD; critical revision of the manuscript for important intellectual content: FC, EG, CC; obtained funding; AD; technical support: RB, AM; study supervision: MM.

Acknowledgements

We thank Luisa Riancho for her valuable help with cell sorting.

ABSTRACT

Background & Aims: Approximately 75% of patients with suspected Lynch syndrome carry variants in *MLH1* or *MSH2*—proteins encoded by these genes are required for DNA mismatch repair (MMR). However, 30% of these are variants of unknown significance (VUS). An assay that measures cell response to the cytotoxic effects of a methylating agent can determine the effects of VUS in MMR genes and identify patients with constitutional mismatch repair-deficiency syndrome. We adapted this method to test the effects of VUS in *MLH1* and *MSH2* genes found in patients with suspected Lynch syndrome.

Methods: We transiently expressed *MLH1* or *MSH2* variants in *MLH1*- or *MSH2*-null human colorectal cancer cell lines (HCT116 or LoVo), respectively. The MMR process causes death of cells with methylation-damaged DNA bases, so we measured proportions of cells that undergo death following exposure to the methylating agent; cells that escaped its toxicity were considered to have variants that affect function of the gene product. Using this assay, we analyzed 88 variants (mainly missense variants), comprising a validation set of 40 previously classified variants (19 in *MLH1* and 21 in *MSH2*) and a prospective set of 48 VUS (25 in *MLH1* and 23 in *MSH2*). Prediction scores were calculated for all VUS according to the recommendations of the American College of Medical Genetics and Genomics, based on clinical, somatic, in silico, population and functional data.

Results: The assay correctly classified 39/40 variants in the validation set. The assay identified 12 VUS that did alter function of the gene product and 28 VUS that did not; the remaining 8 VUS had intermediate effects on MMR capacity and could not be classified. Comparison of assay results with prediction scores confirmed the ability of the assay to discriminate VUS that affected the function of the gene products from those that did not.

Conclusions: Using an assay that measures the ability of the cells to undergo death following DNA damage induction by a methylating agent, we were able to assess whether variants in *MLH1* and *MSH2* cause defects in DNA MMR. This assay might be used to help assess the pathogenicity of VUS in *MLH1* and *MSH2* found in patients with suspected Lynch syndrome.

KEY-WORDS: Lynch syndrome, mismatch repair (MMR), variant of unknown significance (VUS), methylation tolerance

INTRODUCTION

Monoallelic germline mutations in one of the 4 human DNA mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*) are responsible for Lynch syndrome (LS), formerly known as hereditary non-polyposis colorectal cancer. LS is one of the most prevalent inherited cancer syndromes in humans and accounts for 1-3% of all colorectal carcinomas (CRC) ¹. MMR is involved in two pathways that are essential for maintaining DNA fidelity, *i.e.* the correction of replication errors that escape proof-reading by DNA polymerases, and the MMR-dependent cell death following specific DNA damage induced especially by methylating agents ²⁻⁶. Therefore, defective MMR function is linked to increased rates of spontaneous somatic mutation, especially in sequences that comprise multiple repeats, thus giving rise to microsatellite instability (MSI) ⁷⁻¹⁰. Defective MMR also leads to increased cell tolerance to DNA lesions induced by methylating agents, so-called methylation tolerance (MT). In LS patients, somatic inactivation of the wildtype allele results in MMR deficiency, leading to MSI which in turn drives carcinogenesis ¹¹. Accordingly, LS patients have a greatly increased risk of developing CRC in their lifetime, as well as extracolonic cancers such as endometrial, ovarian, stomach, pancreatic and multiple other cancers ¹². Analysis for MSI and for loss of expression of MMR proteins in tumor tissue by immunohistochemical staining can identify patients who may have LS. For a definitive diagnosis of LS, genetic testing should reveal a pathogenic germline mutation in a MMR gene. These are found primarily in *MLH1* (40%) and *MSH2* (34%), and to a lesser extent in *MSH6* (18%) and *PMS2* (8%) ¹³. Identification of a high-risk, disease-causing constitutional mutation in the patient then guides the clinical management of their entire family, with implications for genetic counseling and pre-symptomatic surveillance ¹⁴.

In order to discriminate between variants that are pathogenic and those that are not, the International Society for Gastrointestinal Hereditary Tumours (InSiGHT) undertook a collaborative project to develop a five-tiered, standardized classification. This was based on multiple lines of evidence including variant and family characteristics, as well as the results from various functional assays ¹⁵. Unfortunately, variants with unknown significance (VUS) because their pathological impact is unclear represent 31% and 28% of all *MLH1* and *MSH2* mutations, respectively ¹³. VUS represent a major clinical challenge since the lack of classification precludes LS diagnosis for carriers and their relatives. Variants can be classified as VUS for different reasons. Some are VUS because little clinical data is available without

reaching a sufficient level of confidence, while others show contradictory evidence. In these situations, well-established functional assays can provide further evidence to support a final classification especially since they are considered as strong evidence in the classification system. Numerous functional assays in different experimental systems have been developed¹⁶⁻²⁴ among others, however most are based on model organisms such as bacteria, yeast or mice which have intrinsic limitations. Furthermore, most existing *in vitro* or *in vivo* tests measure specific biochemical or restrictive cell biological properties of variant MMR proteins (expression, stability, subcellular location, ability to dimerize, enzymatic activity), but not a global approach to MMR functions as reflected by mutation rate, MSI or drug response. To be considered relevant, functional studies should be performed using an experimental model that mimics as closely as possible the human organism. They should also include appropriate controls to allow extensive validation, as well as a global approach to MMR function using *in vivo* experiments²⁵.

We recently developed a functional assay that has been used successfully to characterize MMR VUS identified in the context of constitutional MMR deficiency syndrome (CMMRD), a rare and severe form of LS caused by bi-allelic germline mutations in MMR genes^{26, 27}. This functional test made it possible to diagnose CMMRD syndrome, especially when sequencing results were not conclusive due to the presence of VUS^{28, 29}. In the present work we have adapted this functional assay, in order to discriminate VUS identified in suspected LS patients. This assay based on the DNA damage-induced apoptosis function of MMR, was evaluated using a large panel of 88 variants, mainly missense variants, and included a validation set of 40 previously classified variants (19 *MLH1* and 21 *MSH2*) and a prospective set of 48 VUS (25 *MLH1* and 23 *MSH2*). VUS were selected to correspond to distinct prediction scores calculated according to the recommendations of the American College of Medical Genetics and Genomics (ACMG) and were based on available clinical, somatic, *in silico*, population and functional data³⁰. As previously shown for CMMRD, we hypothesized that our method would be of most value for classifying LS missense variants and in particular VUS, with the overall aim of improving the clinical management of LS patients.

MATERIALS AND METHODS

Selection of the variants

The present study comprised 44 *MLH1* and 44 *MSH2* variants. *MLH1* variants included 10 known pathogenic mutations (8 non-synonymous missense mutations and 2 nonsense mutations), 9 known neutral alterations (8 non-synonymous and 1 synonymous missense alteration) and 25 VUS that are non-synonymous missense alterations. *MSH2* variants included 10 known pathogenic mutations (9 missense and 1 nonsense mutations), 11 known neutral alterations (9 non-synonymous and 2 synonymous missense variants) and 23 VUS. For all variants, we checked that no effect on splicing was notified on InSiGHT/French databases, or in the absence of functional data by using splicing prediction programs on Alamut software. We also ensured that mutations were scattered throughout the protein structure. Variants already known to be pathogenic or neutral were selected from the InSiGHT and French databases using the following criteria: (i) missense alterations, except 3 nonsense variants used as controls, (ii) an identical or close classification in both databases where possible (*i.e.* class 4 or 5 for pathogenic and class 1 or 2 for neutral variants), and (iii) a relatively high reported frequency. VUS were selected from the InSiGHT and French databases according to the following criteria: (i) missense alterations classified as class 3 in one or preferably both databases, (ii) a relatively high reported frequency, and (iii) with the most available clinical data. In addition, 5 unreported VUS (4 *MLH1* and 1 *MSH2*) detected in suspected-LS patients from the Pitié-Salpêtrière hospital were also screened. Furthermore, prediction scores were calculated for each VUS in order to select variants with different scores.

Prediction scoring

For each VUS, all available data that matched the ACMG guidelines was compiled as follows. Evidence in favor of pathogenicity: functional studies showing a damaging effect (strong evidence); variant that induces an amino acid change at the same position as another pathogenic variant (moderate evidence); absence or low frequency of the variant in controls (GnomAD database) (moderate evidence); and *in silico* prediction favoring a damaging effect (CADD algorithm) (poor evidence). Evidence of benign impact: functional studies showing a neutral effect (strong evidence); significant frequency of the variant in controls (GnomAD database) (strong evidence); *in silico* prediction in favor of a neutral effect (CADD algorithm) (poor evidence); and co-observation of a pathogenic variant in the individual (poor evidence). We developed a scoring method that reflects the weight of each criterion in the rules for combining criteria to classify sequence variants as follows: strong evidence was scored as 6

points, medium evidence as 2 points and poor evidence as 1 point, with positive values if the evidence was in favor of pathogenicity and negative values if the evidence favored a neutral effect.

Chemicals

Cells were exposed to N-Methyl-N'-Nitro-N-Nitrosoguanidine (MNNG) (TCI Europe, Zwijndrecht, Belgium). To exclude differences in MNNG cytotoxicity due to variations in O6-methylguanine methyltransferase enzyme activity, the latter was abrogated by exposure to 20 μ M O6-benzylguanine (Sigma). All chemicals were dissolved in DMSO to a concentration of 20 mM, protected from light and stored at -20°C until used.

Expression vectors

The complete cDNA for the wild type *MLH1* gene (NM_000249.3) and the *mCherry* reporter gene were subcloned into the pIRES vector, placing the expression of the two proteins under the control of the strong CMV promoter. The complete cDNA for the wildtype *MSH2* gene (NM_000251.2), the *GFP* reporter gene and the IRES sequence were subcloned into the pVAX1 vector, placing the expression of the two proteins under the control of the strong CMV promoter. The expression constructs for the mutant variants were produced from the wildtype vector using site-directed mutagenesis and achieved by ProteoGenix (Schiltigheim France).

Cell transfection and cell sorting

HCT116 (*MLH1*-deficient), HCT116-ch3 (*i.e.* HCT116 cell line that had undergone human chromosome 3 transfer which makes the cell line *MLH1*-proficient)³¹ and LoVo (*MSH2*-deficient) cells were grown in a humidified atmosphere containing 5% CO₂ in DMEM GlutaMAX medium (ThermoFisher) supplemented with 10% fetal calf serum. Exponentially growing cells were seeded into 6-well plates in medium supplemented with 1% Penicillin-Streptomycin and 10% fetal calf serum at a density of 4x10⁵ (HCT116) or 5x10⁵ cells/well (LoVo) for 24 hours then transfected with 6 μ L TurboFect Transfection Reagent (ThermoFisher) mixed with 2 μ g plasmid DNA according to the manufacturer's instructions. Medium was removed 6 hours after transfection and fresh medium supplemented with 20 μ M O6-benzylguanine was added. Cell sorting was performed 24 hours post-transfection based on mCherry (for HCT116) or GFP (for LoVo) fluorescence using a MoFlo Astrios Cell Sorter (Beckman Coulter).

Methylation tolerance assay

Following cell sorting, mCherry positive HCT116 cells or GFP positive LoVo cells were seeded into 12-well plates in medium supplemented with 1% Penicillin-Streptomycin, 20% fetal calf serum and 20 μM O6-benzylguanine at a density of 4×10^3 or 10^3 cells/well, respectively. The day after, extemporaneously reconstituted MNNG solution was added at 0.05, 0.1, 0.5 and 2.5 $\mu\text{mol/L}$ final concentration for the *MLH1* training set, and at 0.1 $\mu\text{mol/L}$ final concentration for the *MLH1* VUS. For *MSH2* variants the final MNNG concentration was 0.1, 0.5, 1 and 2 $\mu\text{mol/L}$ for the training set and 1 $\mu\text{mol/L}$ for the VUS. Because of the short half-life of MNNG in aqueous solution (1 hour), the medium was not replaced after drug treatment and 3 rounds of treatment separated by 12 hours were performed. After a total incubation time of 12-14 days, HCT116 colonies were fixed and colored in a solution containing 6% glutaraldehyde and 0.5% crystal violet. LoVo cells were fixed with methanol and stained with a solution of methanol (25%) and crystal violet (0.5%). Colonies were counted and the survival fraction was expressed as the number of colonies in the treated sample relative to the untreated control sample. All samples were tested in triplicate and each experiment was conducted at least in duplicate.

***In silico* analysis**

All *MLH1* and *MSH2* variants tested in this work were subjected to *in silico* analyses using the Sorting Intolerant From Tolerant (SIFT), Align GVGD, Mutation Taster and CADD algorithms.

Statistical analyses

Statistical analyses were performed using GraphPad Prism. All statistical tests were two-sided except for z-score. *P*-values < 0.05 were considered significant.

RESULTS

Proof of feasibility

Complementation of MLH1 gene defect

Briefly, the principle of the assay is as follows: introduction of a *MLH1* variant in HCT116 cells (*MLH1*-deficient) by transient transfection allows direct investigation of its functional

phenotype (functional complementation). To assess only the subset of transfected cells, cells were sorted 24h-post transfection to recover only those expressing mCherry, a reporter gene contained in the expression vector. Cells were then reseeded, exposed to a methylating agent and MMR activity was assessed by a clonogenic assay (Figure 1A). We first checked the kinetics of expression of the transgene to verify the proportion of transfected cells remained stable for up to 72 hours post-transfection, *i.e.* for the time cells were exposed to MNNG treatments (Supplementary Figure 1A). Then, we investigated the feasibility of the experimental procedure by testing whether restoration of cytotoxicity following methylation damage could be detected in *MLH1*-null human CRC HCT116 cells transiently transfected with a plasmid containing wildtype *MLH1* cDNA sequence. As expected, parental HCT116 cells lacking *MLH1* expression survived mild MNNG selection and were tolerant up to 0.5 μ M MNNG. Interestingly, the introduction of wildtype *MLH1* led to a decreased number of colonies that survived selection, indicating restoration of MMR capacity (Figure 1B, Supplementary Figure 1B). Cells were about 7-fold more sensitive to MNNG than controls (*i.e.* HCT116 cells transfected with a random sequence, LD50=0.25 vs 1.85 following 1 MNNG pulse, respectively) (Supplementary Figure 2A).

Since the literature indicates methylation tolerance is higher than the 7-fold difference observed in our assay, we then confirmed the MT assay by using as a control the HCT116 cell line that had undergone human chromosome 3 transfer (HCT116-ch3), thus making the cells MMR proficient. As expected, we observed that methylation tolerance was up to 62-fold higher in this cell line compared to HCT116 cells that were transiently transfected with the random cDNA (LD50=0.03 vs LD50=1.85 following 1 MNNG pulse, respectively) (Figure 1B, Supplementary figure 2A).

Because transfection is likely to result in multiple copies of the gene in the recipient cells, which could affect their response to the drug, we also addressed this issue by separating cells that expressed low or high fluorescence intensity (Supplementary Figure 3A). This can be considered as an approximate stratification of the cells according to their protein expression and hence copy number. HCT116 cells transfected with *MLH1* or with a random cDNA construct were sorted according to this procedure before screening with the MT assay. No significant differences in response were observed between transfected cells that expressed the *MLH1* plasmid either weakly or strongly (LD50=0.11 vs 0.15 following 1 pulse with MNNG, respectively) or between cells transfected with the random cDNA construct (LD50=0.75 vs 1.14, respectively, P=0.17) (Supplementary Figures 3B and 3C). Therefore, using our

experimental conditions we concluded the level of expression of the transgene had no effect on the cellular response to the methylating agent.

Complementation of *MSH2* gene defect

The human CRC cell line LoVo containing a large *MSH2* homozygous deletion³² was used to assess *MSH2* variants. As expected, LoVo cells survived mild MNNG selection. When transiently transfected with a plasmid containing wildtype *MSH2* cDNA sequence, LoVo cells were about 4-fold more sensitive than LoVo cells transfected with a random sequence (LD50=0.41 vs 1.75 following 1 MNNG pulse, respectively) (Figure 1B and Supplementary Figure 2B).

Validation study

Standardization of the MT assay for *MLH1* variants

We next sought to identify experimental conditions that would best discriminate 10 previously proven pathogenic *MLH1* variants from 9 previously proven non-pathogenic *MLH1* variants. The former consisted of 8 missense and 2 nonsense variants classified as class 5 in the InSiGHT database, while the latter consisted of 8 non-synonymous and 1 synonymous missense variants classified as class 1 in this database (Supplementary Table 1). Following transfection and flow cytometry, sorted cells were reseeded and exposed to a range of MNNG concentrations and to different numbers of treatments. The conditions that allowed the best discrimination of the 2 groups of variants, *i.e.* 2 and 3 pulses at 0.1 μ M MNNG (Figure 2), were used to calculate a mean survival score.

Standardization of the MT assay for *MSH2* variants

The case-control study consisted of 10 previously proven pathogenic *MSH2* variants and 11 previously proven neutral or likely neutral *MSH2* variants. The former group comprised 9 missense and 1 nonsense variants classified as class 4 or class 5 in the InSiGHT database, while the latter group comprised 9 non-synonymous and 2 synonymous missense variants classified as class 1 or class 2 in the InSiGHT or French (Unicancer) Databases³³ (Supplementary Table 1). Following the same experimental approach than that developed for *MLH1* variants, it was determined that the 2 sets of *MSH2* variants were best discriminated using 1 or 2 treatments with 1 μ M MNNG (Figure 3).

MT assay results for MLH1 and MSH2 variants

Using the above experimental conditions to calculate mean survival scores, all pathogenic *MLH1* variants displayed a mean survival score above 82.2%, while all non-pathogenic *MLH1* variants displayed a mean survival score below 55.2% (Figure 4A). A Receiver Operating Characteristic (ROC) curve was constructed assuming a binary classification of the variants and showed the assay yielded perfect discrimination between pathogenic and neutral variants, with 100% sensitivity (69.15-100) and 100% specificity (71.5-100) for a cut-off value of 68.7%. Moreover, all pathogenic *MSH2* variants displayed a mean survival score above 58.90%, whereas the mean survival score of all but one non-pathogenic *MSH2* variant was below 32.2% (Figure 4B). The ROC curve showed 91.67% sensitivity (61.5-99.8) and 100% specificity (71.5-100) with a cut-off value of 45.54%.

Prospective study and comparison with previous available data

Clinical, somatic, *in silico*, population and functional data for the *MLH1* and *MSH2* VUS was retrieved from public databases (Supplementary Tables 2 and 3). For each variant, this information was used to calculate a prediction score in accordance with ACMG recommendations³⁰ and representing the level of evidence in favor of (positive score) or against (negative score) pathogenicity. VUS with different prediction scores were selected for testing with our MT assay.

MT assay results for *MLH1* VUS

The functional assay conditions described above were used to evaluate 25 *MLH1* missense VUS, comprising 21 variants classified as class 3 in the InSiGHT database and 4 unpublished variants identified in suspected LS patients and classified as VUS according to ACMG guidelines (unpublished data) (Supplementary Table 2). The position of all alterations in the *MLH1* protein structure is presented in Supplementary Figure 4A. Unlike the clear differences seen between known pathogenic and non-pathogenic variants, the VUS exhibited a large range of responses to methylation damage (Figure 4A). Two distinct groups were observed, with 6 VUS showing a mean survival score above 82.2% and 13 VUS showing a score below 55.2%. Thus, according to our validation set of variants, these VUS were considered as damaging or neutral, respectively. The mean survival scores of the 6 remaining VUS fell in between these values (*i.e.* 55.2-88.2%) and were significantly different from the reference average (P=0.007 for L85P; P=0.003 for L582F and L622P; P=0.004 for L260F; P=0.02 for H264Y; P=0.03 for V179D; z-score analysis). Using the cut-off value previously

determined by the ROC curve, the first 4 VUS could be classified as “potentially damaging” and the last 2 as “potentially neutral”.

Comparison with previous data for *MLH1* VUS

For 6 VUS (marked with an arrow in Figure 5A), there was moderate evidence in favor of pathogenicity (prediction score $\geq +9$) or against (prediction score ≤ -6) prior to the results of the MT assay, based on previous functional data (see Supplementary Table 2). Interestingly, for all these VUS our MT results were concordant with previous data. A neutral phenotype was indicated for variants A160V, R474Q and E523D and all 3 were classified as non-damaging for MMR in our screen. In contrast, a pathogenic phenotype was indicated for variants G244V, P640L and L582F, with the first 2 classified as damaging and the third as potentially damaging in our screen. Although the MT assay does not confer extra evidence to classify variants according to current ACMG guidelines³⁰, these results support the relevance of our functional assay. For most variants (19/25) there was either little evidence in favor of pathogenicity or contradictory evidence prior to the MT assay ($-6 \leq$ prediction score ≤ 5). For 7 of these variants (L260F, L622P, A539D, W666R, A586D, S106R and L85P), the assay results favor pathogenicity, whereas for the 12 other variants they indicate no pathogenicity (T116K, V179D, Y126N, I565T, E319K, R474W, R474G, N215S, H264Y, N338S, G454R and K618E). The MT assay was therefore clearly able to distinguish this set of variants.

Methylation tolerance assay results for *MSH2* VUS

We next tested a set of 23 *MSH2* missense VUS comprising 22 variants previously classified as class 3 in the InSiGHT or French databases and one new, unpublished variant (Supplementary Table 3). The position of all mutations in the *MSH2* protein is presented in Supplementary Figure 4B. Two distinct groups were observed (Figure 4B), with 6 VUS showing a mean survival score above 58.9% and 15 VUS showing a score below 32.2%. These were considered as pathogenic or neutral, respectively. The response of the 2 remaining VUS fell in between and differed significantly from the reference average ($p=0.017$ for L736P and $p=0.006$ for R711Q; z-score analysis). Using the cut-off value determined by the ROC curve, they could be identified as “potentially damaging” for MMR.

Comparison with previous data for *MSH2* VUS

Only the A305T *MSH2* VUS has been previously assessed with functional assays and this is marked with an arrow in Figure 5B. There was moderate evidence against pathogenicity

(prediction score = -7) for this VUS and in agreement the MT assay strongly indicated a benign impact. For most variants (22/23) there were either little evidence in favor of pathogenicity or else contradictory evidence ($-1 \leq \text{prediction score} \leq 5$) prior to the MT assay. Our assay discriminates these VUS into MMR abrogating (P622T, V470E, L602P, N671Y, A500P, R638G, L736P and R711Q) or neutral (Y856C, G548C, R243Q, E561K, K627N, H610N, Y43C, N583S, T33P, T905I, M152T, V342I, A2T and P349A).

DISCUSSION

Based on our previous successful approach to functionally discriminate MMR VUS in CMMRD syndrome ²⁸, we hypothesized here that MT could also be used to classify LS variants. In a direct comparison with 40 variants previously characterized for pathogenicity, our assay correctly classified 39 (97.5%). These results clearly outperform other currently employed tools such as *in silico* prediction tests (Supplementary Figure 5). The only discordant result was for the *MSH2* (c.593A>G, E198G) mutation, which was classified as neutral in the InSiGHT database but pathogenic in our screen. This mutation was identified in a Swedish family with five affected members (4 siblings and their father) and with a missense *MLH3* (T3826C) variant also present in all 4 siblings ^{34, 35}. The *MSH2* variant segregated with CRC diagnosis at ages 29, 40, 62 and 62 years in the family, and was not found in control subjects. In agreement with our results, previous functional assays also favored pathogenicity for this variant *i.e.* no MMR, defective interaction with all MSH2 partners, and low protein expression in yeast assays ^{17, 36}. However, the CRCs did not show MSI. Therefore, the classification of this *MSH2* variant as neutral is debatable.

The initial results obtained with the MT assay led us to propose a functional and robust test to evaluate the impairment of MMR in confirmed LS patients. When applied to a prospective set of 48 *MLH1* or *MSH2* VUS, 40 (83%) could be classified unambiguously as damaging or neutral for MMR, whereas 8 showed an intermediate capacity to restore MMR and therefore could not be definitively classified. The latter may correspond to weak variants with attenuated (but not abrogated) DNA damage response capacity that confer only a moderate risk of cancer, as suggested for other genes ^{37, 38} or pathologies ³⁹. This hypothesis remains to be tested in the context of LS. We next compared the results of the MT assay with the

calculated prediction scores that reflect the level of evidence for classification of VUS based on all previously available data. For all the VUS in which previous functional data was available, the results of the MT assay were in agreement, thus providing further confirmation of the relevance of our assay. However for most VUS, there was either little previous evidence for pathogenicity or else contradictory evidence. In this subset, our assay makes it possible to discriminate VUS that are damaging from those that are neutral. Interestingly, all VUS with previous evidence favoring a benign effect were also found to be neutral in the MT assay.

The HCT116-ch3 cell line demonstrated a methylation tolerance that was much higher than HCT116 cells transfected with *MLH1* cDNA. This suggests that MLH1 expression might be low in our assay using transient transfection. However, the expression of MLH1 immediately following transfection was found to be much higher than in the HCT116-ch3 cell line (results not shown). Although we have no definitive explanation for the difference observed in MT, one can speculate a possible toxic⁴⁰ or inhibitory⁴¹ effect of over-expressed *MLH1* gene, or a decreased MLH1 expression in subsequent cell divisions 48h post-transfection (*i.e.* during MNNG treatments). The latter situation could lead to a heterogeneous cell population that includes cells with low MLH1 expression. This would result in incomplete MT response when using transient transfection compared to the complete MT phenotype observed when using the HCT116-ch3 cell line.

Our study has several limitations. Firstly, the expression vector used here may not be suitable for assessing the pathogenicity of alleles whose product is functionally equivalent to wildtype, but which is expressed at a lower level. Indeed, the variant is placed under the control of a strong promoter resulting in very high, non-physiological expression, similar to most assays that use transfection models. Secondly, in contrast with CMMRD syndrome²⁸, variants here were evaluated outside the genetic background of the patient and therefore any potential additive effects from several variants could be missed. Thirdly, our test assessed one of the 2 MMR functions. Although both functions (*i.e.* mismatch repair and DNA damage-induced cell death) were found to be linked in the vast majority of MMR mutants explored so far, examples of decoupling have been reported for the Msh2^{G674A} and Msh6^{T1217D} mutations in mice^{42,43} and in MSH-deficient yeast strain models⁴⁴.

We have developed an *in vivo* functional screen that assesses MMR function in a human cell expression model. The MT test has been extensively validated in a case-control cohort for the two main MMR genes responsible for LS, namely *MLH1* and *MSH2*. Overall, the robust read-out demonstrated 95% sensitivity and 100% specificity in classifying MMR variants as pathogenic or neutral. In addition, it allowed functional discrimination of MMR VUS beyond that of previous methods, being able to discriminate 83% of the VUS tested. This assay is simple, reliable, cost-effective and rapid, being performed in 14 to 16 days. Another major advantage of our screen is that patient-derived material is not required.

MMR genes are included in the ACMG list for reporting of incidental findings in clinical exome and genome sequencing⁴⁵ and consequently any methods that help to classify VUS in these genes are particularly relevant. It is more difficult to classify variants identified by sequencing (outside of the known clinical phenotype) without data from functional assays. Functional assays can provide strong evidence for assigning the likely functional impact of a VUS in the MMR gene classification system. However, such assays are not used routinely in the risk assessment of MMR VUS, notably because the reproducibility and performance of almost all current MMR functional assays remains to be determined. To our knowledge, the present study includes the largest case-control cohort evaluated by a functional assay, allowing us to demonstrate its very high sensitivity and specificity. Therefore, we propose that each new *MLH1* or *MSH2* missense variant identified in a suspected LS patient be screened using the MT assay in a prospective manner. A study consortium will be required to independently validate our method using an additional panel of variants with known neutral or pathogenic effect, allowing implementation of this assay in a comprehensive diagnostic procedure for LS-associated *MLH1* or *MSH2* missense variants.

FIGURES

FIGURE 1. Functional screen for the characterization of *MLH1* or *MSH2* variants

A. Schematic diagram. *MLH1* (or *MSH2*) variants were transiently introduced into the *MLH1*-null (or *MSH2*-null) human colorectal cancer cell line HCT116 (or LoVo). Cells were sorted 24 hours-post transfection to recover only those expressing mCherry (or GFP), a reporter gene contained in the expression vector. Cells were then reseeded and exposed to MNNG (N-methyl-N'-nitro-N-nitroso-guanidine), a methylating agent that is highly toxic to MMR-

proficient cells, but only moderately toxic to MMR-deficient cells. Clonogenic survival was used as readout for MMR activity. Variants that confer restoration of cytotoxicity to methylation damage were assumed to be neutral, whereas variants unable to restore apoptosis were assumed to be damaging. All samples were tested in triplicate and each experiment was conducted at least twice independently.

B. Complementation of MMR gene defect. Following transient transfection of HCT116 cells with either the *MLH1*-mCherry or control-mCherry construct (left panels), or LoVo cells with either the *MSH2*-GFP or control-GFP construct (right panels), the transfected cells were sorted, reseeded and exposed to 1, 2, or 3 pulses of MNNG treatment for 48 hours, 60 hours or 72 hours post-transfection. A clonogenic assay was then carried out. Whereas control cells (*i.e.* cells transfected with random construct, in red) survived mild MNNG selection, cells transfected with wildtype *MLH1* (or *MSH2*) containing construct (in green) reversed the MT phenotype, indicating restoration of the DNA damage sensing function associated with MMR. Response of the non-transfected parental HCT116 (or LoVo) cells (in dark blue) or HCT116-ch3 cells (in light blue), as well as that of the HCT116-ch3 cells transfected with the control-random construct (in pink) is shown for comparison. Data represent means \pm SD from at least 2 independent clonogenic experiments, each performed in triplicate.

FIGURE 2. Validation study for *MLH1* variants

Functional screening of a series of proven pathogenic (N=10) and neutral (N=9) *MLH1* variants was performed using increasing concentrations of MNNG. Because of the short half-life of MNNG in aqueous solution, 1, 2, or 3 pulses of MNNG treatment were performed. The best experimental conditions to discriminate pathogenic from neutral variants are indicated (red box). Screen results with the control constructs used in the proof of feasibility step (*i.e.* wildtype *MLH1* in the neutral cluster or random cDNA in the pathogenic cluster) are represented by triangles. All samples were tested in triplicate and each experiment was conducted at least twice independently. Student's t-test was used to compare the means.

FIGURE 3. Validation study for *MSH2* variants

Functional screening of a series of proven pathogenic (N=10) and neutral (N=11) *MSH2* variants was performed using increasing concentrations of MNNG. One, 2, or 3 pulses of MNNG treatment were used. The best experimental conditions to discriminate pathogenic from neutral variants are indicated (red box). Screen results with the control constructs used in the proof of feasibility step (*i.e.* wildtype *MSH2* in the neutral cluster or random cDNA in

the pathogenic cluster) are represented by triangles. All samples were tested in triplicate and each experiment was conducted at least twice independently. Student's t-test was used to compare the means.

FIGURE 4. Functional MT screen for the detection of damaging *MLH1* and *MSH2* variants (validation and prospective sets)

A. A total of 44 *MLH1* variants including 10 proven pathogenic variants and 9 proven neutral variants (validation set), as well as 25 VUS (prospective set) were tested using the functional screen. The mean survival score represents the average survival fraction after 2 and 3 treatments with 0.1 μ M MNNG.

B. A total of 44 *MSH2* variants including 10 proven pathogenic variants and 11 proven neutral variants (validation set), as well as 23 VUS (prospective set) were tested using the functional screen. The mean survival score represents the average survival fraction after 1 and 2 treatments with 1 μ M MNNG.

The grey area corresponds to the 5% values that differed significantly from the average as defined by the z-score. The cut-off value as determined by the ROC curve in the validation set is represented by the dotted line.

FIGURE 5. Evidence for or against pathogenicity for the 48 VUS screened with the MT assay

For each *MLH1* (**A**) or *MSH2* (**B**) VUS, all available data was assigned as strong, moderate or poor evidence in favor of either pathogenicity or a benign impact, according to the recommendations of the American College of Medical Genetics and Genomics (ACMG)³⁰. For graphic representation, a scoring system was established that reflects the weight of each criterion in the rules used for combining criteria to classify sequence variants. *i.e.* +6 points for strong, +2 points for moderate and +1 point for poor evidence of pathogenicity; -6 points for strong and -1 point for poor evidence of a benign impact.

Light blue bars represent the level of evidence in favor of pathogenicity (positive values) or in favor of a benign impact (negative values) for each VUS based on all previously available data. Dark blue bars indicate the level of evidence in favor of pathogenicity (positive values) or in favor of a benign impact (negative values) for each VUS conferred by the results of the MT assay. For the 7 VUS marked with an arrow, there was moderate evidence in favor of or

against pathogenicity prior to the MT assay, due to the existence of previous functional data. Although the MT assay results supported the published data for all of these 7 VUS, it did not confer additional points according to the ACMG classification system. Consequently, the dark blue bar is shown next to the light blue bar for these 7 variants. VUS in italics are classified using the MT assay as potentially damaging or potentially neutral.

REFERENCES

1. Lynch HT, Snyder CL, Shaw TG, et al. Milestones of Lynch syndrome: 1895-2015. *Nat Rev Cancer* 2015;15:181-94.
2. Karran P, Stephenson C. Mismatch binding proteins and tolerance to alkylating agents in human cells. *Mutat Res* 1990;236:269-75.
3. Karran P, Bignami M. Self-destruction and tolerance in resistance of mammalian cells to alkylation damage. *Nucleic Acids Res* 1992;20:2933-40.
4. Hawn MT, Umar A, Carethers JM, et al. Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. *Cancer Res* 1995;55:3721-5.
5. Kat A, Thilly WG, Fang WH, et al. An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair. *Proc Natl Acad Sci U S A* 1993;90:6424-8.
6. Carethers JM, Hawn MT, Chauhan DP, et al. Competency in mismatch repair prohibits clonal expansion of cancer cells treated with N-methyl-N'-nitro-N-nitrosoguanidine. *J Clin Invest* 1996;98:199-206.
7. Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 1993;363:558-61.
8. **Aaltonen LA, Peltomaki P**, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. *Science* 1993;260:812-6.
9. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science* 1993;260:816-9.
10. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998;58:5248-57.
11. Peltomaki P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* 2003;21:1174-9.
12. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med* 2003;348:919-32.
13. Peltomaki P. Update on Lynch syndrome genomics. *Fam Cancer* 2016;15:385-93.
14. Vasen HF, Blanco I, Aktan-Collan K, et al. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. *Gut* 2013;62:812-23.

15. Thompson BA, Spurdle AB, Plazzer JP, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet* 2014;46:107-115.
16. Shimodaira H, Filosi N, Shibata H, et al. Functional analysis of human MLH1 mutations in *Saccharomyces cerevisiae*. *Nat Genet* 1998;19:384-9.
17. Martinez SL, Kolodner RD. Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions. *Proc Natl Acad Sci U S A* 2010;107:5070-5.
18. Drotschmann K, Clark AB, Kunkel TA. Mutator phenotypes of common polymorphisms and missense mutations in MSH2. *Curr Biol* 1999;9:907-10.
19. Blasi MF, Ventura I, Aquilina G, et al. A human cell-based assay to evaluate the effects of alterations in the MLH1 mismatch repair gene. *Cancer Res* 2006;66:9036-44.
20. Trojan J, Zeuzem S, Randolph A, et al. Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology* 2002;122:211-9.
21. Ollila S, Dermadi Bebek D, Jiricny J, et al. Mechanisms of pathogenicity in human MSH2 missense mutants. *Hum Mutat* 2008;29:1355-63.
22. Kariola R, Raevaara TE, Lonnqvist KE, et al. Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. *Hum Mol Genet* 2002;11:1303-10.
23. Andersen SD, Liberti SE, Lutzen A, et al. Functional characterization of MLH1 missense variants identified in Lynch syndrome patients. *Hum Mutat* 2012;33:1647-55.
24. Houlleberghs H, Dekker M, Lantermans H, et al. Oligonucleotide-directed mutagenesis screen to identify pathogenic Lynch syndrome-associated MSH2 DNA mismatch repair gene variants. *Proc Natl Acad Sci U S A* 2016;113:4128-33.
25. Nykamp K, Anderson M, Powers M, et al. Sherlock: a comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med* 2017;19:1105-1117.
26. **Ricciardone MD, Ozcelik T**, Cevher B, et al. Human MLH1 deficiency predisposes to hematological malignancy and neurofibromatosis type 1. *Cancer Res* 1999;59:290-3.
27. Wang Q, Lasset C, Desseigne F, et al. Neurofibromatosis and early onset of cancers in hMLH1-deficient children. *Cancer Res* 1999;59:294-7.
28. **Bodo S, Colas C, Buhard O**, et al. Diagnosis of Constitutional Mismatch Repair-Deficiency Syndrome Based on Microsatellite Instability and Lymphocyte Tolerance to Methylating Agents. *Gastroenterology* 2015;149:1017-1029 e3.
29. Taeubner J, Wimmer K, Muleris M, et al. Diagnostic challenges in a child with early onset desmoplastic medulloblastoma and homozygous variants in MSH2 and MSH6. *Eur J Hum Genet* 2018;26:440-444.
30. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405-24.
31. Jacob S, Aguado M, Fallik D, et al. The role of the DNA mismatch repair system in the cytotoxicity of the topoisomerase inhibitors camptothecin and etoposide to human colorectal cancer cells. *Cancer Res* 2001;61:6555-62.
32. Umar A, Boyer JC, Thomas DC, et al. Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. *J Biol Chem* 1994;269:14367-70.

33. Grandval P, Fabre AJ, Gaildrat P, et al. UMD-MLH1/MSH2/MSH6 databases: description and analysis of genetic variations in French Lynch syndrome families. *Database (Oxford)* 2013;2013:bat036.
34. **Wahlberg S, Liu T**, Lindblom P, et al. Various mutation screening techniques in the DNA mismatch repair genes hMSH2 and hMLH1. *Genet Test* 1999;3:259-64.
35. **Liu HX, Zhou XL**, Liu T, et al. The role of hMLH3 in familial colorectal cancer. *Cancer Res* 2003;63:1894-9.
36. Gammie AE, Erdeniz N, Beaver J, et al. Functional characterization of pathogenic human MSH2 missense mutations in *Saccharomyces cerevisiae*. *Genetics* 2007;177:707-21.
37. Moghadasi S, Meeks HD, Vreeswijk MP, et al. The BRCA1 c. 5096G>A p.Arg1699Gln (R1699Q) intermediate risk variant: breast and ovarian cancer risk estimation and recommendations for clinical management from the ENIGMA consortium. *J Med Genet* 2018;55:15-20.
38. Zerdoumi Y, Lanos R, Raad S, et al. Germline TP53 mutations result into a constitutive defect of p53 DNA binding and transcriptional response to DNA damage. *Hum Mol Genet* 2017;26:2812.
39. Flannick J, Johansson S, Njolstad PR. Common and rare forms of diabetes mellitus: towards a continuum of diabetes subtypes. *Nat Rev Endocrinol* 2016;12:394-406.
40. Zhang H, Richards B, Wilson T, et al. Apoptosis induced by overexpression of hMSH2 or hMLH1. *Cancer Res* 1999;59:3021-7.
41. Shcherbakova PV, Hall MC, Lewis MS, et al. Inactivation of DNA mismatch repair by increased expression of yeast MLH1. *Mol Cell Biol* 2001;21:940-51.
42. **Yang G**, Scherer SJ, Shell SS, et al. Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. *Cancer Cell* 2004;6:139-50.
43. **Lin DP, Wang Y, Scherer SJ**, et al. An Msh2 point mutation uncouples DNA mismatch repair and apoptosis. *Cancer Res* 2004;64:517-22.
44. Drotschmann K, Topping RP, Clodfelter JE, et al. Mutations in the nucleotide-binding domain of MutS homologs uncouple cell death from cell survival. *DNA Repair (Amst)* 2004;3:729-42.
45. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 2013;15:565-74.

Author names in bold designate shared co-first authorship