



**HAL**  
open science

# Syndrome de Lynch : mise au point d'un test diagnostique basé sur la caractérisation fonctionnelle de tolérance à la méthylation des variants des gènes du système de réparation des mésappariements

Delphine Bouvet

## ► To cite this version:

Delphine Bouvet. Syndrome de Lynch : mise au point d'un test diagnostique basé sur la caractérisation fonctionnelle de tolérance à la méthylation des variants des gènes du système de réparation des mésappariements. Génétique humaine. Sorbonne Université, 2018. Français. NNT : 2018SORUS527 . tel-02946934

**HAL Id: tel-02946934**

**<https://theses.hal.science/tel-02946934>**

Submitted on 23 Sep 2020

**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.



## SORBONNE UNIVERSITÉ

Ecole Doctorale Physiologie, Physiopathologie et Thérapeutique (ED394)

Centre de Recherche Saint Antoine UMR\_S 938, INSERM, UPMC

Equipe « Instabilité des Microsatellites et Cancer »

### THÈSE DE DOCTORAT

Spécialité Génétique des Cancers

Syndrome de Lynch : mise au point d'un test diagnostique basé sur la caractérisation fonctionnelle de tolérance à la méthylation des variants des gènes du système de réparation des mésappariements

Présentée par

**Delphine BOUVET**

Dirigée par le Docteur Martine MULERIS

Présentée et soutenue publiquement le 30 novembre 2018

Devant un jury composé de :

Philippe LE ROUZIC  
Marie-Pierre BUISINE  
Dominique LEROUX  
Christine BELLANNÉ-CHANTELOT  
Valérie BORDE  
Martine MULERIS  
Alex DUVAL

Président  
Rapporteuse  
Rapporteuse  
Examinatrice  
Examinatrice  
Directrice de thèse  
Directeur du laboratoire

## Remerciements

Mes premiers remerciements vont aux membres de mon jury de thèse : **Philippe LE ROUZIC**, pour avoir accepté d'être le président de ce jury ; **Marie-Pierre BUISINE** et **Dominique LEROUX** pour avoir étudié et rapporté mon manuscrit avec intérêt et **Valérie BORDE** et **Christine BELLANNÉ-CHANTELOT** pour avoir participé à ma soutenance.

Je remercie **Alex DUVAL** de m'avoir accueilli dans son laboratoire, mais également pour les moyens financiers déployés afin de mener à bien mon projet et me permettre de finaliser ma thèse dans les meilleures conditions.

Je tiens à adresser un merci particulier à **Martine MULERIS**, ma directrice de thèse. Ces années passées ensemble ont été épanouissantes et enrichissantes et c'est en grande partie à toi que je le dois. Merci pour tes conseils, ta disponibilité, ta discrétion et toutes tes autres qualités qui font de toi une super cheffe. Merci pour toutes ces fois où nous avons fermé derrière nous la porte de ton bureau pour parler de boulot, mais pas seulement.

Je remercie l'**Institut Universitaire en Cancérologie**, qui a financé mon doctorat.

Je remercie le service d'anatomopathologie et en particulier **Magali SVRCEK**, qui m'a notamment initié à l'art ésotérique de la lecture des lames d'histologie.

Merci à **Florence COULET** et **Erell GUILLERM**, de l'U.F. d'Oncogénétique et Angiogénétique Moléculaire, pour votre intérêt pour mon travail ; je vous remercie aussi pour toutes les informations techniques et cliniques que vous m'avez apprises.

Un merci plein de tendresse pour **Romane BERTRAND**, mon coup de foudre professionnel et bien sûr amical. Tu es une technicienne consciencieuse et passionnée et il y a toujours quelque chose à apprendre auprès de toi : merci d'avoir fait découvrir le monde des protéines à l'afficionado de génétique que je suis. Merci pour tous les repiquages, les lavages de membranes et autres que tu as pu faire pour moi ainsi que pour tes conseils techniques. Je suis heureuse que nous ayons partagé nos pauses-déjeuner pendant ces 3 années et notre bureau pendant un bout de temps. Merci d'être venue vers moi, quand j'ai tellement de mal à aller vers les autres : tu as su m'appivoiser et je t'en suis reconnaissante. Merci pour cette amitié qui a commencé au laboratoire et qui continuera longtemps après.

Merci à **A'dem BOKHARI** et **Malorie GREENE**, duo de thésards de chic et de choc, qui m'ont montré la voie. A'dem, je te remercie pour ton humeur et ta bonne humeur, mais pas pour ce surnom qui m'aura suivi jusqu'à la fin de ma thèse ! Malorie, tu es pour moi un exemple d'optimisme et de gentillesse : merci de m'avoir montré qu'il y a toujours une bonne raison de sourire en se levant le matin.

Merci à nos stagiaires de cette année, qui ont ramené un peu de fraîcheur dans l'équipe. **Kathleen NOEL** et **Toky RATOVOMANANA** : encore félicitations pour vos bourses de thèse, je vous souhaite bonne continuation dans vos vies privées et professionnelles. **Samuel LANDMAN**, un merci tout particulier car c'est un peu grâce à toi que j'ai décroché mon premier emploi post-thèse !

**Annie MUNIER**, merci pour toutes ces heures passées à se raconter nos vies devant le trieur en regardant les cellules passer. On en aura passé des « minutes de suspens » à attendre que les cellules montent dans les tubulures !

Merci à tous les autres membres de l'équipe MSI qui font vivre le laboratoire tous les jours : **Ada, Vincent, Olivier** (et **Kristell**), **Mouna, Marie-Annick**. Merci également à **Paul-Patrice**, qui a toujours le bon mot pour me mettre de bonne humeur à mon arrivée au CdR !

Merci à **Pauline** et **Damien, Alexia** et **Bastien, Loïc** et **Emilie**, de me rappeler qu'il se passe aussi de jolies choses en dehors d'un laboratoire de recherche. Les années que nous sommes en train de vivre sont celles où nous décidons quelles personnes nous voulons devenir, tant personnellement que professionnellement ; et je suis heureuse de les partager avec vous.

**Maman**, merci pour tout, évidemment. Tu es mon modèle de femme : forte, courageuse et indépendante. Je sais que j'ai hérité de beaucoup de tes défauts, j'espère aussi avoir hérité de certaines de tes qualités. **Cécile**, ma petite sœur : je sais que je pourrai me tourner vers toi en cas de coup dur, toujours, et je te remercie pour ça.

Merci à mes grands-parents qui nous ont permis de réaliser nos études sans contraintes. **Mamie**, merci particulièrement d'avoir accepté de tourner une page de ta vie pour nous laisser écrire une page de la nôtre. **Papy**, j'espère que de là où tu es, tu es un peu fier de moi.

*Last but not least*, merci à mon mari, **Baptiste**. D'abord, merci de croire en moi depuis le début de la décennie que nous avons déjà passée ensemble. Merci aussi de m'avoir convaincue que j'avais les compétences nécessaires pour mener à bien une thèse. Merci surtout d'avoir accepté sans trop broncher mes horaires de travail parfois un peu exotiques, de m'avoir préparé le petit déjeuner quel que soit l'heure du réveil, de m'avoir convaincu que c'était largement faisable de mener de front une thèse et une organisation de mariage. Enfin, merci d'avance pour toutes les prochaines décennies... Ensemble toi et moi, rien ne nous arrêtera.

## Résumé

Le système de réparation des mésappariements (MMR pour *MisMatch Repair*) participe à la signalisation apoptotique des lésions de méthylation sur l'ADN. En conséquence, l'inactivation du système MMR rend les cellules tolérantes à ces lésions (phénotype de tolérance à la méthylation) et leur permet d'échapper à l'apoptose. Le syndrome de Lynch (LS) est la première cause de forme familiale de cancer colorectal. Il est secondaire à des mutations germinales mono-alléliques affectant les gènes du système MMR. Le diagnostic de ce syndrome repose aujourd'hui sur le séquençage et la détection de mutations MMR délétères germinales chez les malades. Cette approche diagnostique pose le problème de l'interprétation fonctionnelle des variants identifiés, qui est souvent équivoque, puisque de telles modifications de séquences peuvent être ou non pathogènes. Les variants de signification incertaine (VSI) représentent 30% des altérations détectées dans les gènes *MLH1* et *MSH2* (gènes les plus couramment altérés dans le LS). Cette problématique est majeure en clinique sur un plan diagnostique et va s'incrémenter avec l'utilisation en routine des techniques de séquençage à haut débit. Notre objectif était de développer un test fonctionnel simple permettant de discriminer les VSI pathogènes et non pathogènes identifiés chez les patients suspectés d'être atteints du LS.

Dans la continuité de travaux récemment publiés par l'équipe, notre stratégie expérimentale repose sur la capacité de variants *MLH1* et *MSH2* à compléter la déficience MMR d'une lignée cellulaire en culture. En bref, une restauration de la sensibilité à la méthylation est attendue si le variant transfecté est non pathogène, alors qu'une conservation de la tolérance à la méthylation indique que le variant est délétère. Nous avons d'abord testé un panel de 40 variants *MLH1* et *MSH2* dont le caractère pathogène ou neutre est connu. Notre test a permis de classer correctement 39/40 variants, montrant la pertinence de notre approche méthodologique. La technique a ensuite été appliquée à une série de 49 VSI (25 *MLH1* et 24 *MSH2*) à caractériser. Globalement, notre technique a permis de proposer une classification pour 41/49 (83%) des VSI : 12 VSI (6 *MLH1* et 6 *MSH2*) sont considérés comme compétents et 29 VSI (13 *MLH1* et 16 *MSH2*) sont considérés comme déficients. La classification des 8 VSI (6 *MLH1* et 2 *MSH2*) restants est incertaine. Les autres données disponibles (données cliniques, somatiques, *in silico* et éventuellement fonctionnelles) confirment nos résultats et ont permis de valider notre approche.

À terme, ce test de signalisation des dommages à l'ADN pourrait être disponible en routine sur une plateforme de diagnostic afin d'étudier en prospectif tous les VSI détectés chez des patients suspectés d'être atteints du syndrome de Lynch.

## Abstract

The MMR system participates in the apoptotic signaling of methylation lesions on DNA. As a consequence, the inactivation of the MMR system makes the cells tolerant to these lesions (methylation tolerance phenotype) and allows them to escape apoptosis. Lynch syndrome (LS) is the main cause of familial colorectal cancer. It is due to monoallelic germ-line mutations affecting MMR genes. The diagnosis of this syndrome is based on the sequencing and detection of deleterious germline MMR mutations in patients. This diagnostic approach raises the issue of the functional interpretation of the detected variants which is ambiguous because those sequence alterations could be pathogenic or not. Variants of Uncertain Significance (VUS) account for 30% of the alterations detected in the *MLH1* and *MSH2* genes (the two genes the most frequently involved in LS). This issue is expected to increase with the routine use of new generation sequencing. Our goal was to develop a simple functional test to discriminate pathogenic from non-pathogenic VUS in LS-suspected patients

In the continuity of the work recently published by our team, our experimental strategy is based on the ability of *MLH1* and *MSH2* variants to complement the MMR deficiency of a cell line in culture. In short, a restoration of methylation sensitivity is expected if the transfected variant is non-pathogenic, while a preservation of methylation tolerance indicates that the variant is deleterious. We first tested a panel of 40 *MLH1* and *MSH2* variants whose pathogenicity or neutrality is known. Our test correctly classified 39/40 variants, showing the relevance of our methodological approach. The technique was then applied to a series of 49 VUS (25 *MLH1* and 24 *MSH2*) to be characterized. Overall, our technique has proposed a classification for 41/49 (83%) of VUS: 12 VSI (6 *MLH1* and 6 *MSH2*) are considered competent and 29 VUS (13 *MLH1* and 16 *MSH2*) are considered deficient. The classification of the remaining 8 VUS (6 *MLH1* and 2 *MSH2*) is uncertain. The other available data (clinical, somatic, *in silico* and possibly functional data) confirm our results and validate our approach.

In the future, we expect that this DNA damage signaling test would be routinely available on a diagnostic platform to prospectively study all the VUS detected in LS-suspected patients.

# Table des matières

<b>REVUE BIBLIOGRAPHIQUE</b> .....	1
Chapitre 1 : Le système MMR .....	3
A. Rôle dans la réparation des mésappariements.....	3
1. Les complexes MutS et MutL .....	4
2. La réparation du mésappariement.....	4
3. L'efficacité de réparation .....	7
B. Rôle dans la signalisation apoptotique.....	7
1. La prise en charge des lésions de l'ADN liées à la méthylation par la MGMT .....	8
2. La prise en charge des lésions de l'ADN liées à la méthylation par le système MMR .....	9
Chapitre 2 : Les cancers de type MSI.....	15
A. L'oncogenèse colorectale MSI.....	15
1. Introduction sur les cancers colorectaux .....	15
2. L'inactivation du système MMR comme origine du phénotype MSI .....	17
3. Les mutations induites par l'inactivation du système MMR.....	18
B. Le traitement des CCR .....	20
1. Le traitement des CCR indépendamment de leur statut microsatellitaire .....	20
2. L'émergence des immunothérapies pour le traitement des CCR MSI .....	21
3. Le phénotype tumoral MSI : valeur pronostique et prédictive .....	23
C. Détermination du statut microsatellitaire d'une tumeur .....	24
1. Approche immunohistochimique.....	24
2. Approche moléculaire .....	25
D. Cancers MSI iatrogènes.....	27
1. Cancers induits par les thiopurines .....	27
2. Cancers induits par les agents méthylants.....	29
Chapitre 3 : Syndromes génétiques prédisposant au développement de tumeurs MSI .....	31
A. Le Syndrome de Lynch.....	31
1. Description clinique.....	31
2. Description moléculaire .....	34
3. Diagnostic du syndrome de Lynch.....	38
B. Le syndrome CMMRD.....	42
1. Description clinique.....	42

2. Description moléculaire .....	43
3. Diagnostic du CMMRD.....	46
Chapitre 4 : La problématique des Variants de Signification Incertaine .....	53
A. Définition générale.....	53
B. Arguments pour l'interprétation des variants des gènes MMR.....	55
1. Données cliniques et études de co-ségrégation .....	55
2. Données épidémiologiques .....	57
3. Les algorithmes de prédiction in silico .....	59
4. Tests fonctionnels.....	61
5. Intégration des données somatiques .....	73
C. La classification des variants en 5 niveaux .....	74
<b>TRAVAUX DE RECHERCHE .....</b>	<b>75</b>
<b>ANNEXES .....</b>	<b>123</b>
<b>CONCLUSION, DISCUSSION ET PERSPECTIVES .....</b>	<b>137</b>
<b>BIBLIOGRAPHIE .....</b>	<b>155</b>

## Liste des Illustrations

<b>FIGURE 1</b>   Mécanisme de la réparation des mésappariements du système MisMatch Repair .....	6
<b>FIGURE 2</b>   Prise en charge des lésions O6-MeG induites par les agents méthylants .....	10
<b>FIGURE 3</b>   L'induction de la mort cellulaire après prise en charge par le système MMR des lésions induites par les agents méthylants .....	11
<b>FIGURE 4</b>   Immunité anti-tumorale, checkpoint immunitaire (ICK) et thérapie anti-ICK.....	22
<b>FIGURE 5</b>   Détermination du statut MSI des tumeurs.....	26
<b>FIGURE 6</b>   Le syndrome de Lynch : gènes impliqués, corrélation génotype-phénotype et risque de développement des différentes tumeurs du spectre Lynch .....	37
<b>FIGURE 7</b>   Algorithme du diagnostic moléculaire du syndrome de Lynch .....	41
<b>FIGURE 8</b>   Le syndrome CMMRD : gènes impliqués, corrélation génotype-phénotype et cancers développés par les patients .....	45
<b>FIGURE 9</b>   Distribution des variants des gènes MMR associés au syndrome de Lynch au sein des différentes classes de pathogénicité .....	54
<b>FIGURE 10</b>   Résumé des tests de caractérisation fonctionnelle des variants des gènes MMR existants .....	73
<b>FIGURE 11</b>   Interprétation des données fonctionnelles pour la caractérisation des VSI proposée par InSiGHT .....	135

## Liste des Tables

<b>TABLE 1</b>   Les protéines du système MisMatch Repair et les protéines accessoires impliquées dans la réparation des mésappariements chez l'Homme .....	5
<b>TABLE 2</b>   Critères d'Amsterdam II et de Bethesda révisés .....	39
<b>TABLE 3</b>   Critères cliniques préconisant la recherche d'un syndrome CMMRD chez un patient atteint de cancer .....	48
<b>TABLE 4</b>   Exemple de résultats discordants dans l'interprétation des VSI par différentes équipes .	146

## Liste des abréviations

**ACMG** : American College of Medical Genetics and Genomics

**ADN** : Acide DésoxyriboNucléique

**ADP** : Adénosine DiPhosphate

**APC** : *Adenomatous polyposis coli*

**ARN** : Acide RiboNucléique

**ATM** : *Ataxia Telangiectasia Mutated*

**ATR** : *Ataxia Telangiectasia and Rad3-related*

**ATP** : Adénosine DiPhosphate

**Aza** : Azathioprine

**CCR** : Cancer ColoRectal

**CHK** : *CHeckpoint Kinase*

**CIMP** : *CpG Island Methylator Phenotype* (phénotype de méthylation des îlots CpG)

**CIN** : *Chromosomal INstability* (instabilité chromosomique)

CMH : Complexe Majeur d'Histocompatibilité

**CMMRD** : *Constitutional MisMatch Repair Deficiency*, syndrome de déficience constitutionnelle du système de réparation des mésappariements

**CpG** : *Cytosine-phosphate-Guanine*

**CTL** : *Cytotoxic T Lymphocyte*, lymphocyte T cytotoxique

**CTLA-4** : *Cytotoxic T Lymphocyte Antigen 4*

**dNTP** : mélange des quatre désoxyribonucléotides : dATP (désoxy-Adénine Tri-Phosphate), dCTP (désoxy-Cytosine Tri-Phosphate), dGTP (désoxy-Guanine Tri-Phosphate), dTTP (désoxy-Thymine Tri-Phosphate)

**EPCAM** : *Epithelial Cell Adhesion Molecule*

**EXO1** : EXOnucléase 1

**HNPCC** : *Hereditary Non-Polyposis Colorectal Cancer*, cancer colorectal héréditaire sans polypose

**ICK** : Immune Checkpoint

**IDL** : *Insertion/Deletion Loop*, boucle d'insertion/deletion

**IHC** : ImmunoHistoChimie

**InSiGHT** : International Society for Gastrointestinal Hereditary Tumours

**IRI** : Irinotécan

**LAM** : Leucémie Aigüe Myéloblastique

**LLB** : Lignée LymphoBlastoïde

**LNH** : Lymphome Non Hodgkinien

**LS** : *Lynch Syndrome*, syndrome de Lynch

**LV** : Leucovorine

**MGMT** : O6-MéthylGuanine-MéthylTransférase

**MICI** : Maladie Inflammatoire Chronique de l'Intestin

**MLPA** : *Multiplex Ligation-dependent Probe Amplification*

**MMR** : *MisMatch Repair*, système de réparation des mésappariements

**MNNG** : N-Méthyl-N'-Nitro-N-NitrosoGuanidine

**MNU** : N-Méthyl-N-NitrosoUrée

**MSI** : MicroSatellite Instable

**MSS** : MicroSatellite Stable

**NF1** : neurofibromatose de type 1 (nom du syndrome génétique et du gène dont la mutation provoque ce syndrome)

**NGS** : Next Generation Sequencing, séquençage de nouvelle génération

**NMD** : *Nonsense-Mediated Decay* (dégradation des ARN non-sens)

**OXA** : Oxaliplatine

**O6-MeG** : Guanine méthylée sur la position O6

**PAF** : Polypase Adénomateuse Familiale

**PCNA** : Proliferating Cell Nuclear Antigen

**PCR** : *Polymerase Chain Reaction* (réaction en chaîne par polymérase)

**PD-1** : *Programmed Death-1*

**RCF** : Replication Factor C

**TIL** : *Tumor Infiltrating Lymphocyte* (lymphocyte infiltrant la tumeur)

**VIH** : Virus de l'Immunodéficience Humaine

**VSI** : Variant de Signification Incertaine

**5-FU** : 5-Fluoro-Uracile

**6-TG** : 6-ThioGuanine

# REVUE BIBLIOGRAPHIQUE



# Chapitre 1 : Le système MMR

Le système de réparation des mésappariements, ou *MisMatch Repair* (MMR), a un rôle capital pour le maintien de l'intégrité génomique à travers deux fonctions. Il permet d'une part la réparation des mésappariements provoqués par les erreurs de la polymérase lors de la réplication de l'ADN et d'autre part la signalisation apoptotique de certains dommages à l'ADN. Le système MMR est par ailleurs impliqué dans les processus de recombinaison pendant la mitose/méiose, de suppression de la recombinaison homéologue (c'est-à-dire entre séquences homologues mais non identiques), de diversité des immunoglobulines et d'expansion des répétitions trinuécléotidiques associées à certaines maladies dégénératives (Jiricny, 2006 ; Li, 2008 ; Reyes *et al.*, 2015). Dans ce chapitre, nous nous intéresserons essentiellement aux deux premières fonctions du système MMR : son implication dans la réparation des mésappariements, rôle le plus connu ; et son rôle dans la signalisation apoptotique de certains dommages à l'ADN, qui est la propriété que nous exploitons pour le test de caractérisation fonctionnelle des variants des gènes MMR que nous avons développé. Le système MMR est hautement conservé des procaryotes aux eucaryotes, mais nous détaillerons exclusivement dans ce travail les complexes présents chez la levure et les mammifères.

## A. Rôle dans la réparation des mésappariements

Au cours de la réplication, les ADN polymérases répliquatives  $\delta$  et  $\epsilon$  font des erreurs tous les  $10^4$  à  $10^5$  nucléotides incorporés (Fortune *et al.*, 2005 ; Shcherbakova *et al.*, 2003). Heureusement, elles contiennent une activité de relecture (« *proof-reading* ») qui leur permet d'éliminer significativement les nucléotides incorrectement introduits, réduisant la fréquence des mésappariements à  $10^{-6}$ - $10^{-7}$  (Reha-Krantz, 2010). Cependant, quelques erreurs subsistent encore, notamment des boucles d'insertion/délétion (IDL, *Insertion/Deletion Loop*) dans les longues séquences répétées, qui ne sont pas corrigées de façon efficace par l'activité de relecture (Kroutil *et al.*, 1996). La correction de l'ensemble des erreurs restantes est à la charge du système MMR, qui permet de diminuer encore le taux d'erreurs d'un facteur 1000, menant à une fréquence d'erreur finale de  $10^{-9}$  à  $10^{-10}$ . Ainsi, le système MMR rend possible la fidélité de la réplication au niveau d'un génome presque entier (qui contient environ  $6,6 \cdot 10^9$  paires de bases sous sa forme diploïde), permettant d'atteindre un taux de mutations acceptable pour la propagation de l'espèce mais qui autorise tout de même quelques variations génétiques (Ganai et Johansson 2016).

## ***1. Les complexes MutS et MutL***

Chez les eucaryotes, le système MMR comprend deux principaux complexes : MutS et MutL. Le rôle du complexe MutS est de reconnaître et de lier le mésappariement, alors que le complexe MutL sert d'interface entre la reconnaissance de l'erreur de réplication et la cascade d'événements en aval qui vont permettre la réparation de cette erreur.

Le complexe MutS est constitué de la protéine MSH2, qui agit en dimère avec MSH1, MSH3 ou MSH6. La protéine MSH1 n'existe que chez la levure et a un rôle dans la réparation des mésappariements dans l'ADN mitochondrial. Le complexe constitué de MSH4 et MSH5 est impliqué dans la recombinaison méiotique : ces protéines sont dépourvues du domaine nécessaire à la réparation des mésappariements et sont exprimées de façon spécifique dans les tissus reproductifs<sup>9</sup>. Le rôle de ces 3 protéines ne sera pas détaillé ici. Les dimères intéressants pour la fonction de réparation des mésappariements du système MMR sont donc MSH2-MSH6 (MutS $\alpha$ ) et MSH2-MSH3 (MutS $\beta$ ). Les protéines MSH6 et MSH3 ont une spécificité de substrat, qui se chevauche partiellement. En effet, le complexe MutS $\alpha$  reconnaît principalement les mésappariements ainsi que les IDL d'une à 2 paires de bases. De son côté, le complexe MutS $\beta$ , bien que capable de reconnaître quelques mésappariements, prend en charge principalement les petites et grandes IDL. Les complexes MutS $\alpha$  et MutS $\beta$  présentent donc une redondance partielle de fonction. Chez l'homme, le complexe MutS $\alpha$  est environ 10 fois plus abondant que le complexe MutS $\beta$  (Reyes *et al.*, 2015). Une surexpression de la protéine MSH3 conduit à un phénotype mutateur fort. En effet, l'excès de MSH3 entraîne une sur-dimérisation de MSH2 avec MSH3, au détriment de la dimérisation de MSH2 avec MSH6, réduisant ainsi la quantité de MutS $\alpha$  dans la cellule et donc l'activité de ce dimère (Drummond *et al.*, 1997).

Le complexe MutL est constitué de la protéine MLH1, en dimère avec les protéines PMS2 (MutL $\alpha$ ), PMS1 (MutL $\beta$ ) ou MLH3 (MutL $\gamma$ ). MutL $\alpha$  est le complexe impliqué dans la réparation des mésappariements et monopolise 90% de la protéine MLH1 (Modrich, 2006). MutL $\gamma$  a un rôle dans la recombinaison méiotique (en combinaison avec MSH4 et MSH5) et MutL $\beta$  ne présente aucun rôle biologique spécifique identifié.

## ***2. La réparation du mésappariement***

La réparation du mésappariement comprend une cascade d'événements, depuis la reconnaissance du mésappariement par MutS, l'activation de la signalisation en aval par MutL, l'excision du brin néosynthétisé contenant le mésappariement, la resynthèse puis la ligation.

Le complexe MutS, sous sa forme lié à l'ADP (Adénosine DiPhosphate), reconnaît le mésappariement. La liaison au mésappariement induit un changement de conformation qui permet

un échange de l'ADP vers l'ATP (Adénosine TriPhosphate). Un deuxième changement de conformation, provoqué par cette liaison à l'ATP, conduit à la formation d'une « pince coulissante » (« *sliding clamp* ») capable de glisser le long de l'ADN (Lee *et al.*, 2014) à la recherche d'un signal de reconnaissance du brin néosynthétisé. En effet, la distance entre le mésappariement et la fin du fragment qui permet de discriminer le nouveau brin peut aller jusqu'à 1000 paires de bases (Fang et Modrich, 1993).

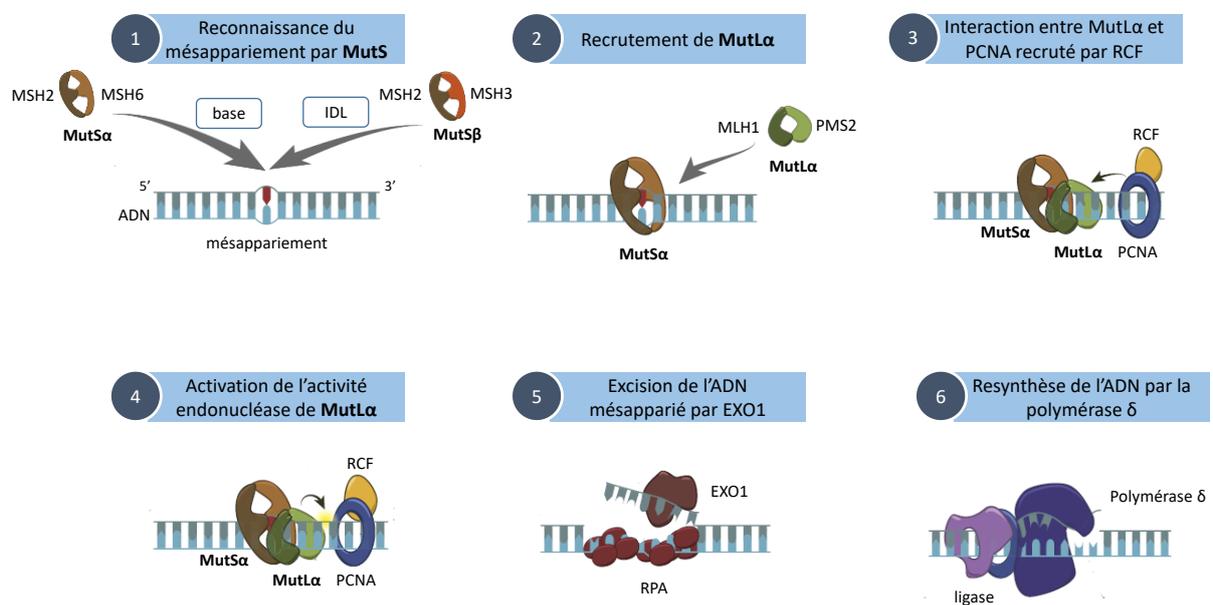
De plus, cette conformation de MutS activée par l'ATP permet son interaction avec MutL $\alpha$ <sup>14</sup>. Ce dernier est ensuite activé par PCNA (*Proliferating Cell Nuclear Antigen*), composant de la machinerie de réplication chargé sur l'ADN grâce à RCF (*Replication Factor C*). Cette activation de MutL $\alpha$  lui permet d'inciser le brin néosynthétisé d'une manière ATP-dépendante grâce à son activité endonucléase (Kadyrov *et al.*, 2006). Cette incision sert de point d'ancrage à l'exonucléase EXO1 qui élimine un fragment d'ADN contenant le mésappariement (Genschel *et al.*, 2002) : l'excision s'arrête environ 150 nucléotides au-delà du mésappariement<sup>13</sup>. La resynthèse de l'ADN est assurée par les polymérase  $\delta$  et/ou  $\epsilon$  (Longley *et al.*, 1997), puis la réaction est terminée par une ligation. Le processus de réparation des mésappariements est schématisé sur la **FIGURE 1** et les protéines qui y sont impliquées sont résumées dans la **TABLE 1**.

Les protéines impliquées dans la réparation des mésappariements et leurs rôles	
<b>Protéines MMR</b>	
MSH2-MSH6 (MutS $\alpha$ )	Reconnaissance des mésappariements et des petites boucles insertions/délétions
MSH2-MSH3 (MutS $\beta$ )	Reconnaissance des boucles insertions/délétions
MLH1-PMS2 (MutL $\alpha$ )	Interaction avec MutS, activité endonucléase, activation de la cascade
<b>Protéines accessoires</b>	
EXO1	Exonucléase : élimination du fragment d'ADN contenant le mésappariement
PCNA	Facteur de processivité de la polymérase, stimule l'activité endonucléase de MutL $\alpha$
RCF	Chargement de PCNA
RPA	Protéine de liaison de l'ADN simple brin
Polymérase $\delta$ et/ou $\epsilon$	Resynthèse de l'ADN excisé
Ligase 1	Jonction de l'ADN après resynthèse

**TABLE 1** | Les protéines du système MisMatch Repair et les protéines accessoires impliquées dans la réparation des mésappariements chez l'Homme  
Adapté de (Reyes *et al.*, 2015)

Les erreurs de réplication sont présentes sur le brin néosynthétisé, et la réparation des mésappariements par le système MMR doit donc être dirigée exclusivement sur ce brin, plutôt que sur le brin parental, pour éliminer systématiquement la base erronée nouvellement incorporée et non pas la base originelle. Le système MMR doit donc pouvoir discriminer les deux brins. Cette discrimination est assurée par la capacité de MutS à former un « *sliding clamp* ». Par ce mécanisme, le complexe MutS

facilite la communication entre deux sites distincts sur l'ADN : celui du mésappariement et celui qui permet la discrimination du brin néosynthétisé (Li, 2008). La reconnaissance du brin à réparer est assurée par la présence de coupures régulières sur le brin néosynthétisé. Concernant le brin de réplication tardive, ces discontinuités régulières sont dues aux intervalles entre les différents fragments d'Okasaki (Pavlov *et al.*, 2003). D'autres coupures sont dues à l'élimination des ribonucléotides incorporés à la place des dNTP pendant la réplication de l'ADN (Ghodgaonkar *et al.*, 2013). Par ailleurs, PCNA peut également présenter un rôle dans la discrimination du brin. En effet, PCNA est chargé sur l'ADN avec une orientation spécifique : cette polarité permettrait l'incision de l'ADN par MutL $\alpha$  sur le bon brin (Pluciennik *et al.*, 2010).



**FIGURE 1 | Mécanisme de la réparation des mésappariements du système Mismatch Repair**

① L'ADN mésapparié peut être un mésappariement base/base ou une boucle d'insertion/délétion (IDL). Selon la nature du mésappariement, il est reconnu par le complexe MutS $\alpha$  (MSH2-MSH6) ou MutS $\beta$  (MSH2-MSH3). ② MutL $\alpha$  (MLH1-PMS2) est recruté au site du mésappariement. ③ PCNA interagit avec MutL $\alpha$  après avoir été chargé sur l'ADN par RCF, ④ ce qui active l'activité endonucléasique de MutL $\alpha$ . ⑤ EXO1 réalise l'excision du fragment d'ADN contenant le mésappariement, pendant que RPA se lie à l'ADN simple brin pour le protéger. ⑥ Un nouveau brin d'ADN est synthétisé par l'ADN polymérase puis est lié à la molécule d'ADN initiale par la ligase. *Adapté de (Li et Martin, 2016).*

Il est intéressant de remarquer que le système MMR requiert la présence de PCNA et des intervalles entre les fragments d'Okasaki pour remplir sa fonction. Cela suggère que son activité soit directement couplée à la réplication, ce qui laisse une fenêtre d'action assez courte. MSH6 et MSH3 sont capables d'interagir avec PCNA, et donc avec la machinerie de réplication<sup>22</sup>. Ainsi, pour éviter que les mésappariements et les IDL incorporés par la polymérase soient introduits dans le génome de manière pérenne, le système MMR doit corriger les erreurs avant le cycle suivant de réplication.

### 3. *L'efficacité de réparation*

L'efficacité de réparation dépend du mésappariement considéré (St Charles *et al.*, 2015). Parmi les huit mésappariements possibles, les G-G, A-T, G-T et A-C sont très bien réparés, les T-T, C-T et G-A sont moins bien réparés et les C-C, les erreurs de réplication les plus rares, sont très mal réparés (Dohet *et al.*, 1985 ; Lahue *et al.*, 1989 ; Modrich 1991). Par ailleurs, il existe une corrélation inverse entre l'efficacité de réparation d'un mésappariement donné et sa capacité à induire une distorsion dans l'hélice d'ADN. Les mésappariements G-T induisent une faible distorsion et sont bien réparés alors que les mésappariements C-C sont les plus déstabilisants et les moins bien réparés (Hsieh et Yamane, 2008). Le fait que les complexes MutS ne reconnaissent pas le mésappariement en lui-même, mais qu'ils détectent plutôt la flexibilité intrinsèque de l'ADN dans la région du mésappariement peut expliquer la grande gamme de mésappariement et de lésions qui peuvent être reconnus par les protéines MSH (Mazurek *et al.*, 2009). Si la capacité de réparation dépend du mésappariement en lui-même, elle varie également en fonction du contexte génétique (Lujan *et al.*, 2012). En effet, l'affinité de liaison de MutS dépend du type de mésappariement mais également de la séquence nucléotidique entourant le mésappariement (Mazurek *et al.*, 2009). Par ailleurs, les erreurs sur le brin retardé sont réparées plus efficacement que sur le brin avancé. Il est possible que la quantité plus importante d'extrémités 5' sur les fragments d'Okazaki et la plus haute densité de PCNA sur le brin retardé puisse stimuler l'activité de réparation du système MMR (Pavlov *et al.*, 2003). Cependant, la Pol  $\epsilon$ , qui réplique le brin avancé, présente une fidélité plus haute que la polymérase  $\delta$ , qui réplique le brin retardé. Ainsi, les deux brins présentent une fidélité de la réplication globalement identique (Ganai et Johansson, 2016).

### B. Rôle dans la signalisation apoptotique

Outre son action dans la réparation des mésappariements, le système MMR joue un rôle dans la réponse cellulaire à certains agents génotoxiques, tels que les agents méthylants, les thiopurines, les dérivés de platine et les agents oxydants. Dans cette étude, nous ne parlerons que des agents alkylants de type  $SN_1$  (substitution nucléophile monomoléculaire) tels que les réactifs MNNG (N-Methyl-N'-nitro-nitrosoguanidine), MNU (Methyl-N-nitrosourea) ou encore le témozolomide, une molécule utilisée en chimiothérapie.

Ces agents alkylants sont capables de modifier les bases déjà incorporées dans l'ADN ainsi que les bases des dNTP, par ajout d'un groupement alkyle sur les positions N3 et N7 de l'adénine, O2 et N3 de la cytosine, O2, N3 et O4 de la thymine et N3, N7 et O6 de la guanine (Saffhill *et al.*, 1985). En fonction de leur position sur l'ADN, les différentes bases modifiées peuvent compromettre l'intégrité

génomique en induisant de la mutagenèse ou en bloquant des processus biologiques essentiels tels que la réplication et la transcription de l'ADN (qui peuvent potentiellement mener à la mort cellulaire). Enfin, certaines lésions peuvent être prises en charge par des mécanismes de réparation de l'ADN, tels que le système MMR, qui aboutissent à l'apoptose (Fu *et al.*, 2012). Ce processus de mort cellulaire induit par un mécanisme de réparation de l'ADN, qui peut paraître contre-intuitif, sera détaillé ci-dessous.

Parmi toutes les modifications précédemment citées, c'est la production d'O6-MeG (méthylation sur l'atome d'oxygène en position 6 de la guanine), pourtant minime (environ 5% des lésions induites), qui confère aux agents méthylants leur cytotoxicité (Swann *et al.*, 1996).

### ***1. La prise en charge des lésions de l'ADN liées à la méthylation par la MGMT***

La première ligne de défense contre les lésions O6-MeG est assurée par l'enzyme MGMT (O6-MéthylGuanine-MéthylTransférase), qui constitue à elle seule un système de réparation de l'ADN qui ne nécessite aucune autre protéine ou cofacteur. La MGMT transfère le groupement méthyle de la guanine modifiée vers un de ses propres résidus cystéine, ce qui l'inactive. Cette protéine peut donc être qualifiée d'enzyme « suicide », puisqu'une molécule de MGMT permet de réparer une unique lésion O6-MeG. Cette réaction est très rapide, mais saturable (Fahrer *et al.*, 2015).

La réponse cellulaire à un agent méthylant a été comparée entre des couples isogéniques de lignées MGMT-déficientes et compétentes. Ces expériences sont basées sur la transfection de cellules MGMT-déficientes par un plasmide porteur de la séquence du gène *MGMT* sauvage de bactérie (Ishizaki *et al.*, 1986 ; Kataoka *et al.*, 1986 ; Samson *et al.*, 1986) ou humain (Kaina *et al.*, 1991 ; Meikrantz *et al.*, 1998 ; Wu *et al.*, 1991). À l'inverse, la perte d'expression de la MGMT peut être forcée dans des cellules initialement MGMT-compétentes par inactivation génétique (invalidation du gène) ou grâce à l'utilisation d'un inhibiteur (O6-benzylguanine) (Branch *et al.*, 1995 ; Dolan et Pegg, 1997 ; Kaina *et al.*, 1997 ; Tominaga *et al.*, 1997). De manière générale, après exposition à un agent méthylant, les cellules MGMT-compétentes sont capables de réparer les lésions et de survivre alors que les cellules MGMT-déficientes présentent une sensibilité accrue et un taux plus élevé d'apoptose. Cette différence de réponse est effectivement due à la seule différence d'expression de la MGMT car ces expériences ont été réalisées dans des couples de lignées cellulaires isogéniques ne différant que par l'expression de la MGMT. D'autre part, les deux types de cellules répondent de la même manière à d'autres agents endommageant l'ADN, dont la cytotoxicité ne passe pas par la production de lésions O6-MeG (Hickman et Samson, 1999). La perte de la MGMT conduit à la persistance des lésions O6-MeG dans l'ADN (Kaina *et al.*, 1990), indiquant que ces lésions sont exclusivement réparées par cette

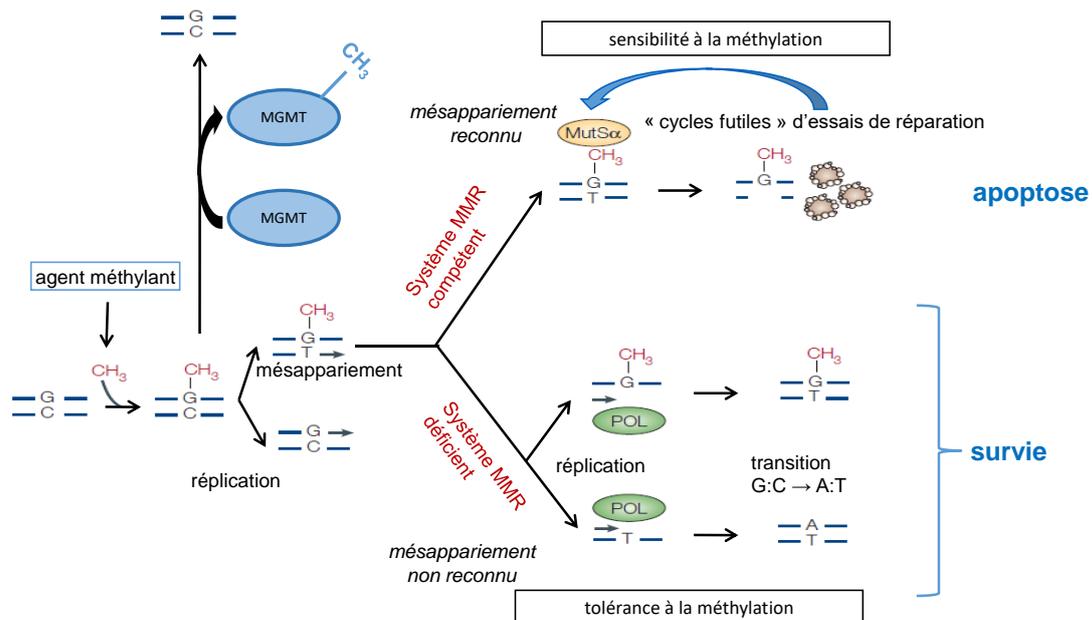
enzyme (Pegg et Byers, 1992). De plus, des fibroblastes de souris *Mgmt*<sup>-/-</sup> sont plus sensibles au témozolomide que des fibroblastes *Mgmt*<sup>+/-</sup> (Glassner *et al.*, 1999) et montrent de hauts niveaux de mort cellulaire dans les tissus à prolifération rapide, tels que la moelle osseuse, l'intestin, le thymus et la rate (Fu *et al.*, 2012).

L'effet protecteur de la MGMT a été mis en évidence chez des souris transgéniques sur-exprimant la MGMT : ces souris présentent une réduction significative du nombre de lymphomes, de tumeurs du foie et de cancers colorectaux (selon l'organe où l'expression de la MGMT est induite) après exposition à des agents méthylants, montrant que la MGMT est bien la première ligne de défense contre les dommages provoqués par ces agents (Allay *et al.*, 1999 ; Becker *et al.*, 1996 ; Dumenco *et al.*, 1993 ; Nakatsuru *et al.*, 1993 ; Zaidi *et al.*, 1995 ; Zhou *et al.*, 2001).

Des données cliniques montrent que l'activation de la MGMT est associée à une résistance au témozolomide, agent méthylant qui est le traitement de référence du glioblastome, tumeur cérébrale maligne la plus courante chez les adultes (Friedman *et al.*, 2000 ; Sarkaria *et al.*, 2008). En effet, des patients dont le promoteur du gène de la MGMT est hyperméthylé et chez qui la MGMT n'est donc pas exprimée, répondent mieux que les patients dont le promoteur n'est pas méthylé (Hegi *et al.*, 2005), faisant de la méthylation du promoteur de la MGMT un biomarqueur de prédiction de réponse au témozolomide (De Carlo *et al.*, 2018 ; Esteller *et al.*, 2000).

## ***2. La prise en charge des lésions de l'ADN liées à la méthylation par le système MMR***

En cas de déplétion de la MGMT, par saturation de l'enzyme ou par inactivation génétique, la prise en charge des lésions O6-MeG nécessite alors l'intervention d'un autre système de réparation : le système MMR (Aquilina et Bignami, 2001 ; Pegg, 2000). En effet, au cours de la réplication, la base modifiée O6-MeG n'est pas correctement reconnue par la polymérase qui est susceptible d'incorporer en face une thymine (Lukash *et al.*, 1991 ; Tomita-Mitchell *et al.*, 2000), ce qui conduira à des mutations en transition (G-C > A-T) au cours du cycle suivant de réplication si la lésion n'est pas reconnue (Abbott et Saffhill, 1979). Le mésappariement O6-MeG/T est détecté par le système MMR, qui sera pourtant incapable de le réparer, ce qui va conduire à terme à l'apoptose de la cellule (**FIGURE 2**).



**FIGURE 2 | Prise en charge des lésions O6-MeG induites par les agents méthylants**

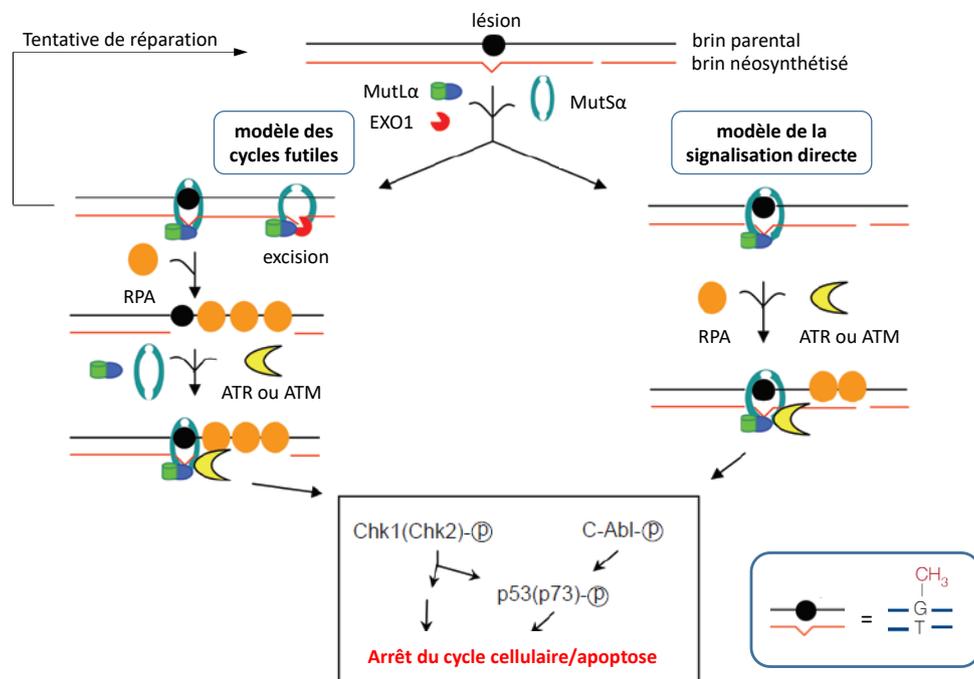
La première ligne de défense contre ces lésions est assurée par la MGMT (O6-MéthylGuanine-MéthylTransférase) qui permet d'éliminer directement la lésion. En cas d'absence de la MGMT (inactivation du gène ou déplétion de la protéine par saturation), la base méthylée sera reconnue au cours de la réplication suivante comme une adénine par la polymérase, qui va introduire une thymidine sur le néo-brin, ce qui provoque l'apparition d'un mésappariement. Dans une cellule MMR-compétente, ce mésappariement sera reconnu par le système MMR (ici représenté par le complexe MutS $\alpha$ ), ce qui va induire la mort cellulaire. Dans une cellule-MMR déficiente, le mésappariement n'est pas reconnu et la cellule est capable de survivre, au prix de l'accumulation de mutations dans le génome susceptibles d'impulser la tumorigenèse. Ce processus, qui permet la survie de la cellule malgré la persistance des lésions, est nommé « tolérance à la méthylation ». *Adapté de* <sup>65</sup>.

Le mécanisme responsable de cette mort cellulaire après reconnaissance par le système MMR du mésappariement induit par un agent méthylant n'est pas encore clairement élucidé et deux modèles non exclusifs sont généralement proposés.

Comme explicité précédemment, le système MMR est spécifique du brin néosynthétisé et ne peut donc pas éliminer la véritable lésion présente sur le brin parental. Par conséquent, c'est la base T qui va être éliminée, mais une nouvelle base T est susceptible d'être à nouveau incorporée au cours de la resynthèse. Dans le modèle de cycles futiles (représenté sur la **FIGURE 2**), ce sont ces tentatives de réparation infructueuses générant des portions d'ADN simple brin puis double brin et un arrêt du cycle cellulaire en phase G2 qui conduiraient, à terme, à l'apoptose (Mojas *et al.*, 2007). Des études ont montré que si la base O6-MeG est sur un brin présentant des cassures (qui mime donc le brin néosynthétisé), en face d'une base T sur un brin ininterrompu (qui mime le brin parental), le mésappariement est correctement corrigé. Cela montre que le système MMR est capable de reconnaître mais également de réparer ces lésions, mais que la réparation est impossible dans le contexte physiologique, à cause de la discrimination de brin (Karran *et al.*, 2003).

Dans le modèle de signalisation directe, le système MMR activerait directement la voie de signalisation qui implique les protéines ATM (*Ataxia Telangiectasia Mutated*) et ATR (*Ataxia Telangiectasia and Rad3-related*) en réponse aux dommages de l'ADN (Yoshioka *et al.*, 2006). L'arrêt du cycle serait ensuite médié par CHK1 et CHK2 (*CHeckpoint Kinase*). Enfin, l'apoptose est provoquée par la phosphorylation de p53 (Duckett *et al.*, 1999). Des études de complexes protéiques ont montré que MSH2 peut interagir directement avec ATR et CHK2 et MLH1 avec ATM (Adamson *et al.*, 2005 ; Brown *et al.*, 2003 ; Wang et Qin, 2003).

Dans les deux modèles, ce ne sont pas les lésions O6-MeG induites par les agents alkylants en elles-mêmes qui sont cytotoxiques pour la cellule, mais leur prise en charge par le système MMR (Carethers *et al.*, 1996 ; Karran et Stephenson, 1990) (**FIGURE 3**).



**FIGURE 3 | L'induction de la mort cellulaire après prise en charge par le système MMR des lésions induites par les agents méthylants**

Dans notre cas, la lésion est une guanine méthylée (O6-MeG) en face de laquelle la polymérase a incorporé une thymine, ce qui provoque un mésappariement (voir encadré en bas à droite). Le modèle des cycles futiles (gauche) suggère que le mésappariement est reconnu par le système MMR, qui va enclencher le processus de réparation. La base modifiée étant sur le brin parental, elle ne peut pas être excisée, ce qui va provoquer l'initiation d'un nouveau cycle de tentative de réparation. Les cycles répétés et futiles provoquent l'apparition de cassures simple puis double brin qui vont à terme signaler la mort cellulaire. Le modèle de signalisation directe (droite) propose que la reconnaissance du mésappariement par MutSα provoque directement une signalisation apoptotique de la lésion. *Adapté de (Li 2008).*

La prise en charge par le système MMR des bases modifiées nécessite que la lésion initiale soit capable d'être répliquée et que le produit qui en découle ressemble à une erreur susceptible d'être reconnue par le système MMR. Sur l'ensemble des bases endommagées produites par les agents méthylants, seule l'O6-MeG remplit ces deux critères. Les autres bases méthylées sont soit mal répliquées (par exemple 3-MeA) ou non perçues comme des mésappariements par le système MMR (7-MeG ou O4-MeT par exemple) (Karran *et al.*, 2003).

Le système MMR a été mis en cause dans ce processus dit de « tolérance à la méthylation » quand il a été observé que des souches de bactéries *Escherichia coli* MMR-déficientes présentaient une diminution du taux de mort cellulaire après exposition aux agents méthylants (Karran et Bignami 1992 ; Karran et Marinus, 1982). On parle de « tolérance » et non de « résistance » car les lésions O6-MeG ne sont pas éliminées : elles restent présentes sur le génome, mais n'induisent plus la mort cellulaire. Ce phénotype a ensuite été mis en évidence dans des lignées cellulaires mammifères MMR-déficientes qui montraient une résistance à la mort cellulaire 100 fois plus importante que des lignées MMR-compétentes ( Bignami *et al.*, 2000 ; Branch *et al.*, 1993 ; Branch *et al.*, 1995 ; Carethers *et al.*, 1996 ; Drummond *et al.*, 1996 ; Fink *et al.*, 1998 ; Hawn *et al.*, 1995 ; Hsieh et Yamane, 2008). Des cellules *MLH1*-, *MSH2*-, *MSH6*- ou *PMS2*-déficientes sont toutes tolérantes à la méthylation (Aquilina *et al.*, 1998). La relation entre la résistance à la drogue et le statut MMR est démontrée par le fait qu'une lignée *MLH1*-déficiente tolérante au MNNG perd sa tolérance à la drogue quand le défaut de *MLH1* est complété par une version du gène sauvage sur le chromosome 3 (Koi *et al.*, 1994). Il en va de même avec un transfert du chromosome 2 (porteur des gènes *MSH2* et *MSH6*) dans une lignée *MSH2* ou *MSH6*-déficiente (HEC59 et HCT15, respectivement) (Tindall *et al.*, 1998 ; Umar *et al.*, 1997). Le gène sauvage peut également être présent sur un plasmide (Buermeyer *et al.*, 1999 ; Umar *et al.*, 1997). L'utilisation d'un plasmide ne portant qu'un seul gène permet de confirmer l'implication de ce gène dans la fonction étudiée. Le principe du test de caractérisation fonctionnelle des variants des gènes MMR sur lequel j'ai travaillé est directement inspiré de ces expériences.

La différence de réponse aux agents méthylants en fonction du statut MMR a également été étudiée *in vivo* chez des souris compétentes ou déficientes pour un gène MMR donné (Sansom *et al.*, 2003 , Sansom *et al.*, 2001 ; Toft *et al.*, 1999).

Ainsi, en l'absence de MGMT, l'inactivation du système MMR confère aux cellules la capacité de survivre à l'exposition aux agents méthylants (P. Branch *et al.*, 1993 ; D'Atri *et al.*, 1998 ; Kat *et al.*, 1993). Cette caractéristique des cellules MMR-déficientes a été utilisée au sein de notre laboratoire pour développer un test de diagnostic du syndrome CMMRD, syndrome de prédisposition au développement de tumeurs dès l'enfance dû à une mutation bi-allélique dans un gène MMR (*cf.* Chapitre 3). Alors que des cellules lymphoblastoïdes issues de patients CMMRD (MMR-déficientes)

présentent une tolérance à la méthylation, des cellules issues d'individus non atteints du syndrome CMMRD sont MMR-compétentes et présentent une sensibilité à la méthylation (Bodo *et al.*, 2015). Dans la continuité de ce travail, nous avons développé un test d'aide au diagnostic du syndrome de Lynch, syndrome de prédisposition au développement de tumeurs dû à une mutation mono-allélique dans un gène MMR. Ce test est basé sur la réponse à un agent méthylant de cellules MMR-déficientes dans lesquelles a été introduit un plasmide porteur de la séquence du gène MMR que l'on veut tester. Dans ces expériences, pour s'assurer que la réponse des cellules est uniquement fonction de leur statut MMR, et pas de l'expression de la MGMT, cette dernière est inactivée par l'utilisation d'un inhibiteur **(ARTICLE)**.



## Chapitre 2 : Les cancers de type MSI

Dans ce chapitre nous nous intéresserons à un type d'oncogenèse particulier : l'oncogenèse de type instable sur les microsatellites, provoquée par une inactivation du système MMR qui conduit à l'accumulation de mutations dans le génome, et ce particulièrement sur les séquences répétées (microsatellites). Ce type d'oncogenèse a été spécialement étudié dans le cancer colorectal (CCR), et dans ce chapitre, excepté dans les deux dernières parties (sections C et D), nous nous focaliserons principalement sur cette localisation tumorale. Les CCR instables sur les microsatellites peuvent se développer dans un contexte sporadique, qui sera détaillé dans ce chapitre, ou dans des contextes de prédisposition génétique, qui seront approfondis dans le chapitre suivant.

### A. L'oncogenèse colorectale MSI

#### 1. *Introduction sur les cancers colorectaux*

Le cancer de façon générale est défini par une instabilité génétique, qui conduit à l'apparition de mutations activatrices dans des oncogènes et des mutations inactivatrices dans des gènes suppresseurs de tumeurs. Selon les gènes impliqués, ces mutations provoquent des altérations dans des processus biologiques essentiels susceptibles de conduire à une autosuffisance aux signaux de croissance, à une insensibilité à l'inhibition de la croissance, à un potentiel de réplication illimité, à une angiogenèse *de novo*, à une inhibition de la mort cellulaire programmée, à une invasion tissulaire et aux métastases (Hanahan et Weinberg, 2011).

Le CCR est l'un des cancers le plus fréquemment diagnostiqué (le deuxième chez les hommes et le troisième chez les femmes) (Chang *et al.*, 2017) et il est également la troisième cause de décès par cancer (Nojadeh *et al.*, 2018). En France en 2015, plus de 43 000 CCR ont été diagnostiqués et presque 18 000 personnes en sont décédées (INCa 2016). Bien qu'il soit l'un des cancers les plus fréquents, le CCR est aussi l'un de ceux dont on guérit le mieux s'il est détecté à un stade précoce (Nojadeh *et al.*, 2018). L'amélioration dans la détection et la prise en charge ainsi qu'une meilleure compréhension moléculaire de la maladie font que le taux de mortalité dû au CCR est en constante réduction (Ryan *et al.*, 2017). Ce cancer reste néanmoins un fort enjeu de santé publique.

Les CCR peuvent être sporadiques ou se développer dans le contexte d'un syndrome de prédisposition génétique, tels que la polypose adénomateuse familiale (due à une mutation dans le gène *APC*), la polypose associée au gène *MUTYH*, le syndrome de Peutz-Jeghers (*STK11/LKB1*) ou encore le syndrome de Lynch (*MLH1, MSH2, MSH6, PMS2*) (Nojadeh *et al.*, 2018).

Dans le cas du cancer colorectal, l'instabilité génétique à l'origine de la tumorigenèse peut avoir deux causes différentes : une instabilité des chromosomes (CIN, *Chromosomal INstability*) ou une instabilité des microsatellites (MSI, *MicroSatellite Instability*). C'est ce deuxième type d'instabilité que je détaillerai dans ce chapitre.

Le phénotype MSI explique 12 à 15% des CCR. Les CCR MSI ont des caryotypes normaux ou peu altérés (Muleris *et al.*, 2008) et se développent principalement à partir d'adénome « sessile dentelé » dans le côlon proximal. De plus, les CCR MSI montrent une histologie mucineuse, une faible différenciation et un grand nombre de lymphocytes infiltrant la tumeur (TIL, *Tumor Infiltrating Lymphocyte*) (Greenson *et al.*, 2003), dont la présence est la conséquence d'une réaction inflammatoire importante. Les CCR MSI sporadiques sont plus souvent observés chez les femmes et chez les fumeurs et sont caractérisés par un âge plus tardif au diagnostic (Richman, 2015).

Le phénotype CIN est retrouvé dans environ 80 à 85% des CCR. Ces CCR sont caractérisés par une instabilité génétique due à des gains ou des pertes de chromosomes ou de segments chromosomiques, résultant en une variabilité du nombre de chromosomes, une fréquence importante de pertes d'hétérozygotie, des translocations de chromosomes et des amplifications de gènes. Les tumeurs CIN sont plus fréquentes dans le côlon distal et ont souvent une histologie non mucineuse, une différenciation modérée et peu de TIL. Ces tumeurs se forment à partir de polypes adénomateux et se développent selon la voie classique de tumorigenèse colique, à savoir avec des mutations séquentielles d'oncogènes et de gènes suppresseurs de tumeurs (tels que *APC*, *KRAS* et *TP53*), selon le modèle développé par Fearon et Vogelstein (Fearon et Vogelstein, 1990). Ces tumeurs ne présentent pas de mutation sur les microsatellites et sont donc également appelées MSS pour *MicroSatellite Stable*.

Les CCR constituent donc une maladie hétérogène. Grâce aux nouvelles techniques de séquençage, une classification moléculaire basée sur une analyse transcriptomique est possible. L'analyse des CCR de 750 patients (tous stades confondus) a permis à Marisa et ses collaborateurs de mettre en évidence six groupes (C1 à C6) en fonction de critères clinico-pathologiques et moléculaires. Dans cette étude, les tumeurs MSI sont majoritairement retrouvées dans le groupe C2, caractérisé par une forte infiltration lymphocytaire et une prolifération cellulaire importante (Marisa *et al.*, 2013). Plus récemment, un consensus nommé CMS (*Consensus Molecular Subtype*) a rassemblé les résultats de 6 études de classification moléculaire des CCR, incluant ceux de Marisa *et al.*, afin de mettre en évidence une signature moléculaire sur une cohorte de 4 151 échantillons de CCR, donnant lieu à une classification plus robuste. Cette étude a permis de mettre en exergue 4 groupes, dont le CMS1 qui regroupe 14% des tumeurs et qui est enrichi en tumeurs MSI caractérisées par une activation immunitaire importante (Guinney *et al.*, 2015).

## 2. *L'inactivation du système MMR comme origine du phénotype MSI*

En 1993, trois équipes indépendantes mettent en évidence une instabilité nucléotidique particulièrement marquée au niveau des séquences répétées (microsatellites) dans 12 à 15% de l'ensemble des CRC (héréditaires et sporadiques) (Aaltonen *et al.*, 1993 ; Ionov *et al.*, 1993 ; Thibodeau *et al.*, 1993). Ce processus de carcinogenèse a été nommé MSI, en opposition au mécanisme CIN (Boland *et al.*, 1998). D'autres types de tumeurs, en dehors du CCR, peuvent présenter un phénotype MSI. Les cancers de l'endomètre et les cancers gastriques présentent un phénotype MSI avec une fréquence du même ordre de grandeur que les CCR ; alors que les cancers du rein, de l'ovaire, de la prostate, du poumon ou du foie peuvent également avoir un phénotype MSI avec cependant une fréquence plus faible (Hause *et al.*, 2016). La tumorigenèse MSI, quelle que soit la localisation tumorale, est définie par un phénotype hypermutateur, particulièrement au niveau des microsatellites.

Les microsatellites sont des répétitions en tandem (généralement 10 à 60 répétitions) d'un motif constitué de un ou deux nucléotides voire plus. Les microsatellites représentent 3% du génome (Nojadeh 2018) et les plus fréquents sont des (CA)<sub>n</sub> (Ellegren, 2004). Au cours de la duplication de l'ADN, ces sites sont particulièrement sujets aux erreurs de réplication dues aux glissements (« *slippage* ») de l'ADN polymérase. Celle-ci intègre parfois trop mais le plus souvent pas assez de nucléotides dans le brin néosynthétisé, ce qui conduit à des boucles d'ADN non apparié (Yoon *et al.*, 2013). Le système MMR est capable de reconnaître et réparer ces erreurs, permettant la réplication fidèle de l'ADN. Cependant, en cas d'inactivation de ce système, les erreurs de réplication non réparées, notamment au sein des microsatellites, s'accumulent au fur et à mesure des divisions. Le phénotype tumoral MSI, conséquence d'une inactivation du système MMR, est donc caractérisé par un phénotype hypermutateur.

Dans 70 à 95% des cas, l'inactivation du système MMR est due à une hyperméthylation du promoteur du gène *MLH1*, médiée par le processus CIMP (*CpG Island Methylator Phenotype*) caractérisé par une hyperméthylation des régions promotrices des gènes dans les régions riches en bases C et G (îlots CpG pour *Cytosine-phosphate-Guanine*) (Cunningham *et al.*, 1998 ; Kane *et al.*, 1997). Cela conduit à une inactivation du système MMR par répression transcriptionnelle du gène *MLH1*. Les tumeurs MSI sont fréquemment associées à un phénotype CIMP. Il a été montré que la perte d'expression de *MLH1* par hyperméthylation du promoteur augmente avec l'âge : l'expression de la protéine est perdue dans environ 50% des patients atteints de CCR de plus de 90 ans (Kakar *et al.*, 2003). Ceci explique l'âge plus tardif au développement des CCR sporadiques MSI par rapport aux CCR sporadiques MSS.

L'inactivation du système MMR peut également être expliquée par des mécanismes alternatifs plus rares. Par exemple, la déficience du système MMR peut être due à des mutations somatiques sur les deux copies d'un gène MMR (Baretti et Le, 2018). Par ailleurs, des tumeurs MSI se sont avérées dépourvues d'altérations génétique ou épigénétique sur les gènes MMR. Il a été montré que des mutations de la triméthyltransférase SETD2, qui permet la triméthylation de la lysine 36 de l'histone 3 (H3K36-me3), sont également susceptibles d'être à l'origine d'un phénotype MSI. En effet, cette marque est nécessaire pour le recrutement de MutS $\alpha$  sur la chromatine. Ainsi, même si le système MMR est théoriquement actif, il ne peut pas remplir sa fonction (Li *et al.*, 2013). Par ailleurs, une surexpression de certains micro-ARN (tels que miR-155 ou miR-21) conduit à une baisse du taux des protéines MMR (Svrcek *et al.*, 2013 ; Valeri *et al.*, 2010a ; Valeri *et al.*, 2010b).

### ***3. Les mutations induites par l'inactivation du système MMR***

L'inactivation du système MMR n'est pas un événement transformant en soi : c'est l'instabilité génomique qu'elle engendre qui constitue un processus oncogénique. Le raccourcissement des microsatellites, qui a donné son nom au phénotype MSI, provoque des décalages du cadre de lecture et l'apparition de codons stop prématurés sur l'ARNm si la séquence est transcrite. Par ailleurs, le système MMR est également impliqué dans la réparation des bases mésappariées et l'inactivation du système MMR provoque donc également des mutations ponctuelles. Ainsi, les cancers MSI sont associés à un taux de mutations (décalage du cadre de lecture et mutations faux-sens) de 100 à 1000 fois plus important que les cancers non MSI (Baretti et Le, 2018).

Les microsatellites sont dispersés sur l'ensemble du génome, et 20% des gènes humains contiennent un microsatellite de 7 nucléotides ou plus au sein de leur région codante (El-Bchiri *et al.*, 2008). Tous ces gènes sont de potentiels gènes cibles de l'instabilité microsatellitaire. Les conséquences de cette instabilité sur les protéines issues de ces gènes sont une perte (totale ou partielle) de la fonction, un effet dominant négatif ou une haplo-insuffisance. Ces gènes sont impliqués dans différentes fonctions cellulaires, telles que la réparation de l'ADN (*MSH3*, *MSH6*, *MRE11A*), la régulation épigénétique (*HDAC2*, *ARID1A*), la signalisation (*TGFBR2*, *IGFR2*, *ACVR2A*), l'apoptose (*BAX*) et le traitement des ARNm (*TARBP2*, *XPO5*) (Baretti et Le, 2018 ; Duval et Hamelin 2002a , Duval et Hamelin, 2002b ; Yamamoto et Imai 2015). Il est intéressant de noter que tous les gènes MMR, à l'exception de *MLH1*, contiennent une ou plusieurs séquence(s) répétée(s) dans leur région codante (Chang *et al.*, 2001). C'est un paradoxe que le système MMR, dont un des rôles principaux est la réparation des erreurs de réplication dans les microsatellites, soit particulièrement vulnérable à ce type d'instabilité. Cette observation est à l'origine du concept du « mutateur qui mute

les autres mutateurs ». Les mutations dans ces gènes semblent subir une pression de sélection positive pendant la tumorigenèse (Malkhosyan *et al.*, 1996). Cette hypothèse a été confirmée récemment par notre équipe dans une étude portant sur l'identification de mutations (sur des séquences répétées et non répétées) qui sont positivement ou négativement sélectionnées dans les CCR MSI. Il a notamment été montré que le gène *MSH2* fait partie des 25 gènes les plus souvent porteurs d'une mutation sur une séquence codante non répétée dans les cancers MSI. Les séquences répétées codantes des gènes *MSH3* et *MSH6* sont également plus souvent mutées qu'attendu, au regard de la longueur et de la composition en paire de base de leur microsatellite (Jonchere *et al.*, 2018). Il est suggéré que ces mutations dans les gènes MMR provoquent une augmentation de l'instabilité génétique et accélèrent l'accumulation de mutations dans les gènes impliqués dans la cancérogenèse.

L'altération des microsatellites dans les régions intra-géniques (régions 5' et 3'-UTR, promoteurs, introns) peut également avoir des conséquences délétères sur l'expression correcte des gènes en influençant la transcription (Lewandowska, 2013) et la traduction (Chatterjee et Pal, 2009). En effet, ces séquences répétées peuvent jouer un rôle dans la régulation de la transcription et l'épissage du transcrit, puis dans la régulation de la traduction (localisation subcellulaire et stabilité du transcrit, reconnaissance du site d'initiation de la traduction...).

Les mutations dans les séquences répétées inter-géniques (nécessaires notamment à la fixation de protéines influençant la transcription et la conformation de la chromatine) et dans les régions non traduites (longs ARN non codants et micro-ARN) ont potentiellement un effet sur le taux d'expression des gènes et/ou la fidélité de la transcription.

L'inactivation du système MMR n'a pas des conséquences que sur les séquences répétées, elle peut aussi conduire à des mutations ponctuelles. Selon une approche de gène candidat, des altérations ont été recherchées dans les principales voies de transformation maligne des cellules épithéliales coliques. Par exemple, la mutation en transversion T>A au niveau du nucléotidique 1799 du gène *BRAF* provoque un changement de la valine en acide glutamique à la position 600 de la protéine (V600E). Cette mutation est retrouvée dans environ 60% des CCR MSI sporadiques alors que sa fréquence n'est que de 5% dans les CCR sporadiques MSS (Parsons *et al.*, 2012).

Il existe des différences dans les taux et les types de mutations des gènes selon que le cancer est MSI ou non, mais également entre différentes localisations tumorales MSI. Le gène *TGFBR2* (qui code pour une sérine/thréonine kinase) peut être utilisé pour illustrer cette idée. Ce gène contient un microsatellite sur l'exon 3. Dans environ 15% des CCR MSS, ce gène est muté par des mutations faux-sens dans le domaine kinase. Dans les cancers MSI, l'inactivation de ce gène ne se fait pas par mutation ponctuelle mais par altération du microsatellite. Cependant, la fréquence de mutation du

microsatellite de *TGFBR2* est différente selon le cancer MSI pris en considération : 90% pour les CCR MSI contre 5% des cancers de l'endomètre MSI (Biswas *et al.*, 2008 ; Cortes-Ciriano *et al.*, 2017).

## B. Le traitement des CCR

### 1. *Le traitement des CCR indépendamment de leur statut microsatellitaire*

Les protocoles de traitement des CCR incluent d'une part la chirurgie et d'autre part les traitements médicamenteux. Ces deux grands types de traitements peuvent être utilisés seuls ou en combinaison, en fonction du stade de la tumeur. Pour les tumeurs de stades précoces et localisés (I et II), la chirurgie seule est le principal traitement. La chirurgie est combinée à la chimiothérapie pour les tumeurs de stade II de haut grade et de stade III : on parle alors de chimiothérapie adjuvante, c'est-à-dire une chimiothérapie administrée à l'issue d'une chirurgie, pour prévenir le risque de rechute en détruisant les cellules cancéreuses qui n'auraient pas été retirées lors de la chirurgie. Enfin, les tumeurs de stade IV bénéficient de la chirurgie et de la chimiothérapie adjuvante, voir néo-adjuvante, c'est-à-dire une chimiothérapie administré préalablement à une intervention chirurgicale afin de réduire la taille de la tumeur.

Les protocoles de chimiothérapie dans le cadre du CCR utilisent 4 agents chimiothérapeutiques principaux : le 5-FluoroUracil (5-FU), l'oxaliplatine (OXA), l'acide folinique (leucovorine : LV) et l'irinotécan (IRI). L'activité anti-tumorale du 5-FU via l'induction de la mort cellulaire a été mise en évidence en 1957 (Heidelberger *et al.*, 1957). C'est une fluoropyrimidine métabolisée en plusieurs molécules actives capables d'inhiber l'activité de la thymidylate synthase (TS) impliquée dans la voie de biosynthèse de la thymidine. D'autre part, son incorporation au niveau de l'ADN et de l'ARN provoque l'arrêt des fourches de réplication et de transcription (Tajima *et al.*, 2004). La LV est un stabilisateur de l'interaction entre la TS et le 5-FU, potentialisant ainsi la cytotoxicité de celui-ci (Longley *et al.*, 2003). L'OXA induit des pontages inter- ou intra-brins entre deux guanines, ce qui empêche la réplication et la transcription (Dilruba et Kalayda, 2016). L'IRI est un inhibiteur de la topoisomérase I, conduisant à l'apoptose après induction de cassures simple brin et d'un blocage de la réplication (Hammond *et al.*, 2016). Ces 4 molécules peuvent être utilisées sous différentes combinaisons, notamment en fonction du stade de la tumeur. Le FOLFOX (association du 5-FU, d'OXA et de LV) est administrée aux patients atteints de CCR de stade II de haut grade et de stade III, alors que les patients atteints de CCR de stade IV reçoivent une chimiothérapie néo-adjuvante composée de FOLFOX et/ou de FOLFIRI (association de 5-FU, de LV et d'IRR).

Certains agents de thérapie personnalisée, ciblant une voie de signalisation dérégulée chez le patient, peuvent également être utilisés. C'est le cas des anticorps monoclonaux cetuximab, ciblant le

récepteur EGFR, et bevacizumab, ciblant le VEGF, qui peuvent être utilisés seuls ou en combinaison avec les molécules de chimiothérapie classiques pour traiter les tumeurs métastatiques (Cunningham *et al.*, 2004 ; Kabbinavar *et al.*, 2003).

Il est important de noter que les traitements décrits ci-dessus proposés aux patients atteints de CCR ne prennent pas en considération le statut microsatellitaire de la tumeur.

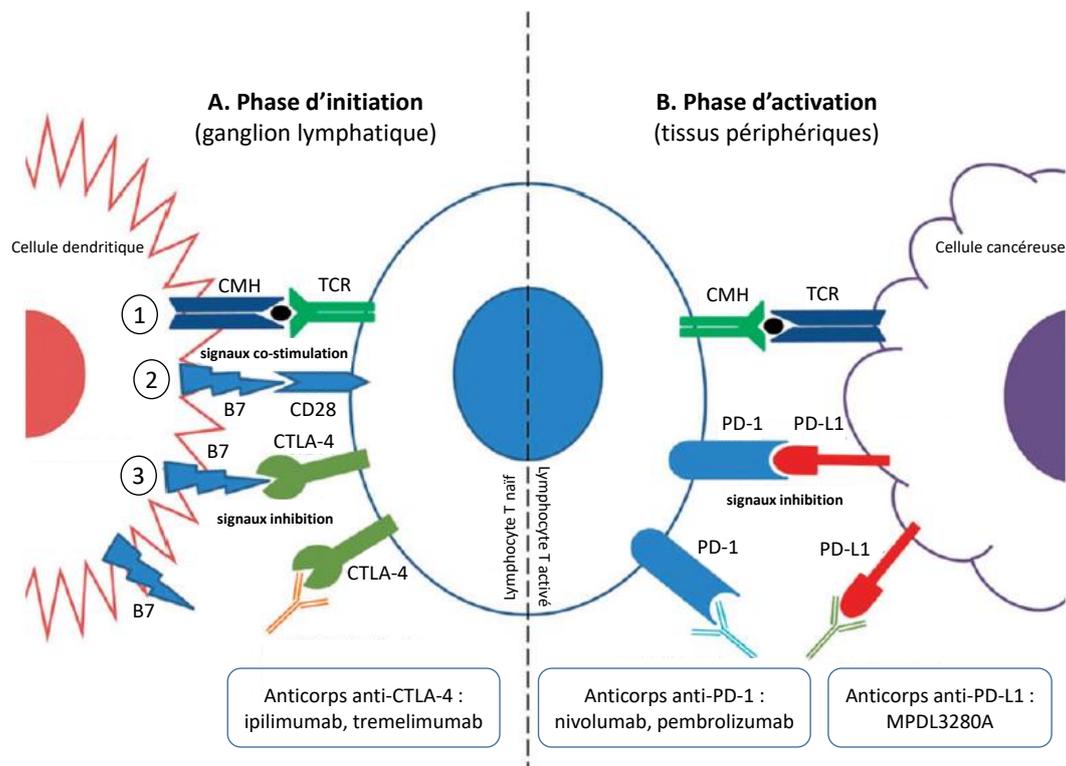
## ***2. L'émergence des immunothérapies pour le traitement des CCR MSI***

Les tumeurs MSI sont caractérisées par un nombre important de mutations dans les microsatellites codants, conduisant à un décalage du cadre de lecture, à l'apparition d'un codon stop prématuré au cours de la traduction et donc à l'expression de protéines tronquées aberrantes, nommées « néo-antigènes » (Linnebacher *et al.*, 2001). Ces néo-antigènes sont présentés à la surface des cellules tumorales via le complexe majeur d'histocompatibilité (CMH). Ils sont uniques à la tumeur et sont donc susceptibles d'être reconnus comme du non-soi par les lymphocytes T cytotoxiques (CTL, *Cytotoxic T Lymphocyte*) qui sont recrutés au niveau de la tumeur. Ceci explique le nombre important de TIL observés dans les tumeurs MSI, et donc l'inflammation, évoquée précédemment (Maby *et al.*, 2016).

Pour contrebalancer cette activation immunitaire importante, les cellules tumorales développent des stratégies d'évasion immune. Ainsi, les cellules cancéreuses MSI présentent entre autres une augmentation de l'expression des régulateurs négatifs du système immunitaire : les points de contrôle immun ou *immune checkpoints* (ICK). En contexte physiologique, ces régulateurs négatifs ont pour rôle d'éviter une réponse immunitaire excessive et donc dommageable. En contexte tumoral, les cellules cancéreuses se lient à la surface des CTL et y activent des inhibiteurs de la réponse lymphocytaire tels que PD-1 (*Programmed Death-1 receptor*), CTLA-4 (*Cytotoxic T Lymphocyte Antigen 4*), LAG-3 (*Lymphocyte Activation Gene 3*) ou TIM-3 (*T Cell Immunoglobulin Mucin 3*). Ce mécanisme est responsable du phénotype d'épuisement des lymphocytes T (**FIGURE 4**) (Colle *et al.*, 2017).

Des approches d'immunothérapie qui permettraient la réactivation fonctionnelle des TIL ont été développées, telles que l'utilisation d'inhibiteurs d'ICK. Ainsi, il a été montré dans une étude de phase I que l'utilisation d'anticorps monoclonaux anti-PD-1 (pembrolizumab) activait l'immunité anti-tumorale et était bénéfique aux patients atteints de cancers MSI métastatiques, CCR ou autres, alors que ce traitement n'avait aucun effet sur les patients atteints de cancers MSS de même stade (Le *et al.*, 2015). Des résultats similaires sont obtenus avec un autre anticorps anti-PD-1 (nivolumab) et un anticorps anti-CTLA-4 (ipilimumab). La présence d'un phénotype MSI est donc un marqueur de bon pronostic à la réponse aux anti-ICK et l'utilisation du pembrolizumab a été approuvée par la FDA (*Food*

and Drug Administration) pour les cancers MSI métastatiques, quelle que soit leur localisation (Killock, 2017).



**FIGURE 4 | Immunité anti-tumorale, checkpoint immunitaire (ICK) et thérapie anti-ICK**

**A.** ① Le lymphocyte T initie la réponse immunitaire par la reconnaissance par le TCR (*T Cell Receptor*) de peptides antigéniques présentés à la surface des cellules présentatrices d'antigènes via le Complexe Majeur d'Histocompatibilité (CMH). ② Cette reconnaissance n'est pas suffisante à l'activation du lymphocyte et des signaux de co-stimulation, médiés par l'interaction entre le récepteur CD28 du lymphocyte et le ligand B7 de la cellule dendritique, sont nécessaires pour sa maturation. ③ Le ligand B7 peut également interagir avec le récepteur CTLA-4 présent à la surface du lymphocyte, ce qui conduit à son inhibition. En contexte immunologique, ce mécanisme permet de moduler la réponse immunitaire afin que sa durée et son intensité ne soient pas dommageables pour l'organisme. **B.** Le couple récepteur PD-1 (lymphocyte) / ligand PD-L1 (cellule tumorale) suit le même mécanisme, pendant la phase d'activation dans les tissus périphériques. Les anticorps anti-CTLA-4 et anti-PD-1/PD-L1 bloquent l'interaction entre les récepteurs et les ligands, empêchant donc la signalisation d'inhibition de la réponse immunitaire. Les lymphocytes activés reconnaissent à nouveau les cellules tumorales et engagent la fonction de défense contre ces dernières. *Adapté de (Singh et al., 2015).*

Les cancers dus à une déficience du système MMR ne sont pas les seuls à présenter une charge néo-antigénique importante et une réponse immunitaire anti-tumorale conséquente, c'est aussi le cas des cancers dus à des mutations dans les gènes *POLE* et *POLD1*. Ces gènes codent pour des sous-unités catalytiques de la polymérase responsables de l'activité de relecture, et donc gardiennes de la fidélité de la réplication. Des mutations somatiques dans ces gènes conduisent à l'émergence de cancers hyper-mutés, qui présentent certaines caractéristiques en commun avec les cancers MSI (*Domingo et al., 2016*). Ces patients, qui ont une expression importante des ICK PD-1, PD-L1 et CTLA-4, pourraient

être des candidats potentiels aux immunothérapies anti-ICK (Bourdais *et al.*, 2017). Ainsi, le « *Tumor Mutational Burden* » (charge mutationnelle tumorale) constituerait un meilleur marqueur de réponse aux anti-ICK que le phénotype MSI (Fabrizio *et al.*, 2018).

### ***3. Le phénotype tumoral MSI : valeur pronostique et prédictive***

Comme nous l'avons vu précédemment, 12 à 15% de l'ensemble des CCR, tous stades confondus, sont de phénotype MSI. Cependant, la prévalence du phénotype MSI diffère selon le stade de la tumeur : environ 20% des CCR de stade II sont MSI, 12% au stade III et environ 4% au stade IV (Goldstein *et al.*, 2014). Cette fréquence qui diminue quand le stade augmente indique la faible agressivité et la faible capacité des CCR MSI à développer des métastases. Les CCR MSI de stade précoce sont aussi de meilleur pronostic (meilleure survie globale et sans récurrence) comparé à un CCR MSS de stade similaire (Gryfe *et al.*, 2000 ; Popat *et al.*, 2005).

Par contre, les CCR MSI lorsqu'ils sont au stade métastatique peuvent présenter un pronostic moins bon que les CCR MSS au même stade (Smith *et al.*, 2013). Ceci peut être corrélé à la présence de mutations *BRAF*, détectées à des taux plus importants dans les CCR MSI métastatiques (35%) que dans les CCR MSI à des stades plus précoces (24%) (Baretti *et al.*, 2018) et qui constitue donc un facteur de mauvais pronostic dans le contexte métastatique (Venderbosch *et al.*, 2014).

Nous avons vu que les CCR MSI ont des caractéristiques moléculaires et clinico-pathologiques différentes des CCR MSS. Il n'est donc pas illogique de penser que les CCR MSI pourraient également présenter une différence de réponse aux traitements. Or, nous avons également constaté que le traitement par chimiothérapie est similaire, que le CCR soit MSI ou MSS (en dehors des nouvelles techniques d'immunothérapies développées très récemment).

Plusieurs études suggèrent que les CCR MSI sont associés à une résistance au 5-FU par rapport aux tumeurs MSS (Ribic *et al.*, 2003 ; Sargent *et al.*, 2010). De façon intéressante, une étude de 2011 indique qu'au sein des patients atteints de CCR MSI, les patients atteints du syndrome de Lynch ont une meilleure survie sans récurrence après traitement au 5-FU que les patients atteints d'un CCR MSI sporadique (Sinicrope *et al.*, 2011). Cette corrélation entre la réponse au 5-FU et le statut MSI a également été mise en évidence *in vitro*. En effet, des lignées cellulaires issues de CCR MSI sont plus tolérantes au 5-FU que des lignées cellulaires issues de CCR MSS (Carethers *et al.*, 1999). Il en va de même pour des fibroblastes embryonnaires humains MMR-déficients et compétents (Meyers *et al.*, 2001). Une lignée initialement MSI en raison d'une hyperméthylation du promoteur de *MLH1* peut être traitée par un agent déméthylant afin de sensibiliser les cellules au 5-FU (Arnold *et al.*, 2003).

A l'heure actuelle, la détermination du statut MSI a donc 4 implications majeures. Premièrement, la présence d'un phénotype tumoral MSI permet de suggérer un syndrome de Lynch : le diagnostic de ce syndrome génétique permet de proposer un suivi particulier aux individus atteints, qui sera détaillé dans le chapitre 3. Ensuite, comme nous l'avons vu précédemment, les patients ayant un CCR MSI ont un meilleur pronostic (au moins aux stades précoces) que les individus atteints d'un CCR MSS mais présentent une moins bonne réponse au 5-FU : le phénotype MSI a donc une valeur pronostique et un rôle prédictif. Enfin, la présence d'un phénotype MSI a également un impact thérapeutique, puisqu'il influence le choix d'un traitement anti-ICK, traitement révolutionnaire des tumeurs MSI (Ryan *et al.*, 2017).

### C. Détermination du statut microsatellitaire d'une tumeur

Deux méthodes complémentaires permettent d'analyser le statut MSI d'une tumeur : l'immunohistochimie (IHC) et la réaction de polymérisation en chaîne (PCR). Ces deux techniques sont mises en œuvre sur des échantillons biologiques issus de biopsies ou de pièces chirurgicales. Actuellement, ces deux méthodes sont réalisées au service d'anatomo-pathologie et de biologie médicale de l'hôpital Saint-Antoine, de façon systématique sur tous les CCR, en parallèle et en routine.

#### 1. Approche immunohistochimique

L'analyse du statut microsatellitaire par IHC est une méthode indirecte puisqu'elle permet de mettre en évidence la perte d'expression des protéines MMR (MLH1, MSH2, MSH6 et PMS2) plutôt que l'instabilité des microsatellites que cette perte engendre. L'IHC est réalisée avec des anticorps dirigés contre ces protéines sur coupe tissulaire. A l'état physiologique, l'expression des protéines MMR est ubiquitaire et localisée dans le noyau. Dans la tumeur, les cellules normales de la muqueuse adjacente ou infiltrant la tumeur (telles que les cellules immunitaires ou les cellules des vaisseaux sanguins) constituent des contrôles positifs internes. L'analyse se fait par comparaison du marquage entre les cellules tumorales et les cellules normales (**FIGURE 5A**).

Les protéines MMR agissent en hétérodimères : MLH1-PMS2 et MSH2-MSH6. MLH1 et MSH2 sont stables en absence de PMS2 et MSH6, respectivement. L'inverse n'est pas vrai : PMS2 et MSH6 sont instables en absence de leur partenaire. Ainsi, une perte d'expression de MLH1 est toujours associée à une perte d'expression de PMS2 (idem pour MSH2 et MSH6), alors qu'une perte de PMS2 est isolée (idem pour MSH6). La majorité des CCR MSI étant des CCR sporadiques dus à une hyperméthylation du gène *MLH1*, la perte protéique la plus fréquemment observée est une perte conjointe de MLH1 et PMS2. A l'inverse, une perte conjointe de MSH2 et MSH6 ou une perte isolée de PMS2 ou MSH6 sont plus suggestifs d'un syndrome de Lynch. Le marquage avec l'anticorps anti-MLH1

est difficile d'interprétation et nécessite la lecture par un pathologiste expérimenté. L'utilisation conjointe de l'anticorps anti-PMS2 permet de simplifier l'interprétation, puisqu'une perte de PMS2 est presque systématiquement concomitante à une perte de MLH1 (De Jong *et al.*, 2004).

L'IHC présente de nombreux avantages puisqu'elle est facile à mettre en place, rapide et peu onéreuse et peut donc être exploitée en routine. Cette technique présente une sensibilité de 92% (Shia, 2008). De plus, l'IHC donne une indication sur le gène impliqué dans la tumorigenèse MSI du patient. En cas de suspicion d'un syndrome de Lynch, ces informations permettent d'orienter l'analyse génétique vers un gène donné.

Cependant, plusieurs limitations sont à noter. Premièrement, le marquage peut porter à confusion dans certains cas, tels qu'un marquage cytoplasmique au lieu de nucléaire. L'interprétation des résultats d'IHC nécessite donc une certaine expérience. De plus, le marquage n'est pas toujours uniforme et la réalisation de l'IHC sur des petits échantillons (tels que des biopsies) peut manquer de sensibilité (Shia, 2008). Par ailleurs, le défaut de sensibilité peut également être dû aux mutations faux-sens qui provoquent l'inactivation fonctionnelle de la protéine sans affecter son expression, sa stabilité et sa capacité à être reconnue par l'anticorps (Baretti et Le, 2018). Des mutations tronquantes et de larges délétions en phase dans le gène *MLH1* peuvent également conduire à un marquage normal en IHC, sans que le mécanisme sous-jacent soit compris (Shia, 2008). Dans les deux cas, il s'agit de faux-négatifs, puisqu'il y a marquage de la protéine alors qu'il y a effectivement une déficience du système MMR.

## 2. *Approche moléculaire*

Le phénotype MSI est caractérisé par une altération de la longueur des microsatellites dans la tumeur. Ainsi, l'analyse du phénotype MSI peut être réalisée grâce à une PCR qui amplifie des répétitions microsatellitaires non-codantes spécifiques (**FIGURE 5B**).

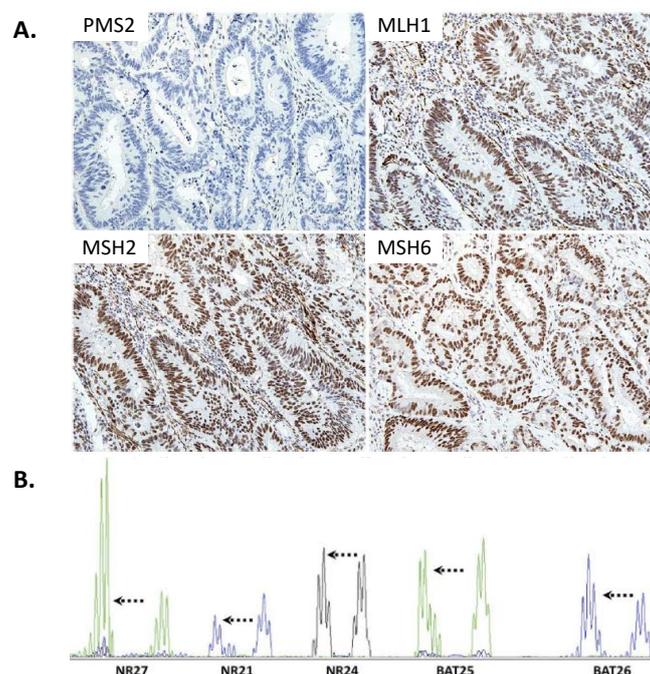
Initialement, chaque équipe utilisait des marqueurs différents, ce qui avait pour conséquence une grande variabilité dans la fréquence rapportée du phénotype MSI, même au sein d'un même type tumoral. Pour pallier à ce biais, un panel consensus de 5 marqueurs mononucléotidiques, développé par notre équipe, a été proposé (NR-21, NR-22, NR-24, BAT-25 et BAT-26) (Buhard *et al.*, 2004; Suraweera *et al.*, 2002). Une tumeur est considérée comme MSI si au moins 2 des 5 marqueurs présentent une variation de taille. Cette technique, connue sous le nom de « pentaplex », constitue le test de référence pour analyser le phénotype MSI en biologie moléculaire, même si beaucoup de laboratoires utilisent encore d'autres panels.

Ce panel de 5 marqueurs mononucléotidiques présente deux avantages. Comme ces marqueurs sont quasi-monomorphes dans la population mondiale, l'analyse des résultats ne nécessite

pas de comparaison avec l'ADN germlinal du patient. D'autre part, le choix s'est porté sur des marqueurs mononucléotidiques plutôt que dinucléotidiques car les premiers sont plus sensibles et spécifiques, en particulier chez les patients ayant une tumeur extra-colique et dans le cas des cancers liés à une mutation dans *MSH6* (Verma *et al.*, 1999). En effet, *MSH6* n'est pas impliqué dans la réparation des microsatellites dinucléotidiques et les tumeurs *MSH6*-déficiantes peuvent donc apparaître faussement stables si on utilise ces marqueurs.

L'étude du phénotype MSI par biologie moléculaire est une expérience très reproductible entre différents laboratoires (Zhang, 2008) et présente une sensibilité de 93% (Shia, 2008).

Cependant, dans certains tissus, l'instabilité peut ne pas être détectée à cause de difficultés techniques comme une trop faible quantité d'ADN tumoral (Hampel *et al.*, 2005). Par ailleurs, les mutations dans le gène *MSH6* provoquent un phénotype MSI moins marqué en biologie moléculaire (Zhang, 2008). Cela s'explique par la redondance de fonction entre les protéines *MSH6* et *MSH3*. Le dimère MutS $\alpha$  (*MSH2-MSH6*) reconnaît les mésappariements ainsi que les petites boucles d'insertion/délétion alors que le complexe MutS $\beta$  (*MSH2-MSH3*) reconnaît l'ensemble des boucles d'insertion/délétion. Ainsi, en cas de déficience de *MSH6*, MutS $\beta$  peut prendre le relais et cacher la perte d'expression de *MSH6*.



**FIGURE 5 | Détermination du statut MSI des tumeurs**

**A.** Immunohistochimie des 4 protéines MMR : exemple d'une perte d'expression isolée de PMS2 dans un CCR, avec conservation des protéines MLH1, MSH2 et MSH6. **B.** Amplification des 5 microsatellites de la pentaplex par PCR puis migration sur capillaire et détection grâce à la fluorescence des amorces. Dans la majorité des cas, l'instabilité des microsatellites se traduit par un raccourcissement des séquences répétées dans l'ADN tumoral par rapport à l'ADN normal (symbolisé par la flèche). *Adapté de (Mojtahed et al., 2011 ; Boland et Goel, 2010).*

Les performances du diagnostic MSI par biologie moléculaire et par ICH sont comparables. Cependant, chaque technique a ses avantages et ses inconvénients et les deux méthodes fournissent des informations complémentaires.

Récemment, notre équipe a suggéré l'étude du microsatellite T17 présent au niveau d'une borne d'épissage du gène HSP110 comme nouveau marqueur diagnostique du phénotype MSI (Buhard *et al.*, 2016). Dans cette étude, il a été démontré que ce marqueur présente un degré de monomorphisme encore plus important que les marqueurs de la pentaplex. En effet, seulement deux allèles sauvages (T16 et T17) ont été mis en évidence dans 1 037 échantillons d'individus de 54 populations issues de 29 régions géographiques à travers le monde (cohorte du CEPH, Centre d'Etude du Polymorphisme Humain). Ce marqueur présente une meilleure sensibilité et une spécificité identique par rapport à la pentaplex. Ainsi, cette technique serait plus simple puisque nécessitant un seul marqueur au lieu de 5.

## D. Cancers MSI iatrogènes

### 1. *Cancers induits par les thiopurines*

Les thiopurines comprennent l'azathioprine (Aza), la 6-mercaptopurine et la 6-thioguanine (6-TG). Les métabolites issus de la transformation des thiopurines agissent selon deux modes d'action : (i) l'incorporation dans les acides nucléiques qui provoque la cytotoxicité et (ii) l'inactivation de la voie de synthèse *de novo* des purines qui occasionne une immunosuppression. Par un processus tout à fait similaire à celui de la signalisation des dommages à l'ADN induits par les agents méthylants, les thiopurines incorporées dans l'ADN induisent également la mort cellulaire suite à leur prise en charge par le système MMR. Le mécanisme de tolérance aux thiopurines implique que le traitement provoque une pression de sélection pour les rares cellules MMR-déficientes qui présentent un avantage sélectif par rapport aux cellules MMR-compétentes. Par ailleurs, la voie de synthèse *de novo* des purines est la voie principale d'obtention des précurseurs puriques dans les lymphocytes : le traitement aux thiopurines entraîne donc un arrêt de la prolifération lymphocytaire. Ce mécanisme peut être mis à profit pour le traitement des leucémies mais également pour induire une immunosuppression, par exemple chez les transplantés d'organes ou encore chez les personnes atteintes de maladies inflammatoires (telles que les maladies inflammatoires chroniques de l'intestin) ou immunitaires (polyarthrite rhumatoïde) (Karran et Attard, 2008).

Les leucémies aiguës myéloblastiques (LAM) représentent une complication fréquente chez les patients transplantés d'organes ayant reçu un traitement immunosuppresseur pour prévenir un rejet de la greffe. Dans une étude de 7 LAM développées par des transplantés, toutes étaient de phénotype

MSI, contre 0/28 des LAM développées *de novo*. De plus, ce résultat est significativement corrélé à la prise d'Aza (Offman *et al.*, 2004).

Une étude du laboratoire a montré que les lymphomes non Hodgkiniens (LNH) présentent un phénotype MSI dans seulement 2% (12/603) des cas. De manière intéressante, la totalité des cas sont observés dans un contexte d'immunosuppression, soit chez des patients porteurs du VIH (Virus de l'Immunodéficience Humaine) (3/128) soit chez des patients transplantés ayant reçu un traitement immunosuppresseur au long cours (9/111) (Duval *et al.*, 2004). Une autre étude réalisée sur 143 LNH développés par des patients transplantés a mis en évidence 13 LNH MSI. Sur l'ensemble des 143 LNH, la prise ou non d'Aza comme traitement immunosuppresseur était connu pour les 13 LNH MSI et pour 20 LNH MSS. Parmi les 13 patients ayant développé un LNH MSI, 11 ont pris de l'Aza ; contre 9/20 chez les patients avec un LNH MSS. Par ailleurs, une association significative a été observée entre la perte d'expression de la MGMT dans les cellules tumorales des LNH et la présence d'un phénotype MSI (Borie *et al.*, 2009).

Les Maladies Inflammatoires Chroniques de l'Intestin (MICI) sont des lésions inflammatoires chroniques atteignant le tractus digestif et les patients atteints sont susceptibles de développer des CCR, dont la fréquence de MSI varie entre 1 et 45% selon les études (Aust *et al.*, 2005 ; Brentnall *et al.*, 1996 ; Cawkwell *et al.*, 2000 ; Fleisher *et al.*, 2000 ; Lyda *et al.*, 2000 ; Schulmann *et al.*, 2005 ; Suzuki *et al.*, 1994 ; Willenbacher *et al.*, 1999). Les gènes cibles de l'instabilité microsatellitaire diffèrent entre les CCR MSI développés dans le cadre d'une MICI et les CCR MSI sporadiques (Schulmann *et al.*, 2005). Notre équipe a décrit une série de 277 néoplasies intestinales développées par 205 patients atteints de MICI : 17 patients ont développé 27 néoplasies MSI (17/205, 8,3%). Dans cette étude, une association entre le développement de CCR MSI et la prise d'Aza, utilisée comme traitement d'entretien des MICI, n'a pas pu être mise en évidence (Svrcek *et al.*, 2016).

Enfin, des études de notre équipe, menées *in vivo* chez la souris, ont permis de mettre en évidence une corrélation entre la prise d'Aza et le développement de tumeurs MSI. Une cohorte constituée de souris de phénotype *Msh2*<sup>+/+</sup>, *Msh2*<sup>+/-</sup> et *Msh2*<sup>-/-</sup> ont reçu de l'Aza ; alors qu'une cohorte contrôle n'a reçu aucun traitement. Les souris *Msh2*<sup>-/-</sup>, traitées ou non à l'Aza, développent des lymphomes MSI. Les souris MMR-compétentes *Msh2*<sup>+/+</sup> et *Msh2*<sup>+/-</sup> contrôles sont asymptomatiques tandis que toutes les souris de même génotype traitées à l'Aza meurent prématurément. Alors que toutes les souris *Msh2*<sup>+/-</sup> développent des lymphomes MSI, les souris *Msh2*<sup>+/+</sup> développent des lymphomes MSS. Chez les souris *MSH2*<sup>+/+</sup>, la cytotoxicité de l'Aza à la dose utilisée dans cette étude n'a pas permis de statuer sur le risque oncogénétique (Chalastanis *et al.*, 2010). Ainsi, une deuxième étude a été menée avec des doses plus faibles. La dose la plus faible permet la survie prolongée des souris *Msh2*<sup>+/+</sup> et est associée à l'émergence de lymphomes dans une fraction significative de souris (7/23), dont une (1/7) a développé un lymphome MSI. De la cyclosporine A, un immunosuppresseur

dont la signalisation est indépendante du système MMR, a été administrée en parallèle à des souris des trois génotypes. Ce traitement à la cyclosporine A n'entraîne pas d'augmentation de l'induction de lymphomes MSI, indiquant que l'Aza, contrairement à la cyclosporine A, conduit à un processus oncogénétique MSI chez la souris (Bodo *et al.*, 2015).

## 2. *Cancers induits par les agents méthylants*

Les cancers MSI peuvent se développer dans un contexte d'immunosuppression et d'inflammation, associé ou non à une prise de thiopurine ; mais ils peuvent également se développer de façon secondaire à la suite d'un traitement chimiothérapeutique d'un premier cancer, utilisant notamment des agents méthylants.

Les LAM secondaires (liées à un traitement) présentent un phénotype MSI dans environ 50% des cas alors que la fréquence du phénotype MSI dans les LAM *de novo* est estimée à moins de 5%. Les thérapies mises en cause sont principalement la chimiothérapie et dans une moindre mesure les rayons (Karran *et al.*, 2003). Il a également été observé qu'un polymorphisme dans le domaine de liaison à l'ADN de MSH2, susceptible de réduire l'efficacité du système MMR, est particulièrement associé aux LAM secondaires à une chimiothérapie avec un agent méthylant (Worrillow *et al.*, 2003).

Un autre exemple est le traitement des gliomes par le témozolomide, un agent méthylant. Des mutations somatiques dans le gène *MSH6* ont été mises en évidence dans des récurrences de glioblastomes après traitement au témozolomide, alors que la tumeur primitive ne présentait pas ces anomalies. Cela suggère que les mutations de *MSH6* permettent une résistance clinique au témozolomide (Cahill *et al.*, 2007 ; Hunter *et al.*, 2006).

Une étude de 2003 suggère que l'expansion clonale des cellules MMR-déficientes serait favorisée par un traitement aux agents méthylants. La moelle osseuse de souris irradiées est reconstituée avec un mélange de cellules *Msh2*<sup>-/-</sup> et *Msh2*<sup>+/+</sup>. Quand les souris sont traitées au témozolomide, la proportion de cellules MMR-déficientes est augmentée de façon significative. Par contre, en absence de pression de sélection, la proportion de cellules MMR-déficiente décroît au fur et à mesure des transplantations sérielles, démontrant par la même occasion le désavantage sélectif des cellules MMR-déficientes dans les conditions physiologiques normales (Reese *et al.*, 2003).

Les cellules MMR-déficientes sont résistantes à certains agents de chimiothérapie mais la déficience du système MMR peut également être acquise pendant la chimiothérapie par des mutations sur les gènes MMR (Li, 2008). Ainsi le risque de la chimiothérapie est double. D'abord pour les patients avec des tumeurs MSI, pour qui le traitement va affecter seulement les cellules MMR-compétentes fortement répliquatives (telles que les cellules sanguines), ce qui va conduire à une forte toxicité sans

effet sur la tumeur. Ensuite, l'utilisation de ces agents peut sélectionner des cellules MMR-déficientes dans une tumeur par ailleurs MMR-compétente, ce qui peut provoquer des cancers secondaires MSI.

## Chapitre 3 : Syndromes génétiques prédisposant au développement de tumeurs MSI

Les tumeurs MSI peuvent se développer de façon sporadique, principalement par hyperméthylation du promoteur du gène *MLH1*, comme nous l'avons vu dans le chapitre précédent. Par ailleurs, l'inactivation constitutionnelle mono- ou bi-allélique d'un gène MMR constitue un contexte génétique de prédisposition au développement de tumeurs MSI dans le cadre de deux syndromes : le syndrome de Lynch (LS) et le syndrome de Déficience MMR Constitutionnelle ou CMMRD (*Constitutional MMR Deficiency*), respectivement. Le développement des cancers MSI sporadiques nécessite l'inactivation somatique des deux allèles du gène *MLH1* et survient vers 65 ans en moyenne. Au contraire, les patients atteints du syndrome de Lynch (patients LS) sont déjà porteurs d'une mutation hétérozygote germinale d'un gène MMR dans l'ensemble de leurs cellules : le développement du cancer ne nécessite qu'un seul événement mutationnel somatique, d'après le modèle des « deux hits » de Knudson (Knudson, 1971), ce qui explique l'âge de diagnostic plus précoce (42 ans, en moyenne) (Jacob et Praz, 2002). Enfin, les patients atteints du syndrome CMMRD présentent une mutation bi-allélique d'un gène MMR. Ils naissent donc avec un système MMR déficient et développent par conséquent des tumeurs dès l'enfance. Ces deux syndromes de prédisposition au développement de tumeurs MSI seront détaillés dans ce chapitre.

### A. Le Syndrome de Lynch

Le LS est, avec le syndrome sein-ovaire dû à une mutation dans les gènes *BRCA1* et *BRCA2*, le syndrome de prédisposition aux cancers le plus fréquent. Leur prévalence respective est de 1 personne sur 250 à 1000 (Tanakaya, 2018) et 1 personne sur 400 à 500. Le LS est responsable de nombreux cancers : il est la cause de 1 à 4% de l'ensemble des CCR (Moreira *et al.*, 2012 ; Peltomäki, 2016) et de 2% des cancers de l'endomètre (Hampel *et al.*, 2006).

#### 1. Description clinique

En 1895, Aldred Warthin décrit la famille de sa couturière (famille G) qui présente de nombreux carcinomes du côlon, gastriques et de l'endomètre à des âges précoces. Une publication parue en 1913 regroupe les arbres généalogiques de 3 familles très bien documentées, dont la famille G, et propose une transmission héréditaire de la susceptibilité au cancer (Warthin, 1985). En 1966, Henry T. Lynch décrit deux autres grandes familles comportant une agrégation de cancers du côlon, de l'endomètre et de l'estomac. Il démontre le mode de transmission autosomique dominant et il différencie ce syndrome de la Polypose Adénomateuse Familiale (PAF) (Lynch *et al.*, 1966). La PAF était

à cette époque le seul syndrome de prédisposition au cancer colorectal décrit. Elle est due à une mutation dans le gène *APC* (*Adenomatous Polyposis Coli*) et les patients atteints présentent des centaines voire des milliers de polypes coliques, contrairement à ce qu'on observe dans le LS. En 1992, le gène *APC* est exclu comme responsable du LS (Peltomäki *et al.*, 1992). En 1993, le phénotype MSI est d'abord détecté dans des CCR sporadiques, et très rapidement après on découvre son association avec le LS (Aaltonen *et al.*, 1993 ; Peltomäki *et al.*, 1993). Suite à cela, les gènes responsables, impliqués dans le système MMR de réparation des mésappariements induits par la polymérase lors de la réplication, sont mis en évidence : *MSH2* (Fishel *et al.*, 1993 ; Leach *et al.*, 1993), *MLH1* (Lindblom *et al.*, 1993 ; Papadopoulos *et al.*, 1994), *PMS2* (Nicolaidis *et al.*, 1994) et *MSH6* (Miyaki *et al.*, 1997).

Ce syndrome a été initialement appelé HNPCC, pour *Hereditary Non Polyposis Colorectal Cancer* (cancer colorectal héréditaire sans polypose), en opposition à la PAF. Mais quand il a été clairement démontré que le LS prédispose au développement de nombreuses tumeurs autres que les CCR, l'appellation HNPCC ne semblait plus appropriée et a été changée. Par ailleurs, la cause génétique de ce syndrome (mutation hétérozygote dans un gène MMR) est maintenant connue et il n'est plus nécessaire d'inclure dans sa dénomination une référence négative à la PAF (Boland, 2005).

Le spectre tumoral associé au LS comprend effectivement les cancers colorectaux et les cancers de l'endomètre mais également d'autres cancers gastro-intestinaux (estomac, intestin grêle, voies biliaires, pancréas) et non gastro-intestinaux (voies urinaires, ovaires) (Tricarico *et al.*, 2017). Au-delà de 70 ans, le risque d'un individu atteint de LS d'avoir développé un cancer du côlon est de 80%, 20-60% pour l'endomètre, 11-19% pour l'estomac et 9-11% pour l'ovaire (tous gènes MMR et sexes confondus). Les risques sont respectivement de 5,5%, 2,7%, <1% et 1,6% dans la population générale (Hsieh et Yamane, 2008) (**FIGURE 6**). Il est intéressant de noter que chez les femmes LS, le risque de développer un cancer de l'endomètre au cours de la vie est plus important que le risque de développer un CCR (Baretti et Le, 2018).

Le LS présente deux variantes dont les patients atteints développent, en plus des tumeurs classiquement associées au LS, d'autres types de cancers. Le syndrome de Muir-Torre décrit en 1981 et caractérisé par des adénomes sébacés et d'autres tumeurs de la peau est principalement associé à des mutations de *MSH2* (Lynch *et al.*, 2015). Le syndrome de Turcot inclut des tumeurs du système nerveux central. Il est dû à une mutation pathogène dans un gène MMR ou dans le gène *APC*. Quand il est associé à une mutation dans un gène MMR, le nombre de polypes est moindre et la tumeur cérébrale est souvent un glioblastome alors que quand il est associé à une mutation dans le gène *APC*, le nombre de polypes est plus important et la tumorale cérébrale est souvent un médulloblastome (Hamilton *et al.*, 1995).

Afin de détecter le plus tôt possible un éventuel cancer, une surveillance est proposée aux patients LS dès le diagnostic. La prise en charge du risque de CCR consiste en une coloscopie, à réaliser tous les deux ans dès l'âge de 20 ans puis tous les ans après 40 ans. Pour les femmes, la surveillance est complétée par une échographie endovaginale tous les deux ans à partir de 30 ans (INCa 2009). L'intérêt de la surveillance n'a été démontré que pour le CCR (Järvinen *et al.*, 2000). Le taux de CCR détecté est réduit de 60% chez les individus LS qui suivent un protocole de surveillance, différence qui peut être expliquée par l'élimination des adénomes à un stade précoce. De plus, les CCR détectés chez les patients suivis sont plus souvent locaux et l'individu a donc plus de chances de guérison.

En plus de la surveillance, qui permet de détecter le cancer de façon précoce, la chirurgie préventive est possible pour éliminer l'organe susceptible de se transformer et ainsi prévenir le développement de tumeurs. La chirurgie prophylactique est envisageable pour éviter les cancers gynécologiques une fois que le projet parental a été rempli : hystérectomie (ablation de tout ou partie de l'utérus) pour prévenir le cancer de l'endomètre et salpingo-oophorectomie bilatérale (ablation des ovaires et des trompes de Fallope) pour prévenir le cancer de l'ovaire (Schmeler *et al.*, 2006). Par contre, la chirurgie prophylactique n'est pas recommandée pour le côlon car la surveillance par endoscopie est une mesure préventive efficace. Cependant, si une chirurgie doit être réalisée pour éliminer un CCR, une colectomie totale ou une résection segmentaire est réalisée, sachant que le risque de CCR métachrone (deuxième tumeur rapidement diagnostiquée après la première) est plus grand après une résection segmentaire (25%) qu'après une colectomie totale (8%). Néanmoins, il n'y a pas d'évidence claire qu'une chirurgie plus étendue confère un bénéfice en terme de survie ; alors que le risque de diarrhée chronique et d'incontinence est plus élevé (Boland *et al.*, 2018). Le choix doit donc être discuté en fonction de l'âge du patient et de son souhait.

Si la surveillance permet de diminuer le risque de CCR, la chimioprévention est également envisageable. En effet, l'étude CAPP2 réalisée sur 937 individus LS a montré que la prise régulière d'aspirine (600 mg par jour pendant au moins 2 ans) entraînait une réduction du risque de CCR de 63% au long terme (recul de 56 mois), même après que le traitement ait été interrompu (Burn *et al.*, 2011). L'étude CAPP3 est actuellement en cours sur 3 000 individus LS pour tester différentes doses d'aspirine (100, 300 et 600 mg par jour) (Burn *et al.*, 2013). A l'heure actuelle, la prise d'aspirine n'est pas proposée systématiquement à tous les individus LS, notamment car elle expose au risque de complications hémorragiques (hémorragies digestives), ce qui pose la question de la balance bénéfice-risque. Par conséquent, il n'existe pas de consensus sur l'indication d'une chimioprévention par aspirine.

Bien que les CCR MSI sporadiques ou associés au LS se développent tous suite à une inactivation du système MMR, il existe des différences notables entre les deux contextes. Le risque de tumeurs synchrones ou métachrones est plus élevé dans le contexte du LS (35% des cas) (Vasen *et al.*,

1999). Certains microsatellites présents dans des séquences codantes ont un taux de mutation comparable (*FLT3LG*, *TEAD2* et *BLM*) dans les CCR MSI sporadiques ou associés au LS alors que d'autres ont des taux de mutations différents (*CHK1*, *B2M*, *Axin2*). Ainsi, bien que les tumeurs MSI associées au LS ou sporadiques présentent des similarités, leurs modes de développement pourraient être distincts (Poulogiannis *et al.*, 2010).

S'il existe des différences entre les CCR MSI sporadiques et associés au LS, il en existe également entre les différents types de cancers du spectre Lynch. En effet, les profils MSI des cancers de l'endomètre et des CCR associés au LS sont différents. Les CCR présentent une instabilité prédominante sur les locus non-codants *BAT*, sur *TGFR11*, sur les répétitions de dinucléotides, sur *MSH3* et sur *BAX*. Au contraire, l'instabilité de *PTEN* est significativement associée aux cancers de l'endomètre (Kuismanen *et al.*, 2002). Cela implique que la tumorigenèse des cancers colorectaux et de l'endomètre passe par des voies différentes même si elles sont toutes les deux dues à l'instabilité des microsatellites (Baretti et Le, 2018).

## 2. Description moléculaire

### i. Mutations dans les gènes MMR

De nos jours, la définition du LS est moléculaire et non plus clinique et le diagnostic définitif repose sur la détection, dans l'ADN germlinal, d'une mutation mono-allélique pathogène dans un des gènes MMR. Les gènes affectés sont *MLH1* (dans 50% des cas), *MSH2* (40%), *MSH6* (7%) et *PMS2* (2%) (**FIGURE 6**).

Les altérations détectées peuvent être de grands réarrangements génomiques (grandes insertions ou délétions, duplications ou inversions) ou des altérations plus locales (petites insertions/délétions en phase ou substitutions). Ce sont les substitutions qui constituent les altérations les plus fréquentes puisqu'elles représentent 63% des variations détectées sur le gène *MLH1*, 55% sur *MSH2*, 58% sur *MSH6* et 70% sur *PMS2*. A l'heure actuelle, plus de 3800 altérations différentes des gènes MMR ont été répertoriées sur la base de données internationale InSiGHT, base de données des variants des gènes qui contribuent aux cancers gastro-intestinaux. Il n'a pas été décrit de point chaud de mutations. Cependant, une cinquantaine de mutations fondatrices ont été mises en évidence (Ponti *et al.*, 2015). Par ailleurs, le taux de mutations *de novo* est faible : la majorité des individus porteurs d'une mutation dans un gène MMR ont hérité cette mutation d'un de leurs parents. Cependant, à cause de la pénétrance incomplète, de l'âge variable au développement du cancer, de la réduction du risque de cancer grâce à la surveillance ou à une chirurgie prophylactique, ou une mort précoce, tous les patients LS n'ont pas obligatoirement un parent qui a développé un cancer.

Les mutations germinales dans les gènes MMR n'expliquent pas l'ensemble des cas de syndromes de Lynch. En effet, certains cas sont dus à des épimutations constitutionnelles des promoteurs des gènes *MLH1* et plus rarement *MSH2*, c'est-à-dire une répression transcriptionnelle due à une hyperméthylation du promoteur (Chan *et al.*, 2006 ; Gazzoli *et al.*, 2002 ; Hitchins et Ward, 2009). La transmission de la méthylation est rare et ne suit pas la loi de Mendel. Chez les patients présentant une méthylation constitutionnelle de *MLH1*, la survenue de cancers est plus précoce que chez les patients ayant une mutation de *MLH1* (Niessen *et al.*, 2009). Par ailleurs, une large délétion de la région 3' du gène *EPCAM* (Epithelial Cell Adhesion Molecule), localisé en amont du gène *MSH2*, peut également aboutir à l'inactivation épigénétique de *MSH2*, et donc être responsable d'un LS (Ligtenberg *et al.*, 2009).

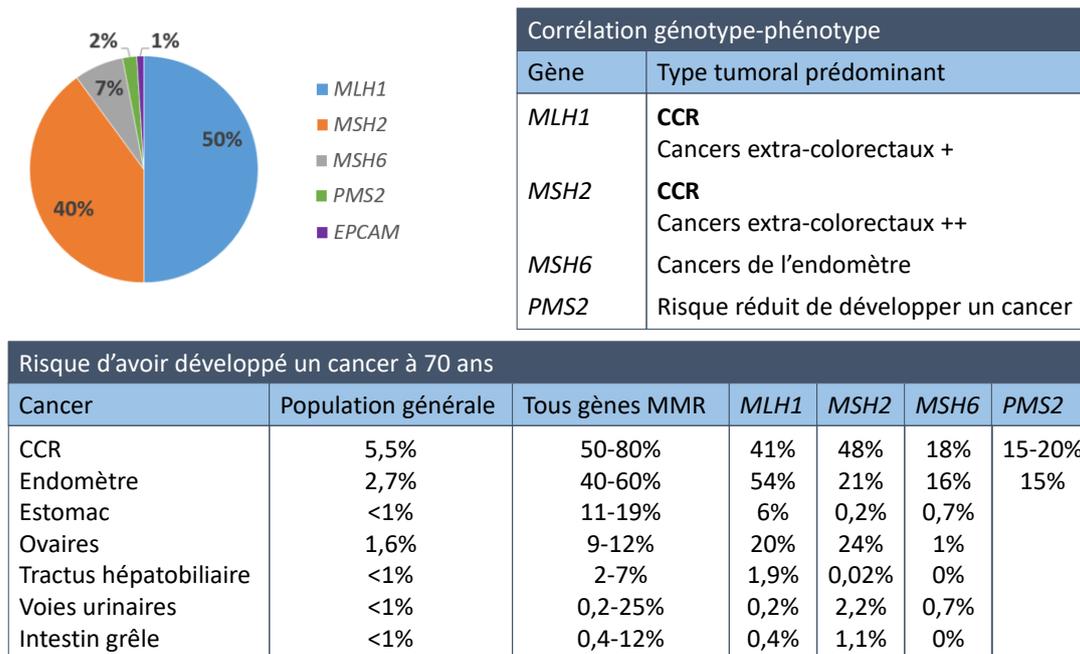
Quel que soit le type d'altération génétique ou épigénétique constitutionnelle portée par l'individu, le développement de cancers nécessite l'inactivation de l'allèle sauvage. Différents mécanismes sont théoriquement possibles : délétion somatique, hyperméthylation du promoteur, mutation ponctuelle (Potocnik *et al.*, 2001) ou encore par conversion génique ou recombinaison mitotique. Pour le gène *MLH1*, les pertes d'hétérozygoties sont le mécanisme le plus fréquent, suivies par les mutations ponctuelles (Porkka *et al.*, 2017).

Par ailleurs, certains individus développent des cancers de phénotype MSI (déterminé par biologie moléculaire ou immunohistochimie) sans hyperméthylation du promoteur du gène *MLH1*, et sans mutation germinale mono-allélique d'un gène MMR. Cette condition est appelée syndrome « *Lynch-like* » car elle imite un LS (développement de tumeurs MSI à un âge jeune) sans que le diagnostic moléculaire puisse être confirmé. Une mutation somatique d'un allèle d'un gène MMR couplée à une perte d'hétérozygotie est le mécanisme le plus fréquent ; mais deux mutations somatiques sur les deux allèles d'un même gène MMR sont également possibles (Carethers et Stoffel, 2015 ; Geurts-Giele *et al.*, 2014 ; Haraldsdottir *et al.*, 2014 ; Mensenkamp *et al.*, 2014). Enfin, un cas de mutation *de novo* mosaïque a été mis en évidence dans un CCR MSI développé à 79 ans. Cette femme ne présentait pas de mutation MMR dans les cellules sanguines. Par contre, une mutation mono-allélique a été détectée dans la muqueuse colique normale adjacente à la tumeur. Cette mutation constitutionnelle, associée à une mutation somatique sur le deuxième allèle, a provoqué le CCR MSI. La mutation constitutionnelle a été transmise à son fils, qui a développé un CCR MSI à l'âge de 54 ans et chez qui la mutation a été détectée dans les cellules sanguines, confirmant le diagnostic de LS chez cet individu (Sourrouille *et al.*, 2013).

ii. Corrélation génotype-phénotype

Le LS présente une hétérogénéité de génotypes (gène atteint, type de mutation et mécanisme d'inactivation de l'allèle sauvage) mais également une hétérogénéité de phénotypes. Il existe une différence de spectre clinique et d'âge au diagnostic entre des individus LS porteurs de mutations dans des gènes MMR différents : une corrélation génotype-phénotype peut donc être mise en évidence. Cependant, il peut également exister des différences de phénotypes entre les individus d'une même famille ou entre deux familles qui portent une même mutation (variabilité intra et inter-familles). Enfin, des altérations différentes sur le même gène peuvent avoir des conséquences phénotypiques différentes. Ces trois points seront détaillés ci-dessous.

Le spectre clinique du LS n'est pas tout à fait le même selon le gène impliqué. Les patients LS porteurs d'une mutation dans le gène *MLH1* sont particulièrement sensibles au développement de cancers colorectaux (à un âge précoce), et dans une moindre mesure de localisations extra-colorectales (Lin *et al.*, 1998). Une mutation dans le gène *MSH2* provoque également des CCR mais aussi un risque accru pour les localisations extra-colorectales. Les patients qui ont une mutation sur le gène *MSH6* ont plus de risque de développer un cancer de l'endomètre que colorectal, et les cancers sont plus tardifs. De plus, le phénotype MSI est plus difficile à mettre en évidence. En effet, la redondance fonctionnelle partielle entre MutS $\alpha$  et MutS $\beta$  peut expliquer l'âge au diagnostic plus tardif (de Wind *et al.*, 1999). Enfin, les patients avec une mutation dans le gène *PMS2* présentent le risque le plus bas de développer un cancer du spectre de Lynch ( Lynch *et al.*, 2015) (**FIGURE 6**).



**FIGURE 6 | Le syndrome de Lynch : gènes impliqués, corrélation génotype-phénotype et risque de développement des différentes tumeurs du spectre Lynch** (Bonadona et al., 2011 ; Giardiello et al., 2014 ; Jasperson et al., 2010 ; Lynch et al., 2015).

Comme évoqué précédemment, le LS a une prévalence de 1 personne sur 250 à 1000 et les mutations des gènes *MLH1* et *MSH2* sont largement plus fréquentes que les mutations des gènes *MSH6* et *PMS2*. Il a été suggéré récemment que les LS associés à *MSH6* et *PMS2* seraient considérablement sous-estimés à cause de leur faible pénétrance et à leur spectre tumoral qui s'éloignerait du spectre LS classique (Espenschied *et al.*, 2017). Historiquement, l'analyse génétique est en effet réalisée chez les individus qui présentent une histoire clinique personnelle et/ou familiale évocatrice du LS, ce qui peut biaiser les résultats vers les variants d'un gène provoquant une pénétrance forte du syndrome au détriment des variants d'un autre gène qui occasionnerait une pénétrance plus faible. Ainsi, si ce sont les individus porteurs d'une mutation sur *MLH1* ou *MSH2* qui développent plus de tumeurs associées au spectre de Lynch et sont donc diagnostiqués comme tels, ce sont les individus porteurs d'une mutation sur *MSH6* ou *PMS2* qui seraient les plus nombreux. En 2016, une analyse constitutionnelle des gènes MMR réalisée sur 5 744 cas de CCR non sélectionnés a rapporté une prévalence globale du LS de 1 personne sur 279 (*MLH1* : 1 sur 1 946, *MSH2* : 1 sur 2 841, *MSH6* : 1 sur 758, *PMS2* : 1 sur 714) (Win *et al.*, 2017). Ainsi, cela laisse suggérer que la fréquence du LS serait plus importante qu'estimée et que la répartition serait beaucoup plus homogène entre les différents gènes, mais que les individus porteurs d'une mutation dans le gène *PMS2* seraient moins souvent au courant de leur condition. Il est estimé qu'aux Etats-Unis, seuls 5% des porteurs de mutations dans un gène MMR sont avertis de leur situation (Lynch *et al.*, 2018).

Par ailleurs, la variabilité phénotypique intra- et inter-familiale peut refléter l'influence de l'environnement et du contexte génétique (Stone et Sidow, 2005). Par exemple, la « famille G » atteinte du LS initialement décrite par Warthin présentait principalement une agrégation de cancers gastriques et de l'endomètre, alors que les générations suivantes présentent principalement des CCR (Boland et Lynch, 2013). Ce changement corrèle avec la diminution de l'incidence des cancers gastriques et l'augmentation de celle des CCR dans la population générale occidentale (Vasen *et al.*, 2013). En outre, les patients LS originaires d'Asie de l'Est (Corée du Sud, Japon) développent plus de cancers gastriques que les patients LS occidentaux (Park *et al.*, 1999), ce qui est en corrélation avec le nombre plus élevé de cancers gastriques que de CCR dans la population générale de ces pays. Enfin, un individu LS en surpoids et/ou fumeur a plus de risque de développer une tumeur alors que la consommation de fruits permettrait de diminuer le risque (Vasen *et al.*, 2013).

Dans le cas d'un LS dû à une substitution conduisant au changement d'un seul acide aminé, la sévérité de la maladie peut être liée à plusieurs paramètres. D'une part, il est probable que les altérations faux-sens affectent l'expression et la fonctionnalité de la protéine à des degrés variables, ce qui peut contribuer au phénotype variable de la maladie observé dans des familles affectées sur le même gène mais avec des altérations différentes (Perera et Bapat, 2008). D'autre part, le contexte génétique (altérations dans d'autres gènes – MMR ou non – susceptibles de moduler l'effet du variant étudié) et l'environnement peuvent également jouer un rôle.

### ***3. Diagnostic du syndrome de Lynch***

On a vu que si les individus LS sont susceptibles de développer différents cancers, la surveillance et la chirurgie prophylactique permettent de réduire la survenue de ces cancers, et donc la mortalité associée. Ainsi, il est important d'identifier ces individus afin de leur proposer des mesures préventives. Ce paragraphe s'intéressera aux critères disponibles pour identifier les individus potentiellement atteints du LS et les techniques mises en œuvre pour confirmer le diagnostic.

#### *i. Les critères cliniques*

Les critères d'Amsterdam II ont été écrits en 1999 (Vasen *et al.*, 1999). Ils ont initialement été rédigés avant la découverte du phénotype MSI et se basent sur des critères cliniques pour identifier les patients suspectés d'être atteints du LS. Trois apparentés doivent présenter une tumeur du spectre de Lynch, deux générations successives doivent être affectées et au moins un des apparentés doit être diagnostiqué avant 50 ans (règle « 3-2-1 ») (Richman, 2015) (**TABLE 2**). Ils sont très spécifiques mais peu sensibles. Par ailleurs, les critères révisés de Bethesda, rédigés en 2004 (Umar *et al.*, 2004) (**TABLE 2**), prennent en compte, en plus de l'histoire personnelle et familiale de cancer, des caractéristiques

clinico-pathologiques spécifiques aux cancers associées au spectre de Lynch. Ils permettent d'identifier les patients qui ont une forte probabilité d'avoir une tumeur de type MSI et pour lesquels une détermination du statut microsatellitaire devrait être effectuée. Malheureusement, les critères d'Amsterdam et de Bethesda ne permettent pas d'identifier à coup sûr les personnes atteintes du LS. En effet, jusqu'à 50% des porteurs de mutations ne remplissent pas les critères d'Amsterdam (manque de sensibilité) alors qu'il s'avère que les tumeurs de 40-45% des familles qui les remplissent ne présentent pas de phénotype MSI (manque de spécificité). Les critères de Bethesda sont plus sensibles, mais échouent quand même à mettre en évidence 12 à 28% des cas de LS (Ryan *et al.*, 2017). Ainsi, les recommandations de Jérusalem de 2009 proposent que tous les CCR développées avant 70 ans ainsi que les CCR développés après 70 ans chez un individu remplissant un des critères de Bethesda soient testés pour le phénotype MSI (Balmaña *et al.*, 2013 ; Boland et Shike, 2010) (**FIGURE 7**). Ces nouvelles directives devraient notamment permettre de détecter plus de patients porteurs de mutations sur les gènes *MSH6* ou *PMS2*, qui ont un âge au diagnostic de la première tumeur généralement plus tardif et qui remplissent donc plus rarement les critères (Ryan *et al.*, 2017). En France, il est recommandé que les tests somatiques recherchant une déficience du système MMR (par biologie moléculaire et/ou IHC) soient réalisés systématiquement chez toute personne ayant développé un CCR avant 60 ans ou un cancer de l'endomètre avant 50 ans (INCa 2016).

Critères d'Amsterdam II
Au moins trois apparentés du cas index sont atteints d'un cancer du spectre LS*
<ol style="list-style-type: none"> <li>1. L'un des trois individus doit être apparenté au premier degré avec les autres</li> <li>2. Au moins deux générations successives sont concernées</li> <li>3. Au moins un des cas de cancer du spectre LS doit être diagnostiqué avant 50 ans</li> <li>4. La polypose adénomateuse familiale (FAP) doit être exclue dans les cas de CCR</li> <li>5. Le diagnostic histologique des tumeurs doit être vérifié</li> </ol>
Critères de Bethesda révisés
Le statut MSI des tumeurs doit être analysé dans les cas suivants :
<ol style="list-style-type: none"> <li>1. CCR diagnostiqué avant 50 ans</li> <li>2. Présence de tumeurs du spectre LS synchrones et métachrones quel que soit l'âge</li> <li>3. CCR avec une histologie évocatrice** du phénotype MSI, diagnostiqué avant 60 ans</li> <li>4. Cancer du spectre LS diagnostiqué chez au moins un parent au premier degré avant 50 ans</li> <li>5. Cancer du spectre LS diagnostiqué chez au moins deux parents de premier ou deuxième degré quel que soit l'âge</li> </ol>

**TABLE 2 | Critères d'Amsterdam II et de Bethesda révisés**

\*Le spectre de Lynch comprend les cancers colorectaux, endomètre, estomac, ovaire, voies urinaires, intestin grêle, voies hépatobiliaires. \*\* L'histologie évocatrice fait référence à la présence de lymphocytes intra-tumoraux, réaction inflammatoire de type Crohn, différenciation de type mucineux ou en bague à chaton ou architecture médullaire. *Adapté de ( Lynch et al., 2015 ; Umar et al., 2004 ; Vasen et al., 1999).*

ii. Mise en évidence du phénotype MSI

Quand un individu remplit les critères nécessaires, une détermination du statut microsatellitaire est effectuée sur la tumeur (par biologie moléculaire et/ou IHC), telle que décrite précédemment (**FIGURE 5**). Cependant, un individu peut être atteint du LS même si sa tumeur ne présente pas de phénotype MSI détecté par biologie moléculaire. En effet, la protéine MSH6 n'est pas impliquée dans la réparation des grandes boucles d'insertion/délétion et une mutation dans le gène *MSH6* conduit donc à une instabilité des microsatellites moindre. De même, un individu peut être atteint du LS même si sa tumeur présente un marquage positif des protéines MMR en IHC. En effet, un marquage positif est preuve de l'expression et non pas de la fonctionnalité d'une protéine (Sijmons *et al.*, 2013) : une protéine porteuse d'une substitution d'un seul acide aminé peut être exprimée au même taux qu'une protéine sauvage tout en n'étant plus capable de remplir sa fonction.

La présence d'un phénotype MSI associé à une perte d'expression de la protéine MLH1 ne suffit pas à poser le diagnostic du LS, étant donné qu'une inactivation épigénétique du gène *MLH1* peut provoquer un cancer MSI sporadique. Des caractéristiques génétiques et épigénétiques supplémentaires doivent être recherchées pour tenter de différencier un cancer MSI sporadique ou associé au LS. En effet, les mutations *BRAF* V600E sont présentes dans 60% des CCR MSI sporadiques (Capper *et al.*, 2013) et sont mutuellement exclusives des mutations *KRAS*, qui elles, sont plus souvent retrouvées dans les cancers MSI associés au LS. Ainsi la présence d'une mutation *KRAS* sera suggestive d'un cancer MSI associé au LS alors qu'une mutation *BRAF* sera suggestive d'un cancer MSI sporadique. Par ailleurs, une hyperméthylation du promoteur du gène *MLH1* fera pencher la balance en faveur d'un cancer MSI sporadique (**FIGURE 7**).

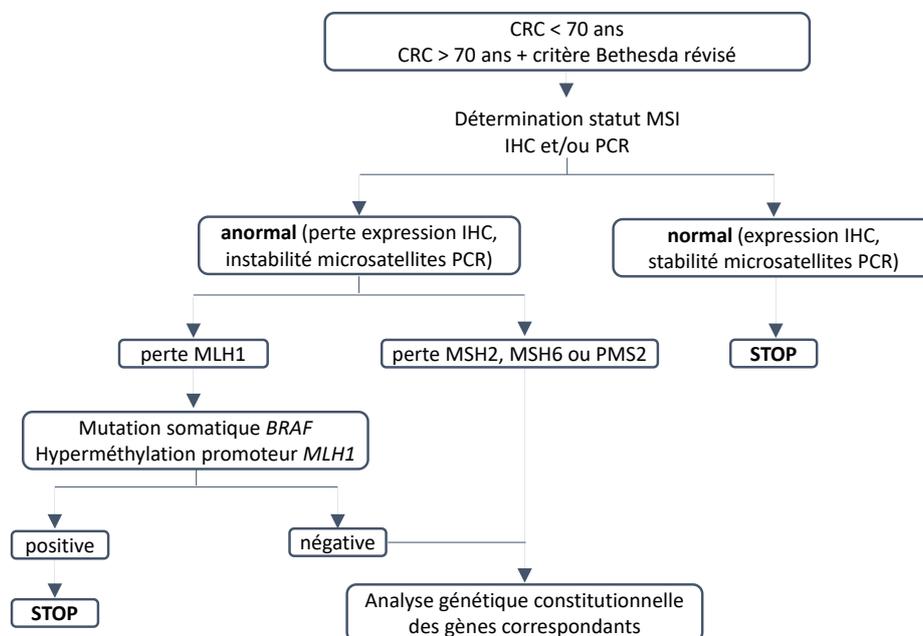
iii. Analyse constitutionnelle des gènes MMR

Si le phénotype MSI et la suspicion d'un LS sont effectivement confirmés, le patient pourra bénéficier d'une consultation d'oncogénétique et la recherche d'une mutation pathogène constitutionnelle sera réalisée pour poser le diagnostic de façon définitive (**FIGURE 7**).

En France en 2015, le délai moyen d'obtention d'un premier rendez-vous en consultation d'oncogénétique pour un cas index est de 12 semaines et l'analyse complète des gènes MMR dure 27 semaines. Cette analyse moléculaire est basée sur l'identification d'une altération génétique dans un gène MMR : la détection des mutations ponctuelles et des petites insertions/délétions se fait par séquençage et la mise en évidence des grands réarrangements se fait le plus souvent par MLPA (*Multiplex Ligation-dependent Probe Amplification*). En France, une quinzaine de laboratoires sont habilités à réaliser la recherche des mutations par séquençage, dont 3 laboratoires sont en mesure de

réaliser les études sur le gène *PMS2*. En effet, le gène *PMS2* présente de nombreux pseudo-gènes qui peuvent en compliquer l'analyse (De Vos *et al.*, 2004 ; Nakagawa *et al.*, 2004). Une délétion d'*EPCAM* ou une épimutation constitutionnelle des promoteurs de *MLH1* ou *MSH2* peuvent également être recherchées.

De manière classique, le séquençage est réalisé par des analyses séquentielles qui commencent par le(s) gène(s) le(s) plus susceptible(s) d'être muté(s). Les gènes *MLH1* et *MSH2* sont responsables à eux deux de 90% des cas de LS ; mais le test génétique peut également être orienté par les résultats de l'IHC. Cependant, le séquençage successif de gènes uniques est une démarche longue. Avec les avancées des techniques de NGS (*New Generation Sequencing*, séquençage de nouvelle génération), le diagnostic génétique utilisant un panel multi-gènes est disponible depuis 2013 (Tanakaya, 2018). Il permet un taux de détection des variants importants, une rentabilité en terme de temps et une clarification de la relation génotype-phénotype. Par exemple, les phénotypes du syndrome de Lynch et du syndrome sein-ovaire se chevauchent partiellement. Une étude a testé des individus suspectés d'être atteints du syndrome de Lynch et a trouvé 114 patients effectivement atteints du syndrome de Lynch mais également 71 individus avec des variants dans d'autres gènes de prédisposition au cancer, dont 15 dans les gènes *BRCA1* ou *BRCA2* (Yurgelun *et al.*, 2015). Cependant, l'utilisation d'un panel multi-gènes ne résout pas le problème de la détection des mutations sur le gène *PMS2*.



**FIGURE 7 | Algorithme du diagnostic moléculaire du syndrome de Lynch**

L'analyse constitutionnelle des gènes MMR est longue et coûteuse. Celle-ci n'est donc réalisée que sur les individus dont les CCR présentent un statut MSI avec une absence de mutation somatique sur le gène *BRAF* et une absence d'hyperméthylation du promoteur du gène *MLH1*. Adapté de (Balmaña *et al.*, 2013).

Quand une altération pathogène a été détectée chez le cas index, les apparentés peuvent également bénéficier d'une consultation d'oncogénétique : le délai d'attente est alors de 7 semaines et le test dure 6 semaines. Ce temps est plus court que pour le cas index car il s'agit d'un test ciblé qui ne s'intéresse qu'à la présence ou l'absence de l'altération détectée chez le cas index et non pas d'une analyse de toute la séquence de l'ensemble des gènes impliqués dans le LS. En France en 2015, 1003 nouvelles personnes (cas index et apparentés) ont été diagnostiquées comme atteintes du LS (INCa 2015). Le diagnostic d'un LS chez un cas index permet en moyenne de mettre en évidence la mutation chez 3 apparentés (Vasen *et al.*, 2013).

## B. Le syndrome CMMRD

Le syndrome de Lynch, dû à une mutation mono-allélique dans un gène MMR, prédispose au développement de tumeurs principalement à partir de la quatrième décennie de vie. Le CMMRD (*Constitutionnal Mismatch Repair Deficiency*) est une variante rare du syndrome de Lynch, due à une mutation bi-allélique d'un gène MMR qui prédispose au développement de tumeurs dès l'enfance. Les aspects cliniques et moléculaires ainsi que les méthodes de diagnostic de ce syndrome seront détaillés dans cette partie.

### 1. Description clinique

Le CMMRD est un syndrome rare qui se transmet selon le mode autosomique récessif. Il prédispose au développement, dès l'enfance, de 3 types principaux de cancers : (i) des hémopathies malignes, (ii) des tumeurs du cerveau et du système nerveux central et (iii) des tumeurs gastro-intestinales ainsi que toutes les autres tumeurs du spectre Lynch (qui représentent souvent la deuxième voir la troisième tumeur développée par les patients). Beaucoup de patients présentent des signes rappelant la neurofibromatose de type I, principalement des taches café au lait (Vasen *et al.*, 2014). Cependant, les taches café au lait développées par les patients CMMRD sont légèrement différentes de celles associées à la neurofibromatose de type I car elles ont des bords irréguliers et elles peuvent varier dans leur degré de pigmentation (Wimmer et Etzler, 2008). Les patients CMMRD peuvent également développer d'autres caractéristiques non tumorales, telles qu'une hypopigmentation de la peau, une déficience modérée en immunoglobuline et des malformations congénitales.

Les hémopathies malignes et les tumeurs cérébrales sont principalement développées pendant la première décennie de vie, à un âge moyen au diagnostic de 6 ans (0,4-21 ans) et de 9 ans (2-40 ans) respectivement. Les cancers gastro-intestinaux et les éventuels autres cancers du spectre Lynch, quant à eux, sont plutôt développés pendant la deuxième et la troisième décennie à un âge

moyen de 17 ans (8-48 ans) (**FIGURE 8**). Une grande proportion des patients CMMRD (environ 40%) développent des tumeurs métachrones, soit beaucoup plus fréquemment que dans le LS. A l'heure actuelle, à peine plus de 200 patients atteints de CMMRD sont décrits dans le monde. Ce syndrome rare est encore méconnu et la plupart des cas de CMMRD n'ont été suspectés qu'après que le patient ait développé un deuxième cancer ou lorsqu'un frère ou une sœur s'est avéré(e) également atteint(e) d'un cancer (Rana et Syngal, 2017). Contrairement aux patients LS, qui ne développent pas de polypose, une oligopolyse est souvent identifiée entre 5 et 10 ans chez les patients CMMRD, ce qui peut conduire à un diagnostic erroné de Polypose Adénomateuse Familiale (Durno *et al.*, 2017).

Bien que des cas de CMMRD aient été décrits depuis la fin des années 1950, la description du syndrome tel qu'on le comprend aujourd'hui a été réalisée par deux équipes indépendantes en 1999. Dans chacune de ces études est décrite une grande famille LS au sein de laquelle des mariages consanguins ont conduit à la naissance d'enfants porteurs de mutations bi-alléliques germinales dans le gène *MLH1* qui ont développées les affections décrites ci-dessus (Ricciardone *et al.*, 1999 ; Wang *et al.*, 1999). Selon plusieurs études récentes, le nombre d'enfants CMMRD issus d'unions consanguines varie entre 39 et 67% (Aronson *et al.*, 2016 ; Durno *et al.*, 2015 ; Lavoine *et al.*, 2015 ; Ripperger et Schlegelberger, 2016).

Comme les patients LS, les patients CMMRD peuvent profiter d'une surveillance pour diagnostiquer le plus précocement possible un éventuel cancer. Une coloscopie annuelle est recommandée à partir de 6 ans, voire tous les 6 mois si des polypes ont été identifiés. Un hémogramme est recommandé tous les 6 mois à partir de 1 an pour détecter les hémopathies malignes. Concernant les tumeurs cérébrales, la surveillance par imagerie par résonance magnétique est recommandée tous les 6 mois à partir de 2 ans. La surveillance du cancer de l'endomètre commence à 20 ans (Durno *et al.*, 2017 ; Rana et Syngal, 2017) (**FIGURE 8**). A l'heure actuelle, on ne sait pas si le bénéfice de ce protocole de surveillance est supérieur au fardeau émotionnel et physique qu'il engendre (Herkert *et al.*, 2011).

## ***2. Description moléculaire***

Le CMMRD est dû à une mutation bi-allélique constitutionnelle homozygote (deux mutations identiques sur les deux allèles du gène) ou hétérozygote composite (deux mutations différentes sur les deux allèles) dans l'un des gènes également impliqués dans le syndrome de Lynch : *PMS2* (affecté dans 60% des cas), *MSH6* (20%), *MLH1* (10%) ou *MSH2* (10%) (Wimmer *et al.*, 2014) (**FIGURE 8**).

On peut constater que le gène le plus fréquemment muté dans le CMMRD est *PMS2*, contrairement au LS où il s'agit du gène le moins fréquemment muté. Deux hypothèses peuvent être avancées pour expliquer cette différence.

1) Le LS et le syndrome CMMRD associés aux gènes *MLH1* et *MSH2* étant plus sévères que ceux associés aux gènes *MSH6* et *PMS2*, il est possible que les patients CMMRD porteurs d'une mutation bi-allélique sur les gènes *MLH1* et *MSH2* soient plus rarement diagnostiqués comme étant atteints du syndrome. En effet, le diagnostic n'a malheureusement pas toujours le temps d'être posé avant qu'ils ne meurent, étant donné que ces patients décèdent souvent de leur première tumeur. Ainsi, la proportion de patients CMMRD due au gène *PMS2* pourrait être surestimée. Par ailleurs, vu la sévérité des mutations dans les gènes *MLH1* et *MSH2*, il a été proposé que des mutations bi-alléliques sur ces gènes seraient plus souvent létales à l'état embryonnaire comparé aux mutations des gènes *MSH6* et *PMS2* (Wimmer et Etzler, 2008).

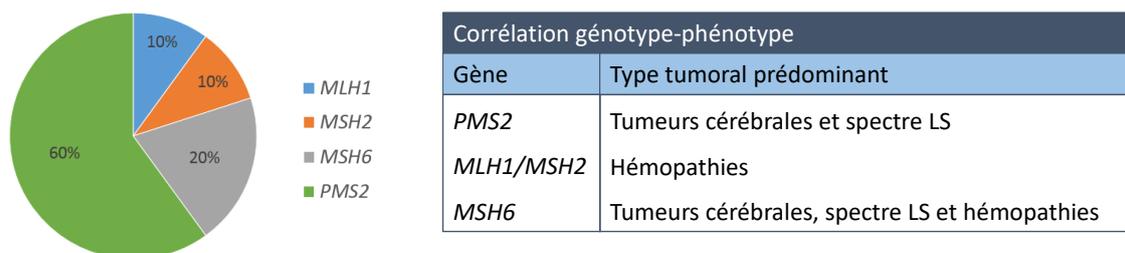
2) La pénétrance du LS est beaucoup plus faible chez les patients porteurs de mutations hétérozygotes dans le gène *PMS2* : moins d'individus développent des cancers donc moins d'individus sont suspectés d'être atteints du LS. Il n'y aurait donc pas forcément moins de porteurs de mutations hétérozygotes dans le gène *PMS2* que dans les autres gènes MMR, mais simplement moins de personnes diagnostiquées comme telles (Truninger *et al.*, 2005).

Etant donné qu'il s'agit d'une maladie à transmission récessive, les deux parents sont porteurs obligatoires d'une mutation dans le même gène MMR. Cette mutation peut être la même (dans le cas d'une union consanguine ou d'une mutation fondatrice) ou elle peut être différente. Cependant, l'histoire familiale d'un patient CMMRD est rarement évocatrice d'un LS. En effet, plus de la moitié des patients CMMRD présentent une mutation du gène *PMS2* qui, à l'état hétérozygote, induit un LS avec une pénétrance faible et un âge au diagnostic tardif. Par conséquent, les patients CMMRD ont rarement des parents qui ont développé des cancers et encore moins fréquemment des parents qui sont au courant de leur condition. Ainsi, les médecins doivent avoir une forte suspicion ou une grande connaissance de la maladie pour envisager le diagnostic du CMMRD devant un enfant se présentant avec une hémopathie maligne ou une tumeur cérébrale et des taches café au lait rappelant une neurofibromatose de type I (Rana et Syngal, 2017).

Comme dans le LS, il existe une certaine corrélation génotype-phénotype dans le CMMRD. Le CMMRD associé au gène *PMS2* prédispose principalement au développement de tumeurs cérébrales et de tumeurs du spectre de Lynch ; et la moyenne d'âge au diagnostic de la première tumeur est de 9 ans. Par ailleurs, le nombre d'enfants qui survivent à leur première tumeur est plus important, ce qui les rend plus susceptibles de développer plusieurs tumeurs au cours de leur vie. Les cas de CMMRD dus à des mutations dans le gène *MLH1* et le gène *MSH2* sont caractérisés par une incidence accrue d'hémopathies malignes et moins de tumeurs du spectre de Lynch. Malheureusement, ces patients développent leur première tumeur à un âge particulièrement précoce (moyenne de 3,5 ans) et sont

plus susceptibles de décéder de leur premier cancer. Enfin, le CMMRD dû au gène *MSH6* présente un phénotype intermédiaire : des tumeurs cérébrales et des tumeurs du spectre de Lynch à un âge légèrement plus tardif (similaire à *PMS2*) mais également des hémopathies malignes ; et une plus grande probabilité de mourir des suites de la première tumeur (similaire à *MLH1/MSH2*) (Lavoine *et al.*, 2015 ; Wimmer et Etzler, 2008) (**FIGURE 8**).

Par ailleurs, certains patients CMMRD ne développent que des tumeurs du spectre de Lynch, à l'adolescence ou à l'âge adulte. Chez ces patients, la fréquence des mutations faux-sens est plus élevée que celle observée dans l'ensemble des patients CMMRD. Ces mutations faux-sens conduisent à l'expression d'une protéine qui peut avoir conservé une partie de sa fonctionnalité, ce qui expliquerait le phénotype CMMRD « atténué » chez ces patients (Wimmer et Etzler, 2008). Un phénotype similaire a été mis en évidence dans des familles Inuits porteuses d'une mutation homozygote fondatrice dans le gène *MSH2* conduisant à un épissage aberrant, qui permet tout de même une expression résiduelle de l'ARN sauvage (Li *et al.*, 2015).



Cancer	Âge moyen diagnostic	Surveillance
Hémopathie maligne	5-8 ans	Hémogramme tous les 6 mois dès 1 ans
Tumeurs cérébrales	9 ans	IRM cerveau tous les 6 mois dès 2 ans
CCR	16 ans	Coloscopie annuelle dès 6 ans (tous les 6 mois si polypes)
Intestin grêle	28 ans	Endoscopie digestive annuelle dès 8 ans
Endomètre	19-44 ans	Surveillance dès 20 ans

**FIGURE 8 | Le syndrome CMMRD : gènes impliqués, corrélation génotype-phénotype et cancers développés par les patients**

Spectre LS : cancer colorectal, endomètre, estomac, intestin grêle, voies biliaires, voies urinaires, ovaires. (Durno *et al.*, 2017 ; Lavoine *et al.*, 2015 ; Wimmer *et al.*, 2014 ; Wimmer et Etzler, 2008).

Comme indiqué précédemment, les patients CMMRD présentent souvent des taches café au lait qui évoquent une autre pathologie : la neurofibromatose de type 1, maladie génétique fréquente à transmission autosomique dominante due à une mutation mono-allélique sur le gène *NF1*. La présence des taches café au lait chez les patients CMMRD pourrait être consécutive à une mutation somatique précoce du gène *NF1* due à l'inactivation constitutionnelle du système MMR. En effet, le gène *NF1* est un gène particulièrement long qui contient de nombreuses séquences microsatellitaires et pourrait donc être un gène cible de l'instabilité microsatellitaire. Cette hypothèse a été vérifiée par

l'équipe de Wang qui a mis en évidence une mutation du gène *NF1* dans 5 lignées cellulaires cancéreuses MMR-déficientes sur 10 alors qu'aucune mutation n'a été détectée dans 5 lignées MMR-compétentes (Wang *et al.*, 2003). Cette hypothèse est corroborée en clinique par la présence, chez un patient, d'une distribution des taches café au lait sur une seule moitié du corps, impliquant un mosaïcisme (Auclair *et al.*, 2007).

Les tumeurs développées par les patients CMMRD ont certaines caractéristiques qui les différencient des mêmes tumeurs développées dans un contexte sporadique ou dans un contexte de LS. Les tumeurs du spectre CMMRD, et en particulier les tumeurs cérébrales, sont connues pour présenter un phénotype exceptionnellement ultra-hyper-muté (*versus* « hyper-muté » pour les tumeurs MSI « classiques »), dû à des mutations somatiques précoces dans les gènes *POLE/POLD*, codant pour une sous-unité des polymérases répliquatives  $\epsilon$  et  $\delta$ . Ces mutations conduisent à une perte de l'activité de relecture de ces polymérases pendant la réplication (Shlien *et al.*, 2015). L'association des mutations sur les gènes *POLE/POLD* et sur les gènes MMR conduit à une absence totale de réparation des erreurs de réplication, que ce soit par l'activité de relecture des polymérases ou par la réparation des mésappariements, et provoque un nombre particulièrement important de mutations dans la tumeur. Cela contraste avec le fait que, en général, les tumeurs sporadiques de l'enfant ont un nombre de mutations beaucoup plus faible que les tumeurs équivalentes chez l'adulte.

Par ailleurs, la vitesse de progression des adénomes des patients CMMRD est plus rapide que celles des patients LS (qui est déjà plus rapide que celle des CCR MSI sporadiques) et est indubitablement la plus rapide de tous les syndromes de CCR familiaux (Durno *et al.*, 2017). Il a été rapporté notamment qu'un CCR avait été identifié chez un patient CMMRD, 6 à 11 mois après la dernière coloscopie normale (Aronson *et al.*, 2016).

### ***3. Diagnostic du CMMRD***

Poser un diagnostic définitif du CMMRD est important pour plusieurs raisons. Premièrement, les patients qui ont développé une première tumeur ont malheureusement un risque important d'en développer une ou plusieurs autre(s). Des mesures de surveillance peuvent donc être proposées aux patients atteints. Deuxièmement, la connaissance d'un défaut constitutionnel du système MMR chez le patient permettrait d'adapter le traitement puisque certaines molécules de chimiothérapie peuvent non seulement présenter une efficacité réduite et une cytotoxicité importante en contexte MMR-déficient, mais également engendrer un risque accru de développer une tumeur secondaire. Enfin, la transmission mendélienne de la maladie fait que les parents d'un patient CMMRD sont obligatoirement atteints du LS et que les frères et sœurs ont 50% de risque d'être atteints du LS et 25%

de risque d'être atteints du CMMRD. La connaissance de la ou des mutation(s) présente(s) chez le patient CMMRD permet de proposer un conseil génétique à la famille mais également de mettre en place un diagnostic prénatal ou préimplantatoire si les parents le souhaitent (Wimmer *et al.*, 2014).

Le CMMRD ayant une présentation clinique variable et un spectre clinique large, il peut être confondu avec d'autres pathologies, telles que la neurofibromatose de type 1, le syndrome de Legius, le syndrome de Li-Fraumeni ou encore la polypose adénomateuse familiale. De plus, le CMMRD est une pathologie rare, encore trop souvent méconnue des cliniciens et son incidence est certainement sous-estimée. Le diagnostic est basé sur des critères cliniques, sur la mise en évidence d'un phénotype tumoral MSI et enfin sur la détection d'une mutation germinale bi-allélique par analyse constitutionnelle des gènes MMR. Malheureusement, certaines situations restent ambiguës. Pour y remédier, l'équipe a récemment publié une méthode fonctionnelle de diagnostic basée sur des caractéristiques spécifiques des cellules MMR-déficientes (Bodo *et al.*, 2015).

*i. Critères cliniques*

Le diagnostic clinique du CMMRD est rendu compliqué par l'absence de signes cliniques pathognomoniques de la pathologie et par le large spectre tumoral du CMMRD, qui chevauche partiellement celui d'autres syndromes de prédisposition au développement de cancers.

Pour pallier à ces difficultés, le consortium Européen C4CMMRD (*Care For CMMRD*), auquel mon équipe d'accueil participe, a proposé un système de notation pour identifier les jeunes patients atteints d'une tumeur chez qui un CMMRD devrait être recherché (Wimmer *et al.*, 2014). Les tumeurs hautement spécifiques du CMMRD se voient attribuer 3 points, 2 points pour les tumeurs surreprésentées dans le CMMRD et 1 point pour les autres tumeurs. La **TABLE 3** récapitule l'ensemble des critères et le nombre de points attribués à chacun. Un CMMRD devrait être suspecté chez un individu qui obtient un score minimum de 3 et une analyse constitutionnelle des gènes MMR doit être proposée à ces patients.

Les critères cliniques préconisant la recherche d'un syndrome CMMRD chez un patient atteint de cancer	
<b>Cancer</b> : un cancer est nécessaire, si le patient a développé plusieurs cancers, additionner les points	
Cancer du spectre LS* avant 25 ans	3 points
Adénomes intestinaux multiples avant 25 ans et absence de mutation <i>APC</i> ou <i>MUTYH</i> ou un seul adénome dysplasique de haut grade avant 25 ans	3 points
Gliome de stade III ou IV avant 25 ans	2 points
Lymphome non hodgkinien à cellules T ou sPNET avant 18 ans	2 points
N'importe quel cancer avant 18 ans	1 point
<b>Autres caractéristiques</b> : optionnel, si le patient en présente plusieurs, additionner les points	
Signes cliniques de la NF1 et/ou ≥ 2 taches cutanées hyper- et/ou hypo-pigmentées de diamètre > 1 cm	2 points
Diagnostic de LS chez un apparenté au premier ou au deuxième degré	2 points
Cancer du spectre LS* avant 60 ans chez un apparenté au premier, au deuxième ou au troisième degré	1 point
Frère ou sœur avec cancer du spectre LS*, gliome de haut grade, sPNET ou lymphome non hodgkinien	2 points
Frère ou sœur avec n'importe quel cancer dans l'enfance	1 point
Multiplés pilomatricomes	2 points
Un pilomatricome	1 point
Agénésie du corps calleux ou cavernome non-induit par un traitement	1 point
Union consanguine chez les parents	1 point
Déficience / taux réduit en IgG2/4 et/ou IgA	1 point

**TABLE 3 | Critères cliniques préconisant la recherche d'un syndrome CMMRD chez un patient atteint de cancer**

Le syndrome CMMRD devrait être suspecté chez un patient présentant un score ≥ 3 points. \* cancers du spectre Lynch : cancer colorectal, endomètre, intestin grêle, voies biliaires, voies urinaires, ovaires. sPNET : tumeur primitive neuro-ectodermique supra-tentorielle. *Adapté de (Wimmer et al., 2014).*

Comme dans le cadre du LS, la recherche d'un phénotype MSI est réalisée par IHC et par biologie moléculaire. Si le principe de ces techniques reste le même entre le LS et le CMMRD, des limitations spécifiques au CMMRD nécessitent que ces méthodes soient adaptées à ce syndrome.

L'immunohistochimie des quatre protéines MMR est réalisée sur le tissu tumoral. Dans le contexte du LS, les cellules normales autour et au sein de la tumeur expriment correctement les protéines MMR et servent de contrôle positif interne à l'expérimentation. Dans le cadre du CMMRD, l'ensemble des cellules du patient (tumoraux mais également non tumoraux) sont déficientes pour l'expression d'une ou plusieurs protéines MMR. Il est donc important de réaliser en parallèle un marquage contrôle d'un autre individu sur la même lame afin de pouvoir considérer une absence totale de marquage comme évocatrice d'un CMMRD et non pas comme un défaut d'expérimentation. Comme dans le cadre du LS, le résultat de l'IHC peut orienter le test génétique vers un gène MMR ou un autre en fonction de la ou des protéine(s) dont le marquage est perdu (Durno *et al.*, 2015). Bien que cette technique ait de nombreux avantages (facilité de mise en œuvre, rapidité, moindre coût), elle présente un risque de faux négatifs dû aux mutations faux-sens qui peuvent conduire à l'expression d'une protéine non fonctionnelle. Cette technique peut être facilement réalisée chez des

individus qui ont déjà développé une tumeur et qui sont suspectés d'être atteints de CMMRD étant donné qu'on dispose alors de tissu tumoral mais également de tissu sain (muqueuse normale adjacente à la tumeur ou cellules normales intra-tumorales). Par contre, l'étude prospective sur les apparentés est plus difficile à mettre en œuvre puisqu'on ne dispose pas de prélèvement tumoral. Une biopsie cutanée est envisageable mais il s'agit d'un geste relativement invasif. Par ailleurs, cette approche n'a pas encore été validée ni standardisée.

Le phénotype MSI peut également être recherché dans la tumeur. Cependant, beaucoup de tumeurs cérébrales et d'hémopathies malignes de patients CMMRD se sont avérées être de phénotype MSS (Bakry *et al.*, 2014). En effet, comme nous l'avons vu précédemment, l'inactivation constitutionnelle du système MMR conduit à une inactivation précoce des gènes *POLE* ou *POLD*. Les mutations induites par la perte d'activité de relecture surpassent les mutations induites par la perte du système MMR et les tumeurs qui en découlent sont MSS (Shlien *et al.*, 2015).

## ii. Analyses constitutionnelles

Les patients CMMRD présentent une déficience du système MMR dans l'ensemble de leurs cellules, donc un phénotype MSI pourrait en théorie être mis en évidence dans des cellules normales, notamment dans les lymphocytes primaires. Cependant, contrairement aux cellules tumorales, ces cellules ne se répliquent pas de façon clonale : la longueur d'un microsatellite spécifique n'est pas nécessairement affectée de la même manière dans toutes les cellules et l'instabilité peut être indétectable si plusieurs génomes sont analysés à la fois, comme c'est le cas dans une PCR classique. Par conséquent, une PCR particulière (« *small-pool PCR* ») où l'échantillon d'ADN est dilué jusqu'à obtenir un maximum de 3 génomes doit être mise en œuvre (Parsons *et al.*, 1995). Ainsi, la difficulté de mise en place et d'interprétation ne font pas de cette technique un bon test diagnostique.

Récemment, une autre technique, nommée « gMSI » (MSI germinal) a été mise au point (Ingham *et al.*, 2013). Elle consiste à amplifier par PCR trois marqueurs dinucléotidiques et à analyser la taille des allèles après migration des fragments fluorescents. Le « ratio gMSI » d'un marqueur est calculé en divisant la taille d'un potentiel allèle supplémentaire par la taille de l'allèle majoritaire et il est considéré comme positif s'il dépasse un certain seuil. Il faut au moins 2 marqueurs positifs sur trois pour que le profil de l'individu soit considéré comme MSI. Cette technique est reproductible et montre une grande spécificité et sensibilité pour les cas CMMRD liés à des mutations *MLH1*, *MSH2* ou *PMS2*. Par contre, les cas de CMMRD dus à des mutations dans le gène *MSH6* ne sont pas détectés par cette technique. Contrairement à la « *small-pool PCR* », elle ne nécessite pas de quantification exacte (nécessaire à la réalisation des dilutions précises). Enfin, elle permet d'analyser un certain nombre d'échantillons à la fois et est donc bien adaptée à un usage en diagnostic de routine. Cependant, la

méthode n'est pas contributive dans le cas des marqueurs hétérozygotes, ce qui arrive dans un nombre non négligeable de situations (environ 22% des patients).

Enfin, comme pour le LS, le diagnostic final du CMMRD est moléculaire et basé sur l'identification d'une mutation bi-allélique délétère sur un gène MMR. L'analyse constitutionnelle des gènes MMR peut être orientée par les résultats d'IHC. Le gène *PMS2* étant le plus difficile à analyser à cause de nombreux pseudo-gènes, son étude n'est entreprise qu'en cas de perte isolée de la protéine *PMS2* en IHC. Enfin, détecter deux mutations et confirmer qu'elles sont présentes chacune sur un allèle peut s'avérer compliqué (Durno *et al.*, 2015).

iii. *Caractérisation fonctionnelle des Variants de Signification Incertaine dans le cadre du CMMRD*

L'analyse constitutionnelle des gènes MMR est non contributive dans environ 30% des cas. En effet, un variant de signification incertaine (voir Chapitre 4) peut être détecté sur l'un des deux allèles ou les deux. On peut également détecter une seule mutation ou pas de mutation du tout. Cela peut s'expliquer en partie par le fait que la détection des altérations sur le gène *PMS2* est particulièrement compliquée due à la présence de nombreux pseudo-gènes et manque de sensibilité. Par conséquent, un test permettant de confirmer ou infirmer le diagnostic du CMMRD sans équivoque serait appréciable.

Notre équipe a développé un test de diagnostic du CMMRD basé sur les deux caractéristiques fonctionnelles spécifiques des cellules tumorales MMR-déficientes : leur phénotype MSI dû à l'incapacité de réparer les mésappariements et leur phénotype de tolérance à la méthylation dû à l'incapacité à signaler certains dommages à l'ADN (Bodo *et al.*, 2015). Ces deux caractéristiques sont étudiées sur des lignées lymphoblastoïdes (LLB), issues de cellules lymphoïdes de patients immortalisées grâce à une infection par le virus Epstein-Barr. Dans un premier temps, une étude cas-contrôle a été réalisée avec des LLB de patients CMMRD avérés et des patients dont les cellules sont MMR-compétentes. Cette dernière catégorie inclut des patients LS et des patients n'ayant aucune mutation MMR constitutionnelle, notamment des patients atteints de neurofibromatose de type 1 ou de Polypose Adénomateuse Familiale, qui sont des diagnostics différentiels du CMMRD. Ce test permet de différencier les patients CMMRD de patients non CMMRD avec 100% de sensibilité et 100% de spécificité.

Dans une deuxième étape, les LLB de 23 patients suspectés d'être atteints du CMMRD ont été étudiés. Ces individus avaient tous une présentation clinique évocatrice du CMMRD mais le diagnostic n'avait pas pu être confirmé par l'analyse mutationnelle des gènes MMR. Cela incluait des individus présentant des altérations MMR bi-alléliques comprenant 1 ou 2 VSI ; des individus chez qui une seule

altération a été détectée et des patients chez qui aucune mutation n'a été mise en évidence. L'avantage de cette méthode est d'être basée sur les cellules issues du patient, l'altération suspectée de mener au phénotype observé n'a donc pas besoin d'être connue. Dans cette série de 23 individus, 5 LLB présentaient deux tests anormaux et les patients sont donc très probablement atteints du CMMRD ; 15 LLB ont montré une réponse normale pour les deux tests et un CMMRD est très improbable. Enfin, deux cas ne présentaient qu'une caractéristique sur les deux et sont donc restés ambigus. Ce test est maintenant disponible en diagnostic de routine à l'hôpital Saint-Antoine et a permis de confirmer le diagnostic de CMMRD chez 9 individus sur 46 patients testés (20%). Le test de caractérisation fonctionnelle des variants de signification que j'ai développé dans mon travail de thèse dans le contexte du LS est dans la continuité directe de cette étude.



## Chapitre 4 : La problématique des Variants de Signification Incertaine

### A. Définition générale

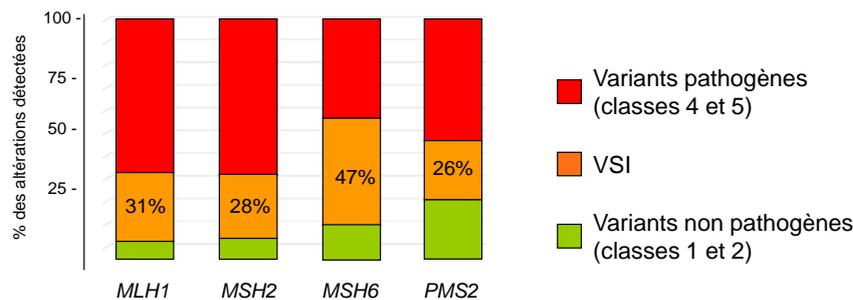
Dans ce paragraphe, nous allons nous intéresser uniquement aux syndromes autosomiques dominants de susceptibilité au développement de cancers, et plus particulièrement au syndrome de Lynch (LS). Dans ce contexte, le test génétique sert à proposer des mesures préventives et thérapeutiques à l'individu affecté mais il permet aussi le conseil génétique afin de déterminer si certains de ses apparentés sont également atteints.

Le diagnostic génétique, par séquençage et MLPA, permet en théorie de poser le diagnostic grâce à la détection d'une altération germinale. Dans la suite de ce chapitre, nous parlerons d'« altération » ou de « variant » plutôt que de « mutation ». En effet, une mutation est définie comme un changement permanent dans la séquence nucléotidique et ce terme ne porte pas d'appréciation sur l'aspect délétère ou non de cette modification de séquence. Un polymorphisme est caractérisé quant à lui comme une variation de séquence ayant une fréquence supérieure à 1% dans la population générale et, de même, cette définition ne préjuge pas de son caractère pathogène ou non. Une confusion classique est de penser qu'une mutation est forcément pathogène et qu'un polymorphisme est forcément non pathogène (Richards *et al.*, 2015). Pour plus de clarté, nous utiliserons donc le terme « variant » ou « altération », éventuellement suivi de « pathogène » ou « neutre » si sa conséquence fonctionnelle est connue ou « de signification incertaine » si elle est inconnue.

Certains types d'altérations conduisent à un épissage aberrant ou provoquent l'apparition d'un codon stop prématuré et engendrent l'expression d'une protéine tronquée non fonctionnelle : ces variants sont dits « pathogènes » et sont à l'origine du syndrome chez l'individu. À l'inverse, des altérations qui n'ont pas d'effet ni sur l'épissage ni sur la protéine peuvent être mises en évidence, telles que certaines variations introniques ou substitutions synonymes : ces variants sont considérés « neutres » et ne sont pas responsables du syndrome chez l'individu. Enfin, dans une proportion non négligeable de cas, un Variant de Signification Incertaine (VSI) est mis en évidence. Il s'agit d'une altération de la séquence génétique pour laquelle il est impossible d'établir formellement le statut délétère ou bénin. Par conséquent, le lien de causalité entre la présence du variant et la pathologie observée chez l'individu ne peut être établi avec certitude.

Les VSI sont principalement des variants non synonymes et non tronquants dans la région codante : substitutions faux-sens, petites insertions/délétions en phase ou encore altérations susceptibles de modifier l'épissage (Peltomäki et Vasen 2004 ; Thompson *et al.*, 2014). En ce qui concerne le LS, les VSI représentaient en 2016 31%, 28%, 47% et 26% des altérations détectées dans

les gènes *MLH1*, *MSH2*, *MSH6* et *PMS2*, respectivement (Peltomäki, 2016) (**FIGURE 9**). Les VSI ne sont pas une problématique restreinte au LS, mais concernent théoriquement tous les gènes impliqués dans des syndromes génétiques. Par exemple, 30 à 50% des variants identifiés dans les gènes *BRCA1* et *BRCA2*, impliqués dans le syndrome sein-ovaire, sont considérés comme des VSI (Lindor *et al.*, 2012). Le manque d'informations ou la présence de données contradictoires dans la littérature scientifique et médicale sur ces variants empêchent d'affirmer avec certitude si ce sont des variations non pathogènes (polymorphisme rare) ou s'il s'agit effectivement de variations qui affectent la fonction du gène, qui sont à l'origine du syndrome génétique et qui par conséquent augmentent le risque de développer un cancer (variant pathogène) (Chao *et al.*, 2008). Par exemple, une substitution d'acide aminé peut être pathogène si elle a lieu dans un domaine de la protéine dont la fonction est importante pour la suppression tumorale ou si elle altère la structure globale de la protéine (Li *et al.*, 2009) ; comme elle peut ne pas être pathogène si elle se situe dans une région qui n'a pas de rôle particulier dans le fonctionnement de la protéine.



**FIGURE 9 | Distribution des variants des gènes MMR associés au syndrome de Lynch au sein des différentes classes de pathogénicité**

L'analyse inclut 932 variants *MLH1*, 842 *MSH2*, 449 *MSH6* et 137 *PMS2*. Adapté de (Peltomäki, 2016).

L'annonce de l'identification d'un VSI peut avoir des conséquences différentes selon les individus. Pour certains, cela peut constituer une source d'anxiété et ces individus se considèrent comme atteints du LS, ou d'une « forme différente » de LS : ils peuvent se voir comme des « bombes à retardement » et penser que le développement d'un cancer fait partie de leur « destinée ». D'autres considèrent ce résultat comme un résultat « intermédiaire », et même s'ils ont conscience que ce variant peut avoir des conséquences sur leur santé, ils considèrent que ce risque est moindre par rapport à celui induit par une altération pathogène.

Les individus porteurs de VSI d'un gène MMR se voient proposer une surveillance comparable à celle des individus effectivement atteints du syndrome de Lynch. Cela constitue évidemment un fardeau physique et émotionnel pour l'individu en question mais également une lourde charge pour le système de santé. La recherche du VSI par analyse génétique et le programme de surveillance ne

sont pas proposés aux apparentés. Ainsi, la détection d'un VSI a des conséquences sur l'individu chez qui il est mis en évidence, mais pas sur les membres de sa famille, qu'ils aient ou non développé un cancer faisant penser qu'ils pourraient potentiellement être atteints du LS. Ainsi, la plupart de ces patients sont dans l'espoir que des informations complémentaires puissent aider à l'interprétation de leur VSI, afin de leur apporter plus de certitude sur leur condition et de servir à la prise en charge de leur famille (Solomon *et al.*, 2017).

L'analyse constitutionnelle des gènes MMR était initialement réalisée par séquençage de type Sanger. Actuellement, les techniques de séquençage à haut débit comprennent le séquençage de gènes uniques, d'un panel de gènes, d'exome, de génome et de transcriptome ou encore des tests épigénétiques (Richards *et al.*, 2015). L'introduction récente de ces technologies permettant d'obtenir des résultats beaucoup plus rapidement et à moindre coût (Metzker, 2010) a fait augmenter le nombre de VSI répertoriés, phénomène qui ne pourra que s'amplifier avec les années (Hinrichsen *et al.*, 2013). D'autre part, le diagnostic génétique est une pratique de plus en plus acceptée par la société, ce qui provoque une augmentation du dépistage systématique (Couch *et al.*, 2008). Actuellement, notre capacité à séquencer le génome humain a dépassé notre capacité à interpréter les variations génétiques. Par conséquent, le goulot d'étranglement dans le diagnostic des syndromes génétiques ne se situe plus sur la détection des altérations génétiques mais bien dans l'interprétation des altérations détectées.

A l'heure actuelle, il n'existe pas de méthode simple qui permettrait de classer avec certitude un VSI présent sur un gène MMR, et toutes les données (cliniques, familiales, épidémiologiques, *in silico* et fonctionnelles) sont collectées pour tenter d'apporter un faisceau d'arguments permettant de caractériser les VSI (Sijmons *et al.*, 2013). Nous allons discuter de ces différents arguments dans les paragraphes suivants.

## B. Arguments pour l'interprétation des variants des gènes MMR

### 1. *Données cliniques et études de co-ségrégation*

Afin d'aider à la classification d'un VSI détecté chez un individu, une étude de son histoire clinique personnelle et familiale est entreprise. La présence d'une instabilité microsatellitaire dans la tumeur, détectée par analyse biomoléculaire et/ou par IHC, est déjà un argument en soi. On cherche ensuite une co-ségrégation entre la présence du variant et le phénotype clinique du syndrome au sein de la famille de l'individu, c'est-à-dire que plus le variant est retrouvé chez les individus qui présentent une histoire personnelle typique du LS, plus il y a de probabilité que ce variant soit pathogène. Pour

chaque individu affecté de la famille, on prend note de l'histoire personnelle de cancer, de l'âge au diagnostic, du phénotype MSI.

Cependant, cette étude de co-ségrégation peut s'avérer compliquée. En effet, comme nous l'avons vu précédemment, dans le cas de la détection d'un VSI, les apparentés du cas index ne profitent pas du conseil génétique et sont donc rarement testés pour la présence du variant. Ceci rend difficile la mise en évidence d'une corrélation entre la présence d'un phénotype évocateur du LS et la présence du variant. Par ailleurs, une étude de co-ségrégation est d'autant plus informative que la famille est grande (la robustesse de ces tests augmente avec le nombre de méioses) ; or, dans nos sociétés et de nos jours, beaucoup de familles sont de taille modeste. Ensuite, on peut se heurter au manque d'informations cliniques ou au manque de matériel biologique (dans le cas d'un individu décédé par exemple). De plus, la plupart des VSI identifiés sont uniques à un individu ou une famille (Hinrichsen *et al.*, 2013).

L'existence de phénocopies et la faible pénétrance relative du syndrome de Lynch complexifient encore la situation même si on se place dans le cas idéal d'une famille où on connaîtrait le phénotype et le génotype de tous les individus. En effet, au sein d'une potentielle famille LS, un CCR peut être développé de façon sporadique par un individu non porteur du variant (non porteur symptomatique), ce qui provoque des doutes sur l'association entre le variant et le phénotype clinique. Par ailleurs, un individu peut être porteur du variant et ne pas avoir développé de tumeurs, ce qui laisse penser que l'individu est porteur sain (porteur asymptomatique), surtout si le variant est présent sur les gènes *MSH6* ou *PMS2*, connus pour prédisposer au développement de cancers à des âges plus tardifs. De plus, il a été suggéré que les altérations non tronquantes pourraient être associées à un phénotype clinique atypique avec un âge au diagnostic de la première tumeur plus tardif et peu ou pas de phénotype MSI dans les tumeurs (Kantelinen *et al.*, 2012).

Enfin, la ségrégation du variant avec le syndrome peut simplement refléter le fait que le variant est en déséquilibre de liaison avec une autre altération pathogène non encore détectée ; c'est-à-dire que le variant identifié co-ségrège dans la famille avec un autre variant, non identifié mais qui est à l'origine du syndrome. En effet, il n'est pas rare qu'un même individu soit porteur de plusieurs variants sur les deux copies du même gène MMR ou sur deux gènes MMR différents. Dans ce cas, il est encore plus compliqué d'estimer à quel variant est dû le phénotype observé ou s'il s'agit d'un effet additif de plusieurs variants. Par ailleurs, le phénotype MSI détecté en biologie moléculaire ou la perte d'expression de la protéine MMR dans le tissu tumoral du porteur du VSI n'est pas forcément imputable au VSI mais peut être dû à une autre mutation non encore identifiée dans le gène MMR ou ses séquences de régulation, ou à une inactivation somatique du gène (Christensen *et al.*, 2009).

## 2. *Données épidémiologiques*

Un autre argument pour aider à la classification des VSI est un argument de fréquence. La fréquence du VSI est recherchée chez les individus atteints ou suspectés d'être atteints du LS (bases de données de patients) mais également dans la population générale (bases de données de populations contrôles). En effet, plus un VSI est fréquent chez des individus/familles suspecté(e)s d'être atteint(e)s du syndrome de Lynch, plus il est probable qu'il soit pathogène (Sijmons *et al.*, 2013). Au contraire, plus un VSI est fréquent dans la population générale, plus il est probable qu'il soit neutre. De manière générale, il est admis qu'un variant dans un gène responsable de syndrome de transmission autosomique dominante ne peut pas avoir une fréquence allélique supérieure à la prévalence du syndrome. Dans tous les cas, il est important de prendre en considération l'origine ethnique ou géographique de l'individu porteur du VSI car les variants sont souvent spécifiques d'une population donnée (Goldgar *et al.*, 2008).

### i. *Les bases de données de patients*

ClinVar (Landrum *et al.*, 2016) est une base de données recensant les variants découverts chez des individus, en relation avec leur phénotype (et n'est donc pas restreinte aux variants des gènes MMR). Chaque soumissionnaire propose une classification du VSI basée sur les arguments à sa disposition. Cette base de données accepte des soumissions à différents niveaux de complexité (allant de la seule fréquence allélique à la liste de tests fonctionnels réalisés en laboratoire, en passant par des données cliniques) afin de permettre la constante évolution de la classification et le développement de connaissances pour tous les VSI répertoriés. Quand de nouvelles informations sont disponibles, le rapport précédent est conservé, afin de pouvoir suivre l'évolution de la classification d'un VSI. De même, les potentielles données conflictuelles sont conservées, afin que chacun puisse se faire sa propre idée. A chaque VSI est attribué un certain nombre d'étoiles, allant de 0 à 4, représentatif du degré de confiance accordé à l'analyse. En effet, la base de données n'est pas curée (toute soumission est mise en ligne sans vérification bibliographique préalable). Ainsi, 4 étoiles signifie que la classification du VSI en question a été proposée par un groupe ayant le statut « *practice guideline* » ; le statut 3 étoiles indique une soumission réalisée par un groupe ayant le statut « panel d'expert » ; le statut 2 étoiles correspond à un VSI ayant des soumissions de plusieurs groupes qui présentent des arguments non contradictoires ; le statut 1 étoile correspond soit à deux soumissionnaires indiquant des données conflictuelles soit à un seul soumissionnaire. Il est important de garder à l'esprit que certains variants sont rapportées avec des informations fausses ou incomplètes, l'utilisateur doit donc vérifier la source des données disponibles.

InSiGHT (*International Society for Gastrointestinal Hereditary Tumor*) est une organisation scientifique multidisciplinaire internationale qui tient une base de données de variants de gènes impliqués dans les cancers gastro-intestinaux (Fokkema *et al.*, 2011). On y retrouve en autres les variants des gènes MMR, impliqués dans le LS, mais également les gènes causant la polypose recto-colique familiale, le syndrome de Peutz-Jeghers, le syndrome de polypose juvénile ou le syndrome des hamartomes dû à une mutation de *PTEN*. Il s'agit d'une base de données « curée » par un comité d'interprétation des variants qui comprend 40 experts multidisciplinaires de 5 continents (Thompson *et al.*, 2014). Cette base de données comprend à l'heure actuelle 1413 variants uniques dans le gène *MLH1*, dont 281 (19,9%) VSI et 1386 variants uniques dans le gène *MSH2*, dont 237 VSI (17,1%).

La base de données française UMD (*Universal Mutation Database*) (Bérout *et al.*, 2000) est disponible pour 25 gènes impliqués dans des syndromes de prédisposition aux cancers (syndrome de Lynch, syndrome sein-ovaire lié à *BRCA1/2*...) ou à d'autres maladies génétiques (syndrome de Marfan, mucoviscidose...). Les bases UMD-*MLH1*, UMD-*MSH2* et UMD-*MSH6* regroupe les données de 17 laboratoires de diagnostic français. Les données sont envoyées au centre de curation à Marseille pour être analysées et incluses dans la base, qui comprend à l'heure actuelle 568 variants uniques dans le gène *MLH1*, dont 257 (45%) VSI et 638 variants uniques dans le gène *MSH2*, dont 288 (45%) VSI.

Un même individu ou une même famille peut être répertorié(e) sur plusieurs bases de données de patients différentes. Ceci peut notamment être le cas en cas de collaborations internationales ou si des individus de la même famille sont suivis par des institutions différentes. Ainsi, la collecte d'informations cliniques sur différentes bases de données nécessite de considérer ce biais afin d'éviter les doublons et de ne pas augmenter artificiellement la fréquence du variant (Richards *et al.*, 2015).

## ii. Les bases de données de populations

La base de données 1000 Genomes contient des données issues de 3511 génomes de sujets non apparentés (1000 Genomes Project Consortium *et al.*, 2015). Cette base regroupe 26 populations différentes et est principalement adaptée à l'interprétation de variants non exoniques. Pour sa classification des VSI, la base de données InSiGHT prend en compte les fréquences rapportées sur la base de données du projet 1000 Génomes ainsi que d'éventuelles données publiées par des centres de recherche sur des populations locales (Tricarico *et al.*, 2017).

La base ESP (*NHLBI Exome Sequencing Project*) rassemble les données issues de 6503 exomes de sujets non apparentés. Les individus séquencés sont stratifiés en deux populations : EA (European-Americans) et AA (African-Americans).

La base ExAC (*Exome Aggregation Consortium*) contient les données issues de 60706 exomes de sujets non apparentés, et incluent les données de 1000 génomes et partiellement d'ESP (Lek *et al.*,

2016). La population générale est génétiquement stratifiée en 6 populations (AFR : population africaine et afro-américaine, AMR : population d'origine mixte américaine hispanique, EAS : population d'Asie de l'est, FIN : population finlandaise, NFE : population européenne non finlandaise, SAS : population d'Asie du sud, OTH : autres). Cette base est principalement adaptée à l'interprétation de variants exoniques. Elle contient les informations de patients atteints de formes héréditaires de cancer. Des données du TCGA (*The Cancer Genome Atlas*) sont également incluses.

Enfin, la base GnomAD (*Genome Aggregation Database*) rassemble les données d'ExAC, de 1000 génomes et partiellement d'ESP, soit un total de 123 136 exomes de sujets non apparentés et de 15 496 génomes. Il s'agit actuellement de la base de données la plus puissante en termes de nombre d'individus séquencés. En particulier, la présence de 15 496 génomes entiers permet d'appréhender la fréquence des variants non-exoniques de manière plus précise que 1000 Genomes. Elle intègre des populations qui jusque-là n'étaient pas représentées dans les bases de données comme la population juive ashkenaze.

Il faut garder à l'esprit que les bases de données de population ne contiennent pas forcément que des individus sains et sont donc susceptibles de contenir des variants pathogènes (associés au LS ou à d'autres maladies génétiques) (Richards *et al.*, 2015).

### **3. Les algorithmes de prédiction *in silico***

Une grande variété d'outils *in silico* peuvent aider à l'interprétation des VSI. L'impact d'une altération faux-sens dépend de différents critères, tels que la conservation du nucléotide ou de l'acide aminé au cours de l'évolution, sa localisation et son contexte au sein des séquences nucléique et protéique ou encore les conséquences biochimiques de la substitution d'acide aminé. La substitution d'une base peut avoir un impact sur la régulation de l'expression du gène ou sur l'épissage ; alors que la modification d'acide aminé qui en résulte peut avoir des répercussions sur la fonction de la protéine.

Bien que les programmes informatiques utilisent des algorithmes de prédiction différents, leur fonctionnement de base reste similaire. Ainsi, la prédiction combinée de différents algorithmes est considérée comme un seul argument plutôt que comme des arguments indépendants. Cependant, l'utilisation de différents programmes est quand même recommandée car chacun présente des forces et des faiblesses, notamment en fonction du gène et de la séquence protéique considérés (Richards *et al.*, 2015). Les résultats des tests *in silico* ne sont que des prédictions et ne constituent qu'un argument parmi d'autres pour aider à la classification d'un VSI, aucune conclusion définitive ne peut être tirée de cet outil seulement (Sijmons *et al.*, 2013).

i. Logiciels de prédiction de l'effet du variant sur l'épissage

Les principaux logiciels de prédiction des aberrations d'épissage sont MaxEntScan, NNSplice, SpliceSiteFinder, GeneSplicer et Human Splicing Finder. Il existe deux grands types de logiciels de prédiction : (i) ceux qui analysent les séquences consensus des sites donneurs et accepteurs d'épissage et les effets de l'altération sur l'affaiblissement d'un site canonique ou sur la création d'un site cryptique d'épissage et (ii) ceux qui étudient les sites exhausteurs ou inhibiteurs d'épissage qui peuvent être reconnus par des facteurs d'épissage (sans pour autant pouvoir préciser quels sont les sites effectivement utilisés dans les conditions physiologiques et dont la modification risque d'avoir une conséquence fonctionnelle). La plupart des programmes ont initialement été développés pour la recherche de gènes, basée sur la détection des séquences consensus d'épissage. Ainsi, ces programmes n'ont pas été conçus pour prédire les conséquences d'une altération de séquence sur l'épissage. De plus, l'analyse est réalisée en examinant la séquence sauvage et la séquence variante indépendamment l'une de l'autre et il n'y a pas d'outil d'interprétation direct (Spurdle *et al.*, 2008). Les outils de prédiction de sites d'épissage ont une sensibilité (90-100%) supérieure à la spécificité (60-80%) (Houdayer *et al.*, 2012), c'est-à-dire qu'ils détectent beaucoup de potentielles aberrations mais la plupart n'ont pas d'effets fonctionnels.

ii. Logiciels de prédiction de l'effet du variant sur la protéine

Les logiciels de prédiction de l'effet des substitutions non synonymes sur la protéine se basent principalement sur l'homologie entre la séquence protéique humaine de référence et les séquences de protéines issues de gènes paralogues et orthologues. L'hypothèse est que les acides aminés importants pour la fonction de la protéine sont conservés au sein des séquences d'une même famille de protéines. Ainsi, plus un acide aminé est conservé à une position donnée, plus la probabilité est grande que la substitution soit délétère. Cependant, ce n'est pas obligatoirement l'acide aminé qui est important en lui-même, mais sa propriété : les logiciels analysent également quel est le type de propriété dominant à une position donnée. Ainsi, si on retrouve à une même position de différentes séquences une isoleucine, une leucine ou une valine, le logiciel est capable d'interpréter que ce n'est pas l'acide aminé qui est nécessaire, mais son caractère hydrophobe. Ainsi, une substitution vers un autre acide aminé hydrophobe sera généralement prédite comme tolérée alors qu'une substitution vers un autre type de résidu sera probablement prédite comme affectant la fonction (Ng et Henikoff, 2003). La diversité d'acides aminés à la position d'intérêt montre la gamme de propriétés physicochimiques qui sont tolérées (indice d'hydropathie, charge, volume de la chaîne latérale, énergie libre en hélice  $\alpha$  ou feuillet  $\beta$ ).

Le logiciel SIFT (*Sorting Intolerant From Tolerant*) (Ng et Henikoff, 2002), qui permet de prédire les effets des variants de n'importe quel gène, ainsi que le logiciel MAPP-MMR (*Multivariate Analysis of Protein Polymorphism*) (Chao *et al.*, 2008 ; Stone et Sidow, 2005) spécifique des protéines MLH1 et MSH2, se basent uniquement sur ces données de conservation. Le programme PolyPhen-2 (Adzhubei *et al.*, 2010) prend en compte, en plus de l'alignement de séquences multiples, les éléments de structure de la protéine de référence et de la protéine variante.

Les logiciels de prédiction de l'effet du variant sur la protéine prennent en considération la conservation de la séquence protéique au cours de l'évolution et les propriétés des acides aminés. La plupart ne considèrent pas d'autres caractéristiques biologiques importantes telles que les sites d'épissage ou de modifications post-traductionnelles ou le biais d'usage du code génétique (un même acide aminé codé à partir de deux codons différents ne sera pas forcément traduit à la même efficacité). Par ailleurs, la dégénérescence du code génétique fait que plusieurs codons peuvent coder pour le même acide aminé. Cette dégénérescence est particulièrement importante sur la troisième base du codon, ce qui induit une conservation entre les espèces souvent bien faible à ces positions. Ainsi, l'observation d'une conservation à ces positions pourrait dans certains cas être plus significative, suggérant l'existence d'une éventuelle séquence régulatrice à ce niveau. Dans ce cas-là, c'est la base, et non pas l'acide aminé, qui est important, et même une substitution synonyme est susceptible d'être délétère. Or, les logiciels ne prédisent que les effets des altérations non synonymes (Bali et Bebok, 2015).

La plupart des algorithmes ont une précision de 65 à 80% quand on analyse des variants contrôles connus pour être pathogènes ou neutres (Thusberg *et al.*, 2011), la précision étant dans ce cas la proportion de résultats corrects (vrais positifs et vrais négatifs) parmi tous les cas examinés. Ils ont une faible spécificité, conduisant à une surreprésentation des variants faux-sens considérés comme pathogènes. De plus, ils ne sont pas très fiables pour prédire les variants faux-sens ayant un effet modéré (Choi *et al.*, 2012). Dans tous les cas, le nombre de variants contrôles (connus pathogènes et connus neutres) dans la cohorte d'entraînement du logiciel doit être important, ce qui peut être problématique, surtout pour les gènes dont peu de variants sont connus pour être clairement neutres ou pathogènes et pour lesquels on a moins d'informations sur le spectre clinique, tels que *PMS2* et *MSH6* (Arora *et al.*, 2017).

#### **4. Tests fonctionnels**

Des tests biologiques réalisés en laboratoire cherchent à évaluer les conséquences du variant, que ce soit au niveau de l'ARN (épissage) ou de la protéine (capacité à remplir ses fonctions), afin d'estimer son caractère pathogène ou non. Le principe de l'utilisation de ces tests fonctionnels est que

la prédisposition au cancer est due à une perte de fonction de gènes suppresseurs de tumeurs, en l'occurrence les gènes MMR. Ainsi, la diminution de l'activité de la protéine associée induit une prédisposition au cancer augmentée (Couch *et al.*, 2008). En ce qui concerne les gènes MMR, le développement de ces tests a débuté à la toute fin des années 1990. A l'heure actuelle, une grande variété de tests ont été mis au point, *in vivo* ou *in vitro* ; sur des modèles de levures, de cellules humaines ou de cellules murines ; afin d'étudier un aspect restrictif de la protéine ou au contraire une fonction de la protéine globale. Un catalogue de ces tests associé à l'évaluation de leurs avantages et limites respectifs sera présenté dans cette partie.

i. Etude de l'épissage

Les substitutions faux-sens synonymes ou non synonymes peuvent avoir des conséquences fonctionnelles, notamment si elles se situent au niveau ou à proximité d'un site d'épissage, ou sur une séquence reconnue par des facteurs favorisant ou inhibant la transcription. Des tests sont donc mis en place pour vérifier que le variant permet l'expression d'un ARN correctement épissé, produit en quantité normale.

Des tests peuvent être réalisés à partir de matériel biologique du patient ou *in vitro* pour analyser le caractère délétère ou non du variant sur l'épissage. Les tests réalisés sur l'ARN issu de cellules sanguines de patients doivent être réalisés sur des échantillons particulièrement frais ou collectés dans des tubes appropriés pour limiter la dégradation et la production de produits de transcription artéfactuels due au stress environnemental (Speevak *et al.*, 2003). L'alternative est d'établir une lignée lymphoblastoïde à partir des cellules sanguines du patient, mais cela est coûteux en termes de temps et de ressources (Spurdle *et al.*, 2008). Une autre possibilité est de recourir à la technique du mini-gène, qui ne nécessite pas de matériel biologique du patient. Un fragment génomique contenant l'exon altéré ainsi que les séquences environnantes de part et d'autre est cloné dans un vecteur d'expression standard puis introduit dans des cellules en culture. Les cellules sont cultivées en parallèle en présence et en absence de puromycine, qui inhibe le système NMD (*Nonsense-Mediated mRAN Decay*) susceptible de conduire à la dégradation des ARN qui portent un codon stop prématuré. Après transcription, les effets sur l'épissage sont étudiés par RT-PCR suivi d'un séquençage (Auclair *et al.*, 2006 ; Tournier *et al.*, 2008). Plusieurs variants nucléotidiques, détectés chez le même individu, peuvent être analysés séparément ou ensemble. Cette technique *in vitro* présente de nombreuses limitations : la séquence du gène n'est pas incluse en globalité et le type cellulaire peut influencer la production des différents produits d'épissage potentiels. En effet, plusieurs produits d'épissage sont naturellement synthétisés à partir de la plupart des gènes, et le type et la quantité de ces produits peut dépendre du type cellulaire. Ainsi, les tests *in vitro* ne reproduisent pas

les subtiles variations d'expression des produits d'épissage dues à l'expression tissu-spécifique des facteurs d'épissage et peuvent donc avoir des difficultés à mimer la production endogène du produit d'épissage variant, et à interpréter ses conséquences.

Quelle que soit la méthode mise en œuvre, l'interprétation de la signification biologique des résultats d'épissage n'est pas toujours évidente, notamment si le variant conduit à la délétion en phase d'un exon sans domaine fonctionnel particulier ou qui provoque de petites insertions/délétions en phase (Spurdle *et al.*, 2008). De plus, un variant d'épissage peut conduire à l'expression totale ou partielle d'un ARN alternatif à partir de l'allèle variant, au détriment de l'ARN sauvage : déterminer la quantité de cet ARN variant, notamment par rapport à celle de l'ARN sauvage, est important pour établir la signification clinique ou biologique (Claes *et al.*, 2002). Par ailleurs, le variant peut aussi conduire à l'expression d'un ARN correctement épissé, mais produit en plus faible quantité que le produit sauvage. Cela peut être vérifié par l'étude de l'expression des ARNm d'intérêt par RT-PCR quantitative (Arora *et al.*, 2017 ; Brieger *et al.*, 2002 ; Hinrichsen *et al.*, 2013). Enfin, le variant peut également avoir des effets sur la traduction ou sur les modifications post-traductionnelles.

## ii. Etude d'un aspect restrictif de la protéine

Les substitutions d'acides aminés peuvent conduire à une protéine instable qui ne peut donc pas remplir sa fonction (qu'elle soit théoriquement active ou non), à une protéine stable mais partiellement ou totalement inactive, ou encore à un effet dominant négatif (la protéine codée par le gène variant, non seulement perd sa fonction, mais interfère aussi avec la fonction du produit de l'allèle normal chez les hétérozygotes) (Couch *et al.*, 2008 ; Lützen *et al.*, 2008). Pour remplir leur rôle dans la réparation des mésappariements et la signalisation des dommages à l'ADN, les protéines MMR doivent être exprimées, capables d'interagir avec leurs protéines partenaires ou encore adressées au noyau par exemple. Une altération de séquence est capable d'abolir n'importe laquelle de ces étapes, ce qui rend la protéine incapable de remplir sa fonction globale. Etudier les différents aspects permet d'avoir une idée du mécanisme d'inactivation de la fonction MMR induite par les VSI.

### a. *Expression et stabilité de la protéine*

Pour pouvoir assurer sa fonction, une protéine doit avant tout être exprimée, et ce de façon stable. Pour étudier l'expression des protéines variantes, une lignée cellulaire humaine déficiente pour le gène MMR d'intérêt est transfectée par un plasmide porteur du VSI à tester. Une expression stable peut également être obtenue grâce à un système d'expression lentivirale (Cyr *et al.*, 2012). Des extractions protéiques sont ensuite réalisées afin d'évaluer par Western Blot l'expression de la protéine d'intérêt et éventuellement de sa protéine partenaire (PMS2 ou MSH6/MSH3). Dans l'étude

de Brieger, les extractions sont réalisées à différents temps post-transfection afin de suivre la cinétique d'expression : la protéine MLH1 (exprimée à partir d'un plasmide) apparaît à 24h et son expression atteint son maximum à 36h, alors que l'expression de son partenaire PMS2 (exprimé à partir du gène endogène) présente un décalage : elle apparaît au bout de 48h et est optimale à 72h. L'expression de la protéine MSH2 (à partir d'un plasmide) est également optimale à 36h ; par contre, sa protéine partenaire MSH6 (exprimée à partir du gène endogène) ne présente pas le décalage observé pour PMS2 (Brieger *et al.*, 2002).

En plus de renseigner sur le taux d'expression de la protéine variante d'intérêt, cette simple expérience de Western Blot permet également de documenter l'interaction de MLH1 avec PMS2 et de MSH2 avec MSH6/MSH3. En effet, PMS2 et MSH6/MSH3 sont instables en absence de MLH1 et MSH2, respectivement. Ainsi, si ces dernières sont exprimées à des taux satisfaisants, une absence d'expression de PMS2 ou MSH6 par Western Blot peut révéler un défaut dans le domaine d'interaction des protéines variantes MLH1 ou MSH2, respectivement, avec leur partenaire. Pour pallier au décalage entre l'expression de la protéine d'intérêt et la stabilisation de sa protéine partenaire, certaines équipes font le choix de co-transfecter un plasmide porteur du VSI d'intérêt et un plasmide porteur du gène MMR de la protéine partenaire (Belvederesi *et al.*, 2006 ; Borràs *et al.*, 2012 ; Raevaara *et al.*, 2005).

Une diminution de l'expression d'une protéine, révélée par Western Blot, peut être due à une diminution de la transcription et/ou de la traduction ou encore une diminution de la stabilité de l'ARN ou de la protéine. Pour étudier ces aspects, l'équipe de Perera a réalisé des expériences de « *pulse-chase* » ou de traitement avec un inhibiteur de la synthèse des protéines qui permettent de déterminer le temps de demi-vie des protéines sauvages ou variantes (Perera et Bapat, 2008).

Par ailleurs, une étude récente indique que certaines protéines MSH2 variantes ne montrent pas de perte de fonction inhérente, mais qu'elles s'avèrent pathogènes car elles présentent des défauts de repliement, sont dégradées par le protéasome et ne peuvent donc pas remplir leur fonction (Nielsen *et al.*, 2017). En effet, certaines protéines issues de variants faux-sens sont plus rapidement dégradées que leurs équivalents sauvages (Capecchi *et al.*, 1974) et MSH2 a été montré comme une cible de ce système chez la levure (Arlow *et al.*, 2013). De manière générale, les protéines issues d'altérations faux-sens sont moins stables structurellement que les protéines sauvages (Tokuriki *et al.*, 2007). Ainsi, le système de contrôle des protéines pourrait être trop zélé et dégrader des protéines structurellement stables qui pourraient avoir conservé leur fonction, comme cela a pu déjà être prouvé dans la mucoviscidose (Vembar et Brodsky, 2008).

La corrélation est loin d'être parfaite entre l'expression d'une protéine variante étudiée par Western Blot et la fonctionnalité de cette protéine : une protéine variante peut être exprimée à un taux normal mais ne pas être fonctionnelle ; alors qu'une autre protéine variante peut être exprimée

à un taux relativement faible mais être quand même en mesure de remplir sa fonction. L'équipe d'Arora a par exemple mis en évidence que le variant MLH1 R725H, qui montre une expression significativement diminuée, répond cependant à des agents génotoxiques de la même façon que la protéine sauvage (Arora *et al.*, 2017). Les expériences de Western Blot ne sont donc pas une démonstration irréfutable de la fonctionnalité d'une protéine variante, mais constituent néanmoins un argument dans l'interprétation des VSI.

Par ailleurs, l'expression des protéines dans ces expérimentations est presque systématiquement induite à partir d'un promoteur fort. Ainsi, une protéine variante qui aurait une faible expression en conditions physiologiques pourrait voir son taux d'expression se rapprocher de celui de la protéine sauvage de manière artificielle.

#### *b. Interaction avec les protéines partenaires*

Comme dit plus haut, un Western Blot peut déjà donner une idée de la capacité d'interaction de MLH1 et MSH2 avec PMS2 et MSH6/MSH3, respectivement. Cependant, une étude *in vitro* par analyse *pull-down*, notamment par co-immunoprécipitation, peut également être entreprise (Andersen *et al.*, 2012 ; Fan *et al.*, 2007 ; Guerrette *et al.*, 1999). Le domaine d'interaction de MLH1 avec PMS2 correspond aux acides aminés 506 à 675 alors que le domaine d'interaction de PMS2 avec MLH1 correspond aux acides aminés 675 à 850 (Guerrette *et al.*, 1999). Néanmoins, les résidus qui ne sont pas compris dans ces intervalles peuvent quand même être testés pour leur capacité d'interaction, car il n'est pas exclu qu'un changement d'acide aminé à une autre position de la protéine induise un mauvais repliement de la protéine qui affecte l'interaction avec sa protéine partenaire (Belvederesi *et al.*, 2008). On peut bien sûr tester l'interaction entre les partenaires au sein des hétérodimères MutS $\alpha$ , MutS $\beta$  et MutL $\alpha$  mais également l'interaction avec d'autres protéines nécessaires à la fonction de réparation des mésappariements, telles qu'EXO1 (Lützen *et al.*, 2008).

L'étude de l'interaction avec les protéines partenaires peut également être menée par test double-hybride *in vivo* dans la levure (Fan *et al.*, 2007 ; Gammie *et al.*, 2007 ; Kondo *et al.*, 2003 ; Zhang *et al.*, 2018). De la même manière que dans le test *in vitro*, l'interaction avec EXO1 peut être analysée (Kondo *et al.*, 2003).

Le test double hybride *in vivo* dans la levure et le test GST pull-down *in vitro* sont tous les deux basés sur des protéines de fusion : le « tag » peut changer la conformation de la protéine, ce qui peut rendre certains sites d'interaction inaccessibles. La limite spécifique au test GST pull-down est la haute concentration de protéines utilisée, qui peut faciliter la liaison entre des protéines qui en conditions endogènes n'arriveraient pas à se lier (Lützen *et al.*, 2008). Enfin, la contrainte principale du test double

hybride est la nécessité que la protéine de fusion soit bien adressée au noyau pour espérer voir une interaction.

### *c. Localisation nucléaire*

Tous les rôles des protéines du système MMR (réparation des mésappariements, signalisation apoptotique des dommages à l'ADN, prévention de la recombinaison homologue...) nécessitent qu'elles soient présentes dans le noyau. Ainsi, il est probable qu'une incapacité de la protéine à être importée au noyau soit un mécanisme possible de pathogénicité d'un VSI. Plusieurs études ont montré que quand la protéine MLH1 (fusionnée à la protéine fluorescente GFP) est exprimée dans une lignée MMR-déficiente à partir d'un plasmide, MLH1 est presque entièrement détectée dans le noyau. Quand un plasmide permettant l'expression de la protéine PMS2 est ajouté, l'intégralité des protéines MLH1 est dans le noyau. Au contraire, quand la protéine PMS2 est exprimée seule, elle est majoritairement présente dans le cytoplasme (Borràs *et al.*, 2012 ; Christensen *et al.*, 2009 ; Raevaara *et al.*, 2005). Certaines protéines MLH1 variantes ne se comportent pas de la même façon que la protéine sauvage et peuvent présenter deux catégories de localisations cellulaires anormales. Soit elles sont majoritairement exprimées dans le cytoplasme en absence de PMS2 et leur proportion nucléaire augmente très peu en présence de PMS2 ; soit elles ont une expression majoritairement nucléaire en absence de PMS2, mais cette proportion n'augmente pas en présence de PMS2 (Raevaara *et al.*, 2005).

Par ailleurs, ce test permet indirectement de déterminer la capacité de protéines MSH2 variantes à dimériser avec MSH6 car MSH2 n'a pas de NLS (*Nuclear Localization Signal*, signal de localisation nucléaire) et nécessite la dimérisation avec son partenaire MSH6 pour être importée au noyau (Belvederesi *et al.*, 2008).

L'avantage de ce test *in vivo* est que les protéines sont exprimées dans un modèle cellulaire mammifère, ce qui permet un repliement et une maturation des protéines théoriquement optimaux. Cependant, les protéines d'intérêt sont fusionnées à une protéine fluorescente, ce qui peut influencer leur conformation ; de plus, elles sont surexprimées, ce qui peut retentir sur leur localisation (Lützen *et al.*, 2008).

### *d. Différentes étapes de la fonction de réparation du mésappariement*

La fonction de réparation du mésappariement, peut être étudiée dans son ensemble (voir ci-dessous), mais peut également être divisée en plusieurs étapes, qui peuvent chacune être limitante pour la fonction globale de réparation.

La première étape est la liaison au mésappariement, assurée par le complexe MutS $\alpha$  ou MutS $\beta$ . La protéine MMR d'intérêt avec le VSI ainsi que sa protéine partenaire sont transcrites et traduites *in vitro* (Lützen *et al.*, 2008) ou produites dans des cellules d'insecte (Ollila *et al.*, 2008a ; Ollila *et al.*, 2008) puis mises en présence d'un substrat d'ADN contenant un mésappariement ou une boucle d'insertion/délétion. Deux types d'expériences permettent d'étudier la liaison à l'ADN. Dans un premier type, le substrat est fixé sur un support et une dose croissante du complexe MutS $\alpha$  ou MutS $\beta$  est ajoutée. L'interaction entre le complexe et l'ADN est analysée par résonance de plasmon de surface (Cyr *et al.*, 2012 ; Heinen et Juel Rasmussen, 2012). Un autre type d'expérience consiste à réaliser un retard sur gel, qui permet de déterminer si les protéines sont capables d'interagir avec le mésappariement (si les protéines sont effectivement en complexe avec l'ADN, elles migreront plus lentement que les protéines non liées) (Lützen *et al.*, 2008 ; Ollila *et al.*, 2008a ; Ollila *et al.*, 2008). L'ajout d'ATP à ces deux types d'expérience permet également d'analyser la capacité de la protéine à former un « *sliding clamp* ».

La deuxième étape est la liaison à l'ATP puis son hydrolyse, qui constitue l'activité enzymatique des complexes MutS. L'activité ATPase globale est mesurée en mettant en présence l'hétérodimère, le substrat à réparer et de l'ATP puis en mesurant la fraction d'ATP hydrolysé (Cyr *et al.*, 2012 ; Heinen *et al.*, 2002). Pour étudier en particulier la liaison de l'ATP, l'hétérodimère est mis en présence d'un analogue de l'ATP qui est faiblement hydrolysable (ATP $\gamma$ S). La quantité d'ATP $\gamma$ S détectée après filtration est proportionnelle à la capacité du dimère à lier l'ATP (Cyr *et al.*, 2012).

#### *e. Recombinaison homologue*

Une équipe des Pays Bas a développé une méthode de caractérisation fonctionnelle des VSI basée sur l'introduction du VSI dans des cellules souches embryonnaires de souris (Wielders *et al.*, 2011 ; Wielders *et al.*, 2013 ; Wielders *et al.*, 2014). Plusieurs essais ont été réalisés, notamment l'étude de la recombinaison homologue. L'expérience consiste à introduire dans les cellules un vecteur 99,4% identique à une séquence du génome et voir si la protéine portant le VSI est capable d'empêcher la recombinaison entre les deux séquences, comme le ferait la protéine sauvage.

#### *iii. Etude d'une fonction du système MMR dans sa globalité*

Les deux grands rôles remplis par le système MMR sont la réparation des mésappariements et la signalisation apoptotique en réponse à certains dommages à l'ADN. Des tests fonctionnels permettent de s'intéresser à la capacité de la protéine variante à remplir l'une de ces fonctions dans sa globalité, sans chercher à s'intéresser au mécanisme sous-jacent (perte d'expression de la protéine,

perte d'interaction avec les protéines partenaires...). En effet, c'est l'abolition de la fonction dans son intégralité qui provoque la tumorigenèse chez le patient.

*a. Etude du rôle de réparation des mésappariements*

L'approche la plus souvent utilisée pour étudier la fonction de protéines MMR variantes est de tester leur capacité à réparer les mésappariements, rôle le plus connu du système MMR. Celle-ci peut être étudiée de trois manières principales : grâce à un test *in vivo* dans la levure *Saccharomyces cerevisiae* ou dans des cellules humaines ; ou avec un test *in vitro*.

Modèle levure :

Le premier test à avoir été développé consiste à réaliser une complémentation fonctionnelle chez une levure MMR-déficiente (suppression du phénotype mutateur) : une levure inactivée sur le gène MMR d'intérêt est transformée par un plasmide contenant le VSI à tester. Si le variant est non pathogène, la protéine exprimée à partir de ce plasmide sera capable de compléter le système MMR de la levure. La capacité ou non à réparer les mésappariements, est détectée grâce à un gène de la levure modifié pour contenir un microsatellite. L'instabilité de ce microsatellite provoque un phénotype particulier et on cherche donc les colonies qui ont perdu ou acquis ce phénotype (Drotschmann *et al.*, 1999a ; Drotschmann *et al.*, 1999 ; Shimodaira *et al.*, 1998). De manière intéressante, l'étude de Wanat, dans laquelle les variants ont été testés en parallèle sur deux souches de levures différentes, montre que les mêmes VSI sont susceptibles de provoquer différents taux de mutations selon la souche, surtout ceux conférant un phénotype faible ou intermédiaire (Wanat *et al.*, 2007).

Un autre moyen d'évaluer la fonction de réparation des mésappariements chez la levure est de tester l'effet dominant mutateur. L'introduction d'un gène MMR humain fonctionnel dans une levure a un effet dominant : le système MMR de la levure n'est plus fonctionnel et celle-ci présente un phénotype MSI. Cet effet dominant mutateur est certainement dû au fait que la protéine humaine fonctionnelle est capable d'interagir avec sa protéine partenaire, mais que cet hétérodimère hybride n'est pas capable de remplir sa fonction (Trojan *et al.*, 2002). Des levures MMR-compétentes sont transformées par un plasmide porteur de la séquence du VSI humain à tester. Le phénotype MSI, qui apparaîtra si le variant est non pathogène, peut être détecté soit sur un gène de la levure modifié pour contenir un microsatellite (Drotschmann *et al.*, 1999a ; Drotschmann *et al.*, 1999b ; Hardt *et al.*, 2011) soit sur un plasmide rapporteur de l'instabilité microsatellitaire (Shimodaira *et al.*, 1998 ; Takahashi *et al.*, 2007). Il est intéressant de tester l'instabilité microsatellitaire sur un plasmide plutôt que sur le génome. En effet, d'une part, le taux de mutations des séquences répétées d'un plasmide épisomal

est plus important que le taux de mutations des gènes chromosomiques. D'autre part, le plasmide est présent en 20 à 100 copies en moyenne dans la levure et il y a donc plus de probabilités qu'une mutation advienne sur le plasmide que sur le génome (Shimodaira *et al.*, 1998). L'inconvénient de cette méthode est que l'interprétation du phénotype dominant mutateur peut dépendre du gène rapporteur (Takahashi *et al.*, 2007) et du vecteur d'expression utilisés (Shcherbakova et Kunkel, 1999).

Enfin, le troisième test développé chez la levure consiste à modifier le gène orthologue de la levure haploïde afin de mimer un VSI et d'étudier l'éventuel phénotype MSI qui en découle grâce à un ou plusieurs gènes de la levure modifié(s) pour contenir des microsatellites (Martinez et Kolodner, 2010 ; Shcherbakova et Kunkel, 1999) ou grâce à un plasmide rapporteur (Polaczek *et al.*, 1998). Au lieu de ne modifier que le seul acide aminé d'intérêt dans le gène endogène de la levure, on peut également remplacer, chez la levure, l'intégralité du gène MMR contenant le VSI ainsi que du gène MMR partenaire par les gènes humains (Vogelsang *et al.*, 2009). Cette méthode permet une expression endogène à partir d'une seule copie du gène d'intérêt sous le contrôle du promoteur originel et permet de préserver le contexte génétique de la levure. De plus, l'absence de plasmides (d'expression du gène MMR d'intérêt et/ou rapporteur) permet de faire pousser les levures en milieu riche sans agent de sélection, ce qui permet de concevoir un test plus sensible et plus précis pour la caractérisation de potentiels phénotypes MSI atténués ou intermédiaires (Martinez et Kolodner, 2010).

L'une des principales limites du modèle levure est que, de manière générale, seuls les variants présents à des positions qui sont conservées entre la levure et l'homme peuvent être analysés. En effet, seules les extrémités N-ter et C-ter des protéines MLH1 humaines et levure présentent une homologie significative (70% pour la région N-ter) alors que la région centrale de la protéine est moins conservée (Kondo *et al.*, 2003 ; Shcherbakova et Kunkel, 1999). Pour pallier à ce problème, l'équipe d'Ellison a développé un gène MMR de fusion qui code pour une protéine hybride homme-levure, c'est-à-dire un gène MMR de levure comportant une portion humaine, capable d'avoir une activité MMR normale *in vivo* dans la levure (Ellison *et al.*, 2001 ; Ellison *et al.*, 2004). On peut alors tester les résidus qui sont présents sur la partie humanisée de ce gène de fusion. Différents gènes de fusion sont produits avec différentes portions du gène humain, pour pouvoir analyser un maximum de positions différentes. Ce gène de fusion est exprimé sur un plasmide, qui permet de réaliser un test de complémentation fonctionnelle dans une souche de levure déficiente pour ce gène MMR.

#### Modèle cellules humaines :

La fonction de réparation des mésappariements peut également être étudiée *in vivo* dans un modèle cellulaire humain. Une lignée cellulaire humaine déficiente pour le gène MMR d'intérêt est transduite par un système lentiviral qui permet l'expression stable du VSI grâce à une sélection à la puromycine. Les cellules sont ensuite transfectées par un plasmide portant le gène de la GFP contenant

un mésappariement dans la séquence codante, qui conduit à l'apparition d'un codon stop prématuré ne permettant pas en l'état l'expression de la protéine fluorescente GFP. La restauration du système MMR est donc détectée par la présence de cellules exprimant la GFP (Mastrocola et Heinen 2010 ; Cyr *et al.*, 2012).

#### Modèle *in vitro* :

Le principe repose sur la production de la protéine MMR variante à tester ainsi que celle de sa protéine partenaire dans des cellules d'insecte *Spodoptera fugiperda* 9. Les hétérodimères sont ensuite extraits et mis en présence d'un extrait nucléaire d'une lignée cellulaire humaine déficiente pour le gène que l'on veut tester (HCT116 MLH1<sup>-/-</sup>, LoVo MSH2<sup>-/-</sup>, HCT15 MSH6<sup>-/-</sup>) et d'un substrat d'ADN contenant un mésappariement ou une boucle d'insertion/délétion au sein d'un site de restriction. Si la protéine variante assure sa fonction, l'altération est réparée et le site de restriction est restauré. Cette restauration est visible après ajout de l'enzyme de restriction correspondante et migration du produit par électrophorèse sur gel (Christensen *et al.*, 2009 ; Kantelinen *et al.*, 2011 ; Kantelinen *et al.*, 2012 ; Kariola *et al.*, 2002 ; Kariola *et al.*, 2003 ; Kariola *et al.*, 2004 ; Nyström-Lahti *et al.*, 2002 ; Ollila *et al.* 2008a ; Ollila *et al.*, 2008b ; Ollila *et al.*, 2006 ; Raevaara *et al.*, 2005). Une variante est de réaliser la transcription et la traduction des protéines grâce à un kit *in vitro*, pour s'affranchir des cellules d'insecte (Drost *et al.*, 2010). De même, le substrat peut varier : le mésappariement peut porter un marqueur fluorescent. Ainsi, le taux de réparation n'est pas détecté par migration sur gel mais par une analyse de fragments fluorescents (Drost *et al.*, 2010 ; Andersen *et al.*, 2012).

Une autre méthode consiste à exprimer un plasmide porteur du gène MMR d'intérêt avec le VSI dans une lignée déficiente pour ce gène. Les extraits cytoplasmiques peuvent être mis en présence du même type de substrat que précédemment (Borràs *et al.*, 2013 ; Hinrichsen *et al.*, 2015 ; Kosinski *et al.*, 2010) ou en présence d'un ADN de bactériophage contenant un mésappariement dans la séquence codante du gène *lacZ*. L'ADN est ensuite purifié et introduit par électroporation dans des bactéries *E. coli* déficientes pour l'homologue du gène MMR d'intérêt : l'efficacité de réparation est proportionnelle au nombre de colonies bleues (Takahashi *et al.*, 2007 ; Trojan *et al.*, 2002).

L'avantage de ces tests *in vitro* est que tous les éléments requis peuvent être produits et stockés en grande quantité et l'expérimentation en elle-même ne requiert que du matériel de laboratoire relativement basique. Par contre, les tests *in vitro* ne détectent pas les variants qui affectent la localisation, l'expression ou la stabilité de la protéine, ce qui est susceptible de résulter en un grand nombre de faux négatifs (variants pathogènes mais qui répondent comme des variants neutres dans cet essai). Il est donc conseillé que les variants capables d'exprimer une fonction MMR dans ces tests soient testés pour d'autres fonctions (Couch *et al.*, 2008 ; Drost *et al.*, 2010).

### Modèle mutagenèse aléatoire :

Dans toutes les études que nous avons mentionnées ci-dessus, l'équipe commence par choisir un certain nombre de VSI qu'elle souhaite analyser avec la méthode qu'elle développe. Une autre approche est de réaliser une mutagenèse aléatoire, de tester les différents variants obtenus (sans connaître précisément leur séquence), de déterminer ceux qui présentent un caractère pathogène, de les séquencer puis de comparer avec les données de la littérature pour voir si certains de ces variants ont déjà été décrits chez des patients suspectés d'être atteints du syndrome de Lynch. L'équipe d'Ellison travaille sur un gène hybride humain-levure, qui est soumis à une mutagenèse sur les fragments humanisés du gène grâce à une PCR sensible aux erreurs : ce gène est ensuite exprimé dans une levure déficiente pour le gène *MLH1* (Ellison *et al.*, 2004). Ces levures sont soumises à un test de réparation des mésappariements *in vivo* : celles qui présentent un défaut MMR sont séquencées pour détecter la mutation qui a causé ce phénotype.

#### *b. Etude du rôle de signalisation des dommages à l'ADN*

Caractériser les VSI grâce à la capacité des protéines MMR variantes à remplir la fonction de signalisation des dommages à l'ADN est une approche beaucoup moins courante et seules quelques équipes s'y sont intéressées. Ces expérimentations reposent sur la réponse de cellules exprimant le variant à des agents méthylants (MNNG ou MNU) ou à la 6-TG, molécules qui induisent des lésions reconnues par le système MMR.

### Modèle humain :

Une lignée cellulaire *MLH1*-déficiente est transfectée par un plasmide permettant l'expression de la protéine *MLH1* variante. Les cellules sont exposées au MNU. Après un certain temps de culture, les colonies de cellules survivantes sont fixées, colorées et comptées (Blasi *et al.*, 2006). Une autre équipe a réalisé des expérimentations similaires avec des cellules *MSH2*-déficientes exprimant une protéine *MSH2* variante à partir d'un plasmide exposées au MNNG (Mastrocola et Heinen, 2010). La signalisation apoptotique en aval de la reconnaissance des lésions par le système MMR peut être analysée en étudiant le cycle cellulaire et l'activation des points de contrôle du cycle cellulaire (la signalisation de lésions à l'ADN par une protéine MMR fonctionnelle entraîne un blocage du cycle cellulaire et donc une augmentation du nombre de cellules en phase S) (Mastrocola et Heinen, 2010) ou encore en recherchant l'activation des acteurs de la voie de signalisation des dommages à l'ADN, tels que ATM/ATR ou  $\gamma$ H2AX (Arora *et al.*, 2017).

### Modèle souris :

L'équipe de te Riele utilise des cellules souches embryonnaires murines. Dans une première série d'études, la séquence du variant d'intérêt est introduit de façon bi-allélique au sein du gène endogène. Les cellules subissent un traitement au MNNG ou à la 6-TG puis les colonies de cellules survivantes sont fixées, colorées et comptées (Wiolders *et al.*, 2011 ; Wiolders *et al.*, 2013 ; Wiolders *et al.*, 2014). La même équipe a ensuite légèrement modifié ce test et réalise une caractérisation du VSI par sélection à la 6-thioguanine, qui n'est toxique que pour les cellules MMR-compétentes. Ainsi, les cellules capables de survivre sont celles qui portent un VSI pathogène. Les cellules porteuses de VSI pathogènes peuvent éventuellement être analysées plus en avant grâce à d'autres tests (expression des protéines, instabilité des microsatellites, réponse au MNNG et à la 6-TG) (Houllberghs *et al.*, 2016 ; Houllberghs *et al.*, 2017). L'avantage de l'utilisation des cellules souches embryonnaires de souris réside principalement dans la création de modèles de souris à partir de ces cellules, afin d'étudier le VSI dans un modèle *in vivo*. Par ailleurs, le variant est exprimé à partir du locus endogène, ce qui est rarement le cas quand on regarde l'ensemble des tests fonctionnels développés. Cependant, de l'avis même des auteurs, la forte expression physiologique des protéines MMR dans un contexte embryonnaire provoque tout de même une légère surexpression par rapport à ce qu'on pourrait observer dans des cellules non embryonnaires

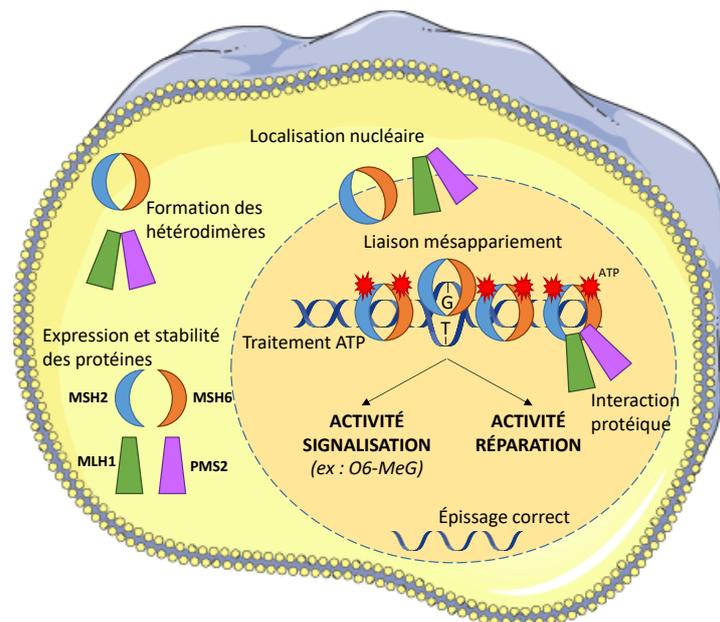
L'équipe d'Avdievich est même allée plus loin : elle a créé une lignée de souris portant la mutation Mlh1<sup>G67R</sup> et l'a caractérisé en termes de survie, de développement de tumeurs, d'expression de la protéine dans les thymocytes, de capacité d'interaction avec PMS2, de capacité à réparer les mésappariements (étude du phénotype MSI), de réponse au cisplatine et de capacité à réaliser la méiose (Avdievich *et al.*, 2008).

### Modèle mutagenèse aléatoire :

L'équipe de Drost travaille sur des cellules souches embryonnaires de souris qui ont déjà une mutation mono-allélique dans le gène *MSH2* (Drost *et al.*, 2013). Ces cellules sont exposées à l'ENU (N-éthyl-N-nitrosourea), mutagène qui provoque des mutations ponctuelles, puis à la 6-TG, qui permet de sélectionner les cellules MMR-déficientes. Le gène *MSH2* est ensuite séquencé dans les cellules tolérantes à la 6-TG pour déterminer la mutation présente.

Dans cette partie sur les tests permettant d'apprécier le caractère pathogène ou non d'un VSI, nous avons vu que de nombreux tests ont été développés à l'heure actuelle, que ce soit sur modèle cellulaire de levure ou de mammifère, *in vitro* ou *in vivo*. Les différents aspects fonctionnels des protéines MMR qui peuvent être étudiés sont résumés sur la **FIGURE 10**. Cependant, il est important de noter qu'il n'existe pas de test fonctionnel « *gold standard* » qui serait utilisé systématiquement

pour aider à la caractérisation des VSI des gènes MMR. L'absence de test de référence rend difficile la validation des tests fonctionnels. Ainsi, le comportement d'un VSI dans un test fonctionnel donné ne permet pas forcément une classification définitive. L'accumulation des résultats de plusieurs tests permet toutefois d'apporter des arguments pour aider la classification (**FIGURE 11 en ANNEXE 2**). Ainsi, une classification optimale nécessite l'intégration de tous les arguments disponibles, tels que les résultats des tests fonctionnels, les prédictions *in silico*, la ségrégation du syndrome dans les familles, les données épidémiologiques et les caractéristiques clinicopathologiques (Couch *et al.*, 2008).



**FIGURE 10 | Résumé des tests de caractérisation fonctionnelle des variants des gènes MMR existants**

Ces essais testent les conséquences du VSI sur l'épissage de l'ARN, le taux d'expression de la protéine, son aptitude à interagir avec ses protéines partenaires, la localisation subcellulaire, la liaison au mésappariement, la liaison et l'hydrolyse de l'ATP, l'activité de réparation et l'activité de signalisation des dommages provoqués par certains agents (blocage du cycle cellulaire et apoptose). *Adapté de (Heinen et Juel Rasmussen, 2012).*

### 5. Intégration des données somatiques

Comme nous l'avons évoqué, un cancer MSI à un âge jeune peut être expliqué par une double mutation somatique (une mutation somatique et une perte d'hétérozygotie ou deux mutations somatiques) et cette condition est appelée « *Lynch-like* ». Ces mutations somatiques ont forcément un aspect délétère, puisqu'elles ont conduit au développement d'une tumeur MSI. Une étude très récente propose donc d'ajouter comme argument dans la classification des VSI des déductions obtenues à partir des données mutationnelles somatiques tumorales (Shirts *et al.*, 2018). En effet, une même altération pathogène peut conduire au même phénotype moléculaire (instabilité microsatellitaire) dans un contexte sporadique et dans un contexte constitutionnel. Ainsi, une altération dont la

classification est considérée incertaine dans un contexte constitutionnel est probable d'être pathogène si elle a aussi été détectée dans une tumeur MSI due à une double mutation somatique d'un gène MMR. Par conséquent, les informations somatiques peuvent aider à l'interprétation des VSI.

### C. La classification des variants en 5 niveaux

L'ACMG (*American College of Medical Genetics and Genomics*), institution interdisciplinaire qui a notamment pour rôle d'établir des normes de politiques de soins et de pratiques de laboratoire, a publié des recommandations pour aider à l'interprétation des variants de n'importe quel gène responsable d'un syndrome génétique ayant un mode de transmission mendélien. Une classification des variants en 5 catégories a été proposée : classe 1 – clairement non pathogène, classe 2 – probablement non pathogène, classe 3 – signification incertaine, classe 4 – probablement pathogène et classe 5 – clairement pathogène (Richards *et al.*, 2015). L'objectif de ces critères de classification est de limiter la variabilité inter- et intra-laboratoire dans l'interprétation des VSI en génétique moléculaire afin que l'interprétation d'un variant soit la même pour tous les patients porteurs de ce variant. Cette classification est présente en **ANNEXE 1**.

Il est toutefois intéressant de développer une classification des variants qui soit spécifique d'un gène ou d'un syndrome génétique donné vu que l'applicabilité et le poids donné à chaque critère peuvent varier selon le gène et le syndrome (Richards *et al.*, 2015). Les critères de classification spécifiques des gènes MMR, détaillés en **ANNEXE 2**, ont été rédigés en collaboration avec le comité d'interprétation des variants InSiGHT et ont été publiés en 2014 (Thompson *et al.*, 2014). Dans le cas où le variant détecté chez le cas index est assigné à la classe 5, le diagnostic génétique est proposé aux apparentés dans un contexte clinique. Cependant, dans le cas où le variant est classé en classe 4, 3 ou 2, le diagnostic génétique peut également être proposé aux apparentés dans un contexte de recherche, pour aider à l'interprétation plus en avant du variant (Plon *et al.*, 2008). Les variants des gènes *MSH6* et *PMS2* sont plus difficiles à classer dû à la plus faible pénétrance et les âges plus tardifs au développement des tumeurs (Thompson *et al.*, 2014). L'interprétation est réalisée dans l'état actuel des connaissances et des informations disponibles. La classe dans laquelle a été assignée un variant est donc susceptible d'être modifiée au fur et à mesure qu'on accumule des données (Plon *et al.*, 2008).

# TRAVAUX DE RECHERCHE



**FUNCTIONAL CHARACTERIZATION OF *MLH1* AND *MSH2* VARIANTS OF UNKNOWN SIGNIFICANCE BASED ON METHYLATION TOLERANCE**

BOUVET Delphine<sup>1</sup>, BODO Sahra<sup>1</sup>, MUNIER Annie<sup>2</sup>, GUILLERM Erell<sup>1,3</sup>, BERTRAND Romane<sup>1</sup>, COLAS Chrystelle<sup>1,4</sup>, COULET Florence<sup>1,3</sup>, DUVAL Alex<sup>1</sup>, MULERIS Martine<sup>1</sup>

1. Sorbonne Université, Inserm, Centre de recherche Saint-Antoine, Paris, France

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

3. Laboratoire d'Oncogénétique et d'Angiogénétique, GH Pitié-Salpêtrière, Paris, France

4. Department of Genetics, Institut Curie, Paris, France



## Introduction

Mon équipe d'accueil s'intéresse aux cancers de type MSI, caractérisés par une déficience du système MMR. Au sein de cette thématique, je me suis intéressée au cours de ma thèse au syndrome de Lynch, syndrome génétique de prédisposition aux tumeurs MSI dû à une mutation délétère mono-allélique germinale d'un gène MMR. Le diagnostic repose sur l'analyse constitutionnelle de l'ADN du patient pour mettre en évidence une telle altération pathogène. Cependant, environ 30% des altérations détectés dans les gènes *MLH1* et *MSH2* (gènes les plus couramment affectés dans le syndrome de Lynch) sont des variants de signification incertaine (VSI) qui ne permettent pas le diagnostic pour l'individu et ses apparentés.

Le laboratoire a récemment développé un test fonctionnel d'aide au diagnostic du syndrome CMMRD, variante rare du syndrome de Lynch dû à une mutation bi-allélique constitutionnelle d'un gène MMR. Ce test a fait l'objet d'un brevet et est actuellement disponible en diagnostic de routine dans le laboratoire commun de biologie et génétique moléculaire de l'hôpital Saint-Antoine. Cette méthode repose sur l'étude des deux caractéristiques spécifiques des cellules MMR-déficientes : leur phénotype d'instabilité microsatellitaire et de tolérance à la méthylation.

## Résultats

Dans la continuité directe de cette étude, nous avons adapté cette approche pour proposer un test de caractérisation fonctionnelle des VSI des gènes *MLH1* et *MSH2* dans le contexte du syndrome de Lynch, basé sur le mécanisme de tolérance à la méthylation. Ce test s'appuie sur le principe de la complémentation fonctionnelle : la capacité d'un VSI *MLH1* (ou *MSH2*) à restaurer la fonction de signalisation des dommages à l'ADN du système MMR et donc la sensibilité à la méthylation dans la lignée cellulaire HCT116 *MLH1*-déficiente (ou LoVo *MSH2*-déficiente) est évaluée. Pour cela, les cellules sont transfectées par la séquence du VSI à tester, triées par cytométrie en flux pour ne conserver que celles qui expriment le VSI puis exposées au MNNG, un agent méthylant. Un test de clonogénicité permet d'évaluer la survie cellulaire des cellules transfectées en réponse à l'agent méthylant. Nous avons d'abord testé un panel de 40 variants *MLH1* et *MSH2* dont le caractère pathogène ou neutre est connu. Notre test a permis de classer correctement 39/40 variants, montrant la pertinence de notre approche méthodologique. La technique a ensuite été appliquée à une série de 49 VSI (25 *MLH1* et 24 *MSH2*) à caractériser. Globalement, notre test a permis de proposer une classification pour 41/49 (83%) des VSI : 12 VSI (6 *MLH1* et 6 *MSH2*) ont restauré la cytotoxicité vis-à-vis des dommages liés à la méthylation et sont considérés comme compétents et 29 VSI (13 *MLH1* et 16 *MSH2*) étaient incapables de restaurer l'apoptose à l'agent méthylant et sont considérés comme déficients ; 8 VSI (6 *MLH1* et 2 *MSH2*) présentent un taux de survie intermédiaire entre celui des contrôles pathogènes et celui des contrôles neutres et leur classification demeure incertaine.

La comparaison avec les autres éléments (données cliniques et somatiques des individus porteurs de ces variants, prédictions *in silico* et données fonctionnelles déjà existantes pour une petite proportion des VSI étudiés) confirment nos résultats et ont permis de valider notre approche.

### Conclusion

A l'heure actuelle, de nombreux tests de caractérisation fonctionnelle des VSI des gènes MMR ont été développés. Cependant, ces méthodes ont très rarement été testées au préalable sur une large cohorte cas-contrôles et ils n'ont donc pas été validés en termes de sensibilité, spécificité et reproductibilité. En pratique aucun test n'est, à l'heure actuelle, réalisé systématiquement sur tous les VSI mis en évidence et cet argument d'interprétation n'est pas disponible pour la grande majorité des VSI. Or les données fonctionnelles, quand elles sont disponibles, représentent un argument fort dans les critères de classification des VSI. Ainsi, disposer de données fonctionnelles pour l'ensemble des VSI représenterait un atout majeur pour aider à caractériser leur pathogénicité. Nous avons donc développé un test relativement simple et rapide qui a l'avantage d'utiliser un modèle cellulaire humain, contrairement à d'autres tests réalisés dans des cellules murines ou de levure, et d'analyser une fonction du système MMR dans sa globalité plutôt qu'un aspect restrictif de la fonctionnalité de la protéine. Néanmoins, la principale limitation de notre approche est l'expression ectopique des protéines variantes à partir d'un promoteur fort qui conduit à un taux protéique largement supérieur au niveau endogène et qui est susceptibles de masquer les effets d'un variant hypomorphe. À terme, ce test de signalisation des dommages à l'ADN pourrait être, en complément d'un test de réparation des mésappariements à définir, disponible en routine sur une plateforme de diagnostic afin d'étudier en prospectif tous les VSI détectés chez des patients suspectés d'être atteints du syndrome de Lynch.

## ABSTRACT

**Background and aims:** The DNA mismatch repair (MMR) pathway is essential in maintaining genomic stability through its role in DNA repair and DNA damage sensing. Loss of MMR underlies the hereditary cancer disease Lynch syndrome (LS). Germline mutations in *MLH1* or *MSH2* account for approximately 50% and 40 % of LS patients, respectively and about 30% of these are variants whose functional implications are unclear (called variants of unknown significance or VUS). Lack of classification of these VUS precludes LS diagnosis for carriers and their relatives. We aimed to develop a methylation tolerance assay in a human model system in order to propose a functional characterization of *MLH1* and *MSH2* VUS.

**Methods:** We transiently introduced *MLH1* (or *MSH2*) variants into the *MLH1*- (or *MSH2*-) null human colorectal cancer cell line HCT116 (or LoVo). After selection of transfected cells using flow cytometry, we evaluated the MMR function *in vivo* through the cell response to DNA-methylation damage. A clonogenic assay provided a readout for MMR activity. A panel of 89 (44 *MLH1* and 45 *MSH2*) variants was tested including 40 proven pathogenic or neutral variants (validation set), and 49 (25 *MLH1* and 24 *MSH2*) VUS identified in suspected-LS patients (prospective set). The screening results were compared to clinical data from suspected-LS patients carrying these variant alleles, experimental *MLH1* protein expression data, published functional and *in silico* prediction results.

**Results:** Our assay correctly classified 39/40 variants from the validation set. Among the prospective set, 12 VUS classified as damaging, showed values comparable to known pathogenic mutants, 29 classified as neutral, restored cytotoxicity to methylation with values comparable to known non-pathogenic variants, and 8 showed intermediate restoration of MMR capacity precluding their classification. Available clinical, somatic and expression data are in line with our screen results.

**Conclusion:** This proof of concept study validates our approach for the functional characterization of VUS for the 2 MMR genes mainly involved in LS. Our screen may constitute an important tool for the comprehensive diagnostic procedure for LS-associated *MLH1* or *MSH2* VUS.

## INTRODUCTION

Monoallelic germline mutations in one of the 4 human DNA mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*) are responsible for the Lynch syndrome (LS) (formerly known as hereditary non-polyposis colorectal cancer), one of the most prevalent inherited cancer syndromes in man, which accounts for 1-3% of all colorectal carcinomas incidence<sup>1</sup>. One of the main functions of MMR pathway is to correct replication errors that escape proofreading by replicative DNA polymerases. As a consequence, a defective MMR results in increased spontaneous somatic mutation rates, especially in sequences comprising multiple repeats (so called microsatellite instability or MSI)<sup>2-5</sup>. In LS patients, somatic inactivation of the wild-type allele makes the cell MMR deficient, resulting in elevated mutation rates in microsatellites, which drives carcinogenesis<sup>6</sup>. Accordingly, LS patients have a great increased risk of developing colorectal cancer in their lifetimes in addition to certain extracolonic cancers such as endometrial, ovarian, stomach, pancreatic and multiple other cancers<sup>7</sup>. Analysis of microsatellite instability and loss of immunohistochemical staining for MMR proteins in tumor tissue can identify patients who may suffer from LS. However, sequence analyses must reveal a pathogenic mutation in one MMR gene for a definitive LS diagnosis. Germline variants are found primarily in either *MLH1* or *MSH2* genes, and to a smaller extent in *MSH6* or *PMS2* genes, which account for 40, 34, 18 and 8% of the unique germline sequence alterations, respectively<sup>8</sup>. Identification of a high-risk disease causing constitutional mutation in the patient guides the clinical management of his whole family with implications for genetic counseling and presymptomatic surveillance<sup>9</sup>. 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 work to develop and apply a five-tiered standardized classification based on multiple lines of evidence including variant and family characteristics and results from various functional assays<sup>10</sup>. Classes 5 and 4 indicate a pathogenic and likely pathogenic variant (mostly nonsense and frameshift mutations resulting in non-functional proteins that unambiguously abrogate MMR activity) respectively; classes 1 and 2 indicate a pathogenic and likely non-pathogenic variant (mostly intronic variants, non-synonymous and synonymous missense variants) respectively; and class 3 (primarily non-synonymous missense changes) indicates a variant of unknown significance (VUS) of which the pathological significance is unclear. VUS, which represent 31% and 28% of the *MLH1* or *MSH2* mutations respectively<sup>8</sup>, present a considerable clinical issue since the lack of classification of these variants precludes LS diagnosis for carriers and their relatives.

To help classify VUS, functional assays are needed. Indeed, they can demonstrate alterations in the activity of MMR gene that can be used to predict whether the variant predisposes to disease or alternatively has no significant influence on cancer risk. A multitude of functional assays in different

experimental systems have been developed <sup>11-13</sup> for review. Most have been performed in model organisms such as bacteria, yeast or mouse, which has intrinsic limitations. Characterization of LS mutations using a human expression system is more limited <sup>14-17</sup> for instance. 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 the whole MMR function (mutation rate, microsatellite instability, drug response). Moreover, in order to translate a functional effect to a clinical risk, it is of outmost importance to study a set of known deleterious and neutral variants in order to establish sensitivity, specificity and predictive values of the test. To date, the reproducibility and performances of almost all current MMR functional assays still remain to be determined and although they can provide functional evidence, they are not routinely used.

In order to overcome these limitations, we developed an *in vivo* screen, using a human expression model, for the functional characterization of *MLH1* and *MSH2* VUS that assesses a whole MMR function, i.e. its DNA damage-induced apoptosis function. Indeed, besides its role in removal of replication errors, MMR is also involved in the processing of O6-methylguanine (O6-meGua), a DNA adduct induced by methylating agents <sup>18-22</sup>. However, MMR processing does not result in lesion repair but instead it signals DNA damage to apoptosis. Therefore, defective MMR function is also linked to cell tolerance to killing by methylating agents (so called methylation tolerance). Relying on this specific property, we recently developed a functional assay which has been successfully used for the characterization of MMR VUS identified in the constitutional MMR deficiency syndrome (CMMRD), a rare and severe variant of LS due to bi-allelic germline mutations in MMR genes <sup>23, 24</sup>. This functional test made it possible to diagnose CMMRD syndrome, especially when the sequencing was not contributive due to the detection of VUS <sup>25, 26</sup>. Unlike germinal cells of patients with CMMRD syndrome, which harbour a bi-allelic mutation of an MMR gene and therefore are MMR-deficient, the germ cells of LS patients are MMR-proficient since they retain one wild-type MMR allele. It is therefore not possible to evaluate the MMR function directly on the patient's cells as previously performed <sup>25</sup>. As a result, our experimental strategy was to evaluate *in vivo* the consequence of ectopic expression of *MLH1* (or *MSH2*) VUS on the restoration of the DNA damage response function of MMR in a human *MLH1*- (or *MSH2*)-null cell model. This assay was evaluated using a large panel of 89 variants which includes a validation set of 40 (i.e. 19 *MLH1* and 21 *MSH2*) previously characterized variants and a prospective set of 49 (i.e. 25 *MLH1* and 24 *MSH2*) VUS. Comparison of our screen results with clinical, somatic and protein expression data confirmed the relevance of such functional approach.

## RESULTS

### Proof of feasibility

#### Complementation of *MLH1* gene defect

The human colorectal adenocarcinoma cell line HCT116 presents only one single *MLH1* copy, which contains a mutation resulting in a truncated, non-functional protein<sup>27</sup>. Therefore, HCT116 cells have defective mismatch repair and they are known to be tolerant to methylation damage<sup>28</sup>. Hence, introduction of a variant in the cells 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 and exposed to MNNG (N-methyl-N'-nitro-N-nitroso-guanidine), a methylating agent that is highly toxic to MMR-proficient but only moderately toxic to MMR-deficient cells. A clonogenic assay provided a readout for MMR activity (Figure 1).

We first investigated whether restoration of cytotoxicity to methylation damage could be detected in the *MLH1*-null human HCT116 cells that have been transiently transfected with a plasmid that contains the wild-type (WT) *MLH1* cDNA sequence. As expected, parental HCT116 cells, which lack the expression of *MLH1*, survived mild MNNG selection and are tolerant up to 0.5  $\mu$ M MNNG. Interestingly, introduction of the WT *MLH1* led to a decrease of colonies that survived selection, indicative for a tangible restoration of MMR capacity. Cells were found on average about 2-fold more sensitive to 0.5 $\mu$ M MNNG than the control, i.e. HCT116 cells transfected with a random sequence (survival fraction of 39.4% and 89.6% respectively, after 2 pulses of 0.5  $\mu$ M MNNG) (Figure 2A). The kinetics of expression of the transgene allowed to control that the rate of transfected cells remained stable up to 72h-post transfection, i.e. at the time cells are exposed to MNNG treatments (Suppl Figure 1).

#### Complementation of *MSH2* gene defect

The human colorectal adenocarcinoma cell line LoVo, which contains a large *MSH2* homozygous deletion (from exon 3 to 8)<sup>29</sup>, was used to assess *MSH2* variants. As expected, LoVo cells survived mild MNNG selection. When transiently transfected with a plasmid that contains the wild-type (WT) *MSH2* cDNA sequence, LoVo cells were found on average about 3-fold more sensitive than LoVo cells transfected with a random sequence (survival fraction of 21.2% and 77.3% respectively, after 1 pulse of 1  $\mu$ M MNNG) (Figure 2B).

## Validation study

### MLH1 variants

We next sought to identify experimental conditions that would best discriminate 10 previously proven pathogenic (including 8 missense and 2 nonsense variants classified as class 5 in InSiGHT database) from 9 previously proven non-pathogenic (including 8 non-synonymous and 1 synonymous missense variants classified as class 1 in InSiGHT database) *MLH1* variants (suppl Table 1).

We exposed the sorted cells to a range of MNNG concentration and to different numbers of treatments and chose the conditions that allowed the best splitting of the 2 groups of variants (i.e. 2 and 3 pulses at 0.1  $\mu$ M MNNG) (Figure 3A), and calculated the mean survival score. Using these experimental conditions, all pathogenic variants displayed a mean survival score above 82.2% and the mean survival score of all non-pathogenic variants was below 55.2% (Figure 4A). Assuming a binary classification of the variants, i.e. variants can be either pathogenic or neutral but not hypomorphic, the Receiver Operating Characteristic (ROC) curve was constructed. It showed that our assay demonstrated a perfect discrimination with 100% sensitivity (69.15-100) and 100% specificity (71.51-100) between pathogenic and neutral variants for a cut-off value of 68.71%.

### MSH2 variants

The case-control study comprised 10 previously proven pathogenic (including 9 missense and 1 nonsense variants classified as class 4 or class 5 in InSiGHT database) and 11 previously proven neutral or likely neutral (including 9 non-synonymous and 2 synonymous missense variants classified as class 1 or class 2 in InSiGHT or UMD (Universal Mutation Database) databases) *MSH2* variants (suppl Table 1). We determined that the 2 sets of variants were the best discriminated using 1 or 2 treatments with 1  $\mu$ M MNNG (Figure 3B). Using these experimental conditions all pathogenic variants displayed a mean survival score above 58.90% and the mean survival score of all but one non-pathogenic variants was below 32.17% (Figure 4B). The ROC curve showed a 91.67% sensitivity (61.52-99.79) and 100% specificity (71.51-100) with a cut-off value of 45.54%.

## Prospective study

### MLH1 VUS

The functional assay conditions described above were used to evaluate 25 missense *MLH1* VUS including 21 variants classified as class 3 in the InSiGHT database and 4 private variants) (See Suppl Table 2). The position of all alterations on the MLH1 structure is presented in Suppl Figure 2A. Unlike the situation between the known pathogenic and non-pathogenic variants, two distinct groups were

not clearly observed since the VUS exhibited a large dispersion of responses to methylation damage (figure 4A). Six and 13 VUS showed a mean survival score within the values found in our validation set of variants (i.e. above 82.2% or below 55.2%, respectively). Thus, these variants could be considered as damaging or neutral, respectively. The mean survival scores of the 6 remaining VUS fell in between. Using the cut-off value previously determined by the ROC curve, 4 could be identified as damaging (they will further be called as “potentially damaging”) and 2 as neutral (further called as “potentially neutral”).

### MSH2 VUS

A set of 24 missense VUS including 23 variants classified as class 3 in the InSiGHT or UMD database and 1 private variant (See Suppl Table 3). The position of all mutations on the MSH2 structure is presented in Suppl Figure 2B. Two distinct groups were clearly observed (figure 4A) with 16 and 6 VUS showing a mean survival score within the values found in our validation set of variants. They were thus considered as neutral or damaging, respectively. The response of the 2 remaining VUS fell in between. Using the cut-off value determined by the ROC curve, they could be identified as damaging (further called as “potentially damaging”). An overview of the results for all VUS is presented in Table 1.

### **MLH1 and MSH2 protein expression**

Results of the functional screen were further compared to the MLH1 (or MSH2) protein levels assessed by Western blotting on sorted cells.

All but one known non-pathogenic variants resulted in MLH1 expression levels above 47% in our transfection system. This only exception was the V716M alteration, which was associated with a lower expression (22%) already noticed in previous report<sup>30</sup>. Accordingly, the 15 VUS that were detected as neutral or potentially neutral by our screen, retained moderate to high MLH1 expression level, i.e. 32-142%.

Besides, 8 of the 10 VUS that were identified as damaging or potentially damaging by our screen showed reduced MLH1 protein levels below 43%, which is consistent with the 8-34% MLH1 protein levels observed in most of the known pathogenic variants (Table 1, Suppl Figure 3A). Overall, there was a moderate inverse correlation ( $r^2=0.5$ ,  $P<0.000$ , Pearson correlation) between mean survival scores assessed by our functional screen and MLH1 expression levels (Figure 5).

The pattern was quite different for MSH2 since there was no correlation between protein expression and survival scores in the methylation assay ( $r^2=0.07$ ). A large overlap in protein expression was

observed between pathogenic and neutral variants, and between damaging and neutral VUS (Table 1, figure 5, Suppl Figure 3B).

### **Comparison with clinical, somatic and previous functional data**

Clinical and somatic data available from suspected-LS patients carrying the VUS screened in our study, together with previous functional results, are presented in Suppl Table 2 (for *MLH1* VUS) and Suppl Table 3 (for *MSH2* VUS).

#### *MLH1* variants

Interestingly, the damaging or potentially damaging phenotype found for the P640L VUS and the L622P, L260F and L582F VUS, respectively, were detected in patients with MSI LS-related tumors, often with an early age and with a family history evocative of LS. Moreover, the last 3 variants are the only ones of the series for which the absence of *MLH1* promoter methylation was reported, thus excluding a sporadic *MLH1*-related origin. In contrast, clinical data is not truly evocative of LS (non LS-related tumors and/or late onset CRC) for variants Y126N, A160V, V179D, E319K, K618E, which were classified as neutral or potentially neutral in our screen. For the later, methylation of the proximal part of *MLH1* promoter was also noted, which favors a sporadic MSI CRC. Thus, clinical and somatic data support our screening results.

Previous functional data, from cell free or yeast assays, were available for 7 VUS. All are also in line with our screen results. They support a neutral phenotype for variants A160V, N338S, R474Q and E523D, all of them being classified as non MMR damaging in our screen whereas they favor a pathogenic phenotype for variants G244V, P640L and L582F, the 2 former being classified as damaging, and the later as potentially damaging in our screen.

With the exception of 4 private unreported VUS, all other 21 VUS are classified as class 3 in the InSiGHT database. Most of them (14/21) are also classified as VUS in the UMD database and 4/21 do not appear. The remaining 3 are noted as pathogenic (P640L) or neutral (G454R and R474Q) in UMD. Accordingly, we found the former as MMR damaging unlike the 2 others.

#### *MSH2* variants

The damaging or potentially damaging phenotype found for the T33P, A500P, L602P, P622T, R638G, N671Y VUS and the L622P and R711Q VUS respectively, were detected in patients with MSI LS-related tumors, often with an early age in a family evocative of LS. Unfortunately, clinical and/or somatic data were absent or scarce for all VUS identified as neutral in our screen, except the A2T VUS. Moreover, this variant has been classified as non-pathogenic in the UMD database thus in agreement with our

result. Previous functional data were available for only 3 variants, all of them identified as non-damaging in our screen. For the T33P and A305T VUS, various previous assays led to the same interpretation. However, for the S723F VUS, several previous functional results supported an abrogated MMR function contrary to our screen result.

### **Performance of *in silico* prediction programs for the 49 VUS that have been experimentally screened**

We used 4 very popular *in silico* prediction programs (Align GVGD, SIFT, Mutation Taster and CADD) to predict the function of the 49 VUS that have been experimentally tested (Suppl Figure 4). It is uneasy to perform direct comparison because *in silico* prediction programs (i) frequently generate discordant results and (ii) tend to overestimate the pathogenicity of the variants (see the performance of *in silico* prediction programs for the 40 variants with known functional effects in Suppl Figure 5). Of note, all VUS with at least 2 predictions in favor of a non-pathogenic phenotype, i.e. 5 *MLH1* and 7 *MSH2* variants, belong to the group of VUS that have been classified as non-damaging in our screen.

## **DISCUSSION**

The problem of VUS is a major issue in the field of tumor genetics, regardless of the impacted gene or its function, since the lack of classification of these VUS precludes diagnosis for carriers and their relatives. Furthermore, the prevalence of VUS is expected to increase with routine use of next-generation sequencing techniques. To date, many functional characterization studies of MMR genes VUS have been performed in model organisms such as bacteria, yeast or mouse, or *in vitro*. Characterization of LS mutations using a human expression system is more limited. However, the performance (sensitivity, specificity) of most assays remains unknown due to the frequent lack of consistent case-control series. So, although they can provide functional evidence, they are not routinely used and there is no “gold standard” functional assay.

Most of the functional assays that study an MMR function as a whole, and not one or other restrictive aspect of the variant protein (expression, stability, ...), assessed the mismatch repair function of MMR. However, it is generally assumed that both the DNA repair and DNA damage response MMR functions are important for the suppression of tumorigenesis. For several years now, our laboratory has been interested in the DNA damage sensing function of MMR and its role in MSI-driven oncogenesis. Indeed, through the pathophysiological mechanism of methylation tolerance, MMR deficiency confers a selective advantage to the cells in the specific context of exposure to methylating or thiopurine agents. We and others have reported the occurrence of MMR-deficient cancers after

treatment with alkylants or thiopurines in various clinical contexts<sup>31-34</sup>, a similar mechanism is likely to play a major role in MSI colorectal cancers since compounds such as *N*-nitroso compounds (NOC) that are present in the diet or can arise endogenously in the colon lumen may lead to the formation of O<sup>6</sup>meG lesions<sup>35 for review</sup>.

In the assay we developed, the *MLH1* and *MSH2* variants were transiently expressed in a complementary defective human cell line and the resulting MMR activity of the transfected cells in DNA damage sensing was examined. In a direct comparison using 40 variants previously characterized according to pathogenicity, our assay correctly classified 39/40 variants, which yields 100% sensitivity and specificity for *MLH1*, and 91.67% sensitivity and 100% specificity for *MSH2*. These results outperform several currently employed tools such as *in silico* prediction tests and demonstrated the relevance of our methodological approach. The only discordant result concerns the *MSH2* (c.593A>G, E198G) mutation, which is classified as neutral in the InSiGHT database but was found as damaging in our screen. This mutation was identified in a Swedish family with five affected members (4 sibs and their father). The variant segregated in the family with MSI-negative colorectal cancers (age: 29, 40, 62 and 62 years old) and was not found in control subjects. A missense *MLH3* (T3826C) variant, inherited from the mother, was also present in all 4 sibs<sup>36, 37</sup>. Previous published functional assays favored the pathogenicity of the variant (no mismatch repair, defective interaction with all MSH2 partners, low protein expression in yeast assays)<sup>38, 39</sup> in agreement with our results. Thus, the arguments that led to the qualification of this variant as neutral may be debatable.

When the assay was applied to a prospective set of 49 *MLH1* or *MSH2* VUS, 12 (24%) and 29 (59%) could unambiguously be classified as damaging and neutral, respectively. For the 8 remaining VUS that showed intermediate restoring MMR capacity, the classification remains questionable. Clinical/somatic data, as well as published results from functional tests when available, are in line with the results of our screen. The only notable inconsistency concerned the *MSH2* S723F VUS, which we identified as neutral contrary to several published assays that supported a pathogenic phenotype (no mismatch repair in yeast and mammals, 59% expression in yeast)<sup>38, 40, 41</sup>. Such discrepancy can be related to the different cell model systems used in the studies. Another hypothesis is that the *MSH2* S723F mutation might uncouple the DNA mismatch repair (tested in the published studies) and the DNA damage-induced apoptosis (considered in our assay) function. Despite the fact that both functions were linked in the vast majority of the MMR mutants explored so far, examples of decoupling have been reported for the Msh2<sup>G674A</sup> and Msh6<sup>T1217D</sup> mutations in mice<sup>42, 43</sup> and in MSH-deficient yeast strain models<sup>44</sup>. Interestingly, all these variants exhibit a mutation in the ATPase domain, as is the case for the *MSH2* S723F mutation in our study. It has been proposed that DNA damage response and DNA repair pathways have different requirement for ATP processing, the MSH2-ATPase domain being

dispensable for the apoptosis signaling function of MutS $\alpha$ . As a result, for an overall functional characterization of LS-associated *MLH1* and *MSH2* VUS, we recommend the implementation of our methylation tolerance assay in combination with a mismatch repair assay (Figure 6).

We acknowledge that our study has a number of limitations. First of all, the expression vector we used may not be suitable for assessing the pathogenicity of hypomorphic alleles since the variant is placed under control of a strong promoter, which results in a very high expression unequaled in the physiological conditions (see the endogenous expression of MLH1 protein in Lovo compared to that found in HCT116 cells transfected with *MLH1* WT Suppl Figure 3A). Therefore, it is likely that in the case of a hypomorphic allele with attenuated (but not abrogated) DNA damage response capacity that may confer only a moderate risk of cancer, our screen could conclude to a misclassification of this variant as neutral. Our screen should be optimized to address these conditions. Another limitation of our model system is that the variant is evaluated outside the genetic background of the patient and therefore, potential additive effects of several variants, such as previously described for MMR genes<sup>39, 45, 46</sup>, could be missed.

This proof of concept study validates our approach for the functional characterization of VUS for the 2 MMR genes mainly involved in LS. With this screen, we have developed a simple, rapid, cost-effective, and reliable method. The full assay (transfection of the mutants, cell sorting and *in vivo* MMR assay) is performed in 14 to 16 days, and many mutants can be tested in parallel. Another major benefit of our screen is that patient-derived material is not required. Further validation of our assay using a much larger group of accepted neutral and pathogenic mutations from well-studied patients is required to aid in better defining a cutoff value between MMR damaging and non-damaging variants. Moreover, testing a larger panel of VUS will allow searching for potential correlations between the functional screen quantitative data and the clinical phenotype. We are confident that our validated and standardized functional test will soon have a place of choice in a comprehensive diagnostic procedure for LS-associated *MLH1* or *MSH2* VUS.

## **MATERIALS AND METHODS**

### **Selection of the variants**

The present study comprised 44 *MLH1* and 45 *MSH2* variants. *MLH1* variants include 10 known pathogenic mutations (8 non-synonymous missense mutations and 2 non-sense mutations), 9 known

neutral alterations (8 non-synonymous and 1 synonymous missense alteration) and 25 VUS that are non-synonymous missense alterations (including 3 different alterations at the R474 protein position). *MSH2* variants include 10 known pathogenic mutations (9 missense and 1 nonsense mutations), 11 known neutral alterations (9 non-synonymous and 2 synonymous missense variants) and 24 VUS. Variants already known as pathogenic or neutral were selected on the basis of (i) identical or close classification in the InSiGHT and UMD databases, and (ii) a relatively high reported frequency. VUS were selected on the basis of (i) a classification as class 3 in InSiGHT database, (ii) a relatively high reported frequency and (iii) the most clinical data available. In addition, 4 private VUS were selected. For all variants, the absence of an effect of the alteration on splicing has been checked on InSiGHT/UMD databases or in default of data, using Alamut software. We also ensure that the mutations were scattered throughout the MLH1 structure.

## **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  $\mu$ 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 in the pIRES vector, placing the expression of the two proteins under control of the strong CMV promoter. The complete cDNA for the wild type *MSH2* gene (NM\_000251.2) and the *GFP* reporter gene were subcloned in the pVAX1 vector, placing the expression of the two proteins under control of the strong CMV promoter. The expression constructs for the mutant variants were designed from the wild type vector, using site directed mutagenesis, and achieved by ProteoGenix (Schiltigheim France).

## **Cellular transfection and cell sorting**

HCT116 *MLH1*-deficient and LoVo (*MSH2*-deficient) cells were grown in a humidified atmosphere containing 5% CO<sub>2</sub> in DMEM GlutaMAX medium (ThermoFisher) supplemented with 10% fetal calf serum. Exponentially growing HCT116 cells were seeded into 6-well plates in medium supplemented with 1% Penicillin-Streptomycin and 10% fetal calf serum at a density of 4x10<sup>5</sup> (HCT116) or 5x10<sup>5</sup>

cells/well (LoVo) for 24 hours then transfected with 6  $\mu$ L TurboFect Transfection Reagent (ThermoFisher) mixed with 2  $\mu$ g plasmid DNA according to the manufacturer's instructions. Medium was removed 6 hours after the transfection and fresh medium complemented with 20  $\mu$ M O6-benzylguanine was added. Cell sorting was performed 24 hours post-transfection based on the mCherry fluorescence using a MoFlo Astrios flow cytometer (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 20% fetal calf serum and 20  $\mu$ M O6-benzylguanine, at a density of  $4 \times 10^3$  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  $\mu$ mol/L final concentration for the *MLH1* training set, and at 0.1  $\mu$ mol/L final concentration for the *MLH1* VUS. For the *MSH2* variants, final MNNG concentration was 0.1, 0.5, 1 and 2  $\mu$ mol/L for the training set, and 1  $\mu$ mol/L final concentration for the *MSH2* 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, colonies were colored in a solution containing 6% glutaraldehyde and 0.5% crystal violet, and counted. For LoVo cells, 25% methanol was added in the staining solution. Survival fraction was expressed as the number of colonies of treated sample relative to untreated control sample. All samples were tested in triplicate and each experiment was conducted at least in duplicate.

### **Western Blot analysis**

Proteins extracted from mCherry- or GFP-positive sorted cells were separated on a 4-12% SDS-PAGE gels, transferred to nitrocellulose membranes and incubated overnight with either mouse anti-MLH1 (BD Pharmingen), or mouse anti-MSH2 (Calbiochem), or rabbit anti- $\beta$  actin (LI-COR) antibody followed by goat anti-mouse or goat anti-rabbit antibody (LI-COR) and then revealed and quantified with Odyssey (LI-COR) imaging system. Band intensity was normalized with the corresponding loading control ( $\beta$  actin) and expressed as a percentage of the expression observed in HCT116 cells transfected with WT *MLH1* or LoVo cells transfected with WT *MSH2*.

### ***In silico* analysis**

All *MLH1* 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.

## FIGURES

### FIGURE 1. Functional screen for the characterization of *MLH1* or *MSH2* variants (schematic diagram)

We transiently introduced *MLH1* (or *MSH2*) variants into the *MLH1*-null (or *MSH2*-null) human colorectal cancer cell line HCT116 (or LoVo). After selection of transfected cells using cell sorting to recover only those expressing mCherry (or GFP), a reporter gene contained in the expression vector, cells were reseeded and exposed to MNNG (methylating agent). Clonogenic survival was used as readout for MMR activity: variants that confer restoration of cytotoxicity to methylation damage were classified as neutral whereas variants unable to restore apoptosis were classified as damaging. All samples were tested in triplicate and each experiment was conducted at least in duplicate.

### FIGURE 2. Complementation of MMR gene defect

After transient transfection of (A) HCT116 cells with either the *MLH1*-mCherry or the control-mCherry construct, or (B) LoVo cells with either the *MSH2*-GFP or the control-GFP construct, the transfected cells were sorted and exposed to 1, 2, or 3 rounds of MNNG treatment performed 48h, 60h and 72h post-transfection, respectively. A clonogenic assay was then carried out. While control cells (i.e. cells transfected with random construct, in red) survived mild MNNG selection, cells transfected with the WT *MLH1* (or *MSH2*) containing construct (in green) reversed the methylation tolerance phenotype, which indicates a restoration of the DNA damage sensing function associated with MMR. Response of the non-transfected parental HCT116 (or LoVo) cells (in blue) is shown for comparison. Data are means  $\pm$  SD from at least 2 clonogenic experiments (each in triplicate).

### FIGURE 3. Validation study

The functional screen of a series of proven pathogenic and neutral *MLH1* (A) or *MSH2* (B) 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 WT *MLH1* (or WT *MSH2*) or the control constructs are represented by triangles. All samples were tested in triplicate and each experiment was conducted at least twice independently.

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

**A.** 44 *MLH1* variants including 10 proven pathogenic and 9 proven neutral variants (validation set) and 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  $\mu$ M MNNG.

**B.** 45 *MSH2* variants including 10 proven pathogenic and 11 proven neutral variants (validation set) and 24 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  $\mu$ M MNNG.

Screen results with the WT or the control constructs are represented by triangles. The area between the response of pathogenic and neutral variants found in the validation set is in gray. The cut-off value as determined by the ROC curve in the validation set is represented in dotted line.

**FIGURE 5.** Correlation between *MLH1* (A.) or *MSH2* (B.) protein expression assessed by Western blot in transfected sorted cells and methylation tolerance screen results

Protein expression of *MLH1* and *MSH2* is expressed relative to the expression observed in HCT116 cells transfected with WT *MLH1* or LoVo cells transfected with WT *MSH2*, respectively.

Red dot: known pathogenic mutation; green dot: known neutral mutation; red square: damaging VUS; green square: neutral VUS; orange square: potentially damaging VUS; light green square: potentially neutral VUS

**FIGURE 6.** Decision tree for the functional characterization of VUS of *MLH1* and *MSH2* genes in suspected LS patients

Since a single functional assay may not be definitive, an optimal characterization may require integration of data from 2 functional assays each evaluating a different MMR function, i. e. DNA damage signalization or mismatch repair. Because variants that induce aberrant splicing, exon skipping or mRNA stability will be overlooked in the setting of most functional assays, a comprehensive evaluation of splicing through a combination of *in silico* prediction algorithms and RT-PCR analysis of RNA should be included.

**SUPPLEMENTARY FIGURE 1. A.** Results of the clonogenic assay showing decrease of survival colonies after MNNG treatment in HCT116 cells transfected with the *MLH1* containing construct contrary to

HCT116 cells transfected with the random sequence containing construct. Each column of three wells constitutes a triplicate. Colonies are stained with crystal violet.

**B.** Kinetics of expression of the *MLH1* transgene that allowed to control that the rate of transfected cells remained stable up to 48h- and 72h-post transfection, i.e. when cells are exposed to MNNG treatments. Data are means  $\pm$  SD from two independent experiments.

**SUPPLEMENTARY FIGURE 2.** Protein localization of all *MLH1* (A) or *MSH2* (B) variants screened in our study.

**SUPPLEMENTARY FIGURE 3.** Protein expression assessed by Western blots on transfected sorted cells for all *MLH1* (A) and *MSH2* (B) variants

**A.** *MLH1* variants comprised 10 known pathogenic variants (including 2 nonsense mutations marked with \*), 9 known neutral variants and 25 VUS

**B.** *MSH2* variants comprised 10 known pathogenic variants (including 1 nonsense mutation marked with \*), 11 known neutral variants and 24 VUS

Protein expression of *MLH1* and *MSH2* is expressed relative to the expression observed in HCT116 cells transfected with WT *MLH1* or LoVo cells transfected with WT *MSH2*, respectively.

**SUPPLEMENTARY FIGURE 4.** Comparison of the functional screen results with *in silico* predictions for the 49 VUS from the prospective set, which includes 25 *MLH1* (A) and 24 *MSH2* (B) VUS

**SUPPLEMENTARY FIGURE 5.** Comparison of the functional screen results with *in silico* predictions for 40 variants from the validation set, which includes 19 *MLH1* (A) and 21 *MSH2* (B) variants known to be pathogenic or neutral

**SUPPLEMENTARY TABLE 1.** List of the proven pathogenic and proven neutral *MLH1* and *MSH2* variants of the validation set

**SUPPLEMENTARY TABLE 2.** Clinical data from the literature and MMR databases for the 25 *MLH1* VUS of the prospective set.

**SUPPLEMENTARY TABLE 3.** Clinical data from the literature and MMR databases for the 24 *MSH2* VUS of the prospective set.

## BIBLIOGRAPHY

1. Lynch HT, Snyder CL, Shaw TG, et al. Milestones of Lynch syndrome: 1895-2015. *Nat Rev Cancer* 2015;15:181-94.
2. 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.
3. Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. *Science* 1993;260:812-6.
4. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science* 1993;260:816-9.
5. 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.
6. Peltomaki P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* 2003;21:1174-9.
7. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med* 2003;348:919-32.
8. Peltomaki P. Update on Lynch syndrome genomics. *Fam Cancer* 2016;15:385-93.
9. 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.
10. 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.
11. Ou J, Niessen RC, Lutzen A, et al. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. *Hum Mutat* 2007;28:1047-54.
12. Rasmussen LJ, Heinen CD, Royer-Pokora B, et al. Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. *Hum Mutat* 2012;33:1617-25.
13. Heinen CD, Juel Rasmussen L. Determining the functional significance of mismatch repair gene missense variants using biochemical and cellular assays. *Hered Cancer Clin Pract* 2012;10:9.
14. Belvederesi L, Bianchi F, Galizia E, et al. MSH2 missense mutations and HNPCC syndrome: pathogenicity assessment in a human expression system. *Hum Mutat* 2008;29:E296-309.
15. Ollila S, Sarantaus L, Kariola R, et al. Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. *Gastroenterology* 2006;131:1408-17.
16. Mastrocola AS, Heinen CD. Lynch syndrome-associated mutations in MSH2 alter DNA repair and checkpoint response functions in vivo. *Hum Mutat* 2010;31:E1699-708.
17. 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.
18. Karran P, Stephenson C. Mismatch binding proteins and tolerance to alkylating agents in human cells. *Mutat Res* 1990;236:269-75.
19. Karran P, Bignami M. Self-destruction and tolerance in resistance of mammalian cells to alkylation damage. *Nucleic Acids Res* 1992;20:2933-40.

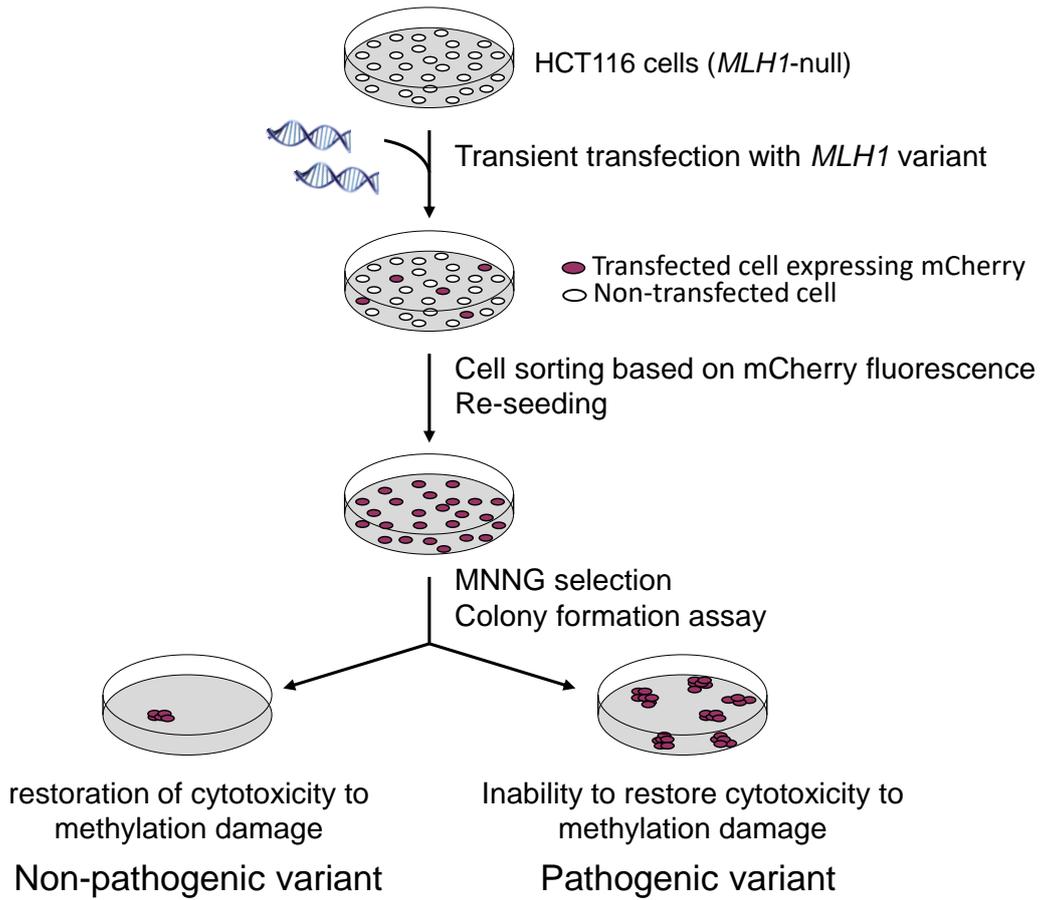
20. Hawn MT, Umar A, Carethers JM, et al. Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. *Cancer Res* 1995;55:3721-5.
21. Kat A, Thilly WG, Fang WH, et al. An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair. *Proc Natl Acad Sci U S A* 1993;90:6424-8.
22. 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.
23. 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.
24. Wang Q, Lasset C, Desseigne F, et al. Neurofibromatosis and early onset of cancers in hMLH1-deficient children. *Cancer Res* 1999;59:294-7.
25. 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.
26. Taebner 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.
27. Papadopoulos N, Nicolaides NC, Wei YF, et al. Mutation of a mutL homolog in hereditary colon cancer. *Science* 1994;263:1625-9.
28. Branch P, Hampson R, Karran P. DNA mismatch binding defects, DNA damage tolerance, and mutator phenotypes in human colorectal carcinoma cell lines. *Cancer Res* 1995;55:2304-9.
29. 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.
30. Hinrichsen I, Brieger A, Trojan J, et al. Expression defect size among unclassified MLH1 variants determines pathogenicity in Lynch syndrome diagnosis. *Clin Cancer Res* 2013;19:2432-41.
31. Offman J, Opelz G, Doehler B, et al. Defective DNA mismatch repair in acute myeloid leukemia/myelodysplastic syndrome after organ transplantation. *Blood* 2004;104:822-8.
32. Duval A, Raphael M, Brennetot C, et al. The mutator pathway is a feature of immunodeficiency-related lymphomas. *Proc Natl Acad Sci U S A* 2004;101:5002-7.
33. Borie C, Colas C, Dartigues P, et al. The mechanisms underlying MMR deficiency in immunodeficiency-related non-Hodgkin lymphomas are different from those in other sporadic microsatellite instable neoplasms. *Int J Cancer* 2009;125:2360-6.
34. Bodo S, Svrcek M, Sourrouille I, et al. Azathioprine induction of tumors with microsatellite instability: risk evaluation using a mouse model. *Oncotarget* 2015;6:24969-77.
35. Fahrer J, Kaina B. Impact of DNA repair on the dose-response of colorectal cancer formation induced by dietary carcinogens. *Food Chem Toxicol* 2017;106:583-594.
36. 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.
37. Liu HX, Zhou XL, Liu T, et al. The role of hMLH3 in familial colorectal cancer. *Cancer Res* 2003;63:1894-9.
38. Gammie AE, Erdeniz N, Beaver J, et al. Functional characterization of pathogenic human MSH2 missense mutations in *Saccharomyces cerevisiae*. *Genetics* 2007;177:707-21.

39. 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.
40. Drost M, Zonneveld JB, van Hees S, et al. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. *Hum Mutat* 2012;33:488-94.
41. Sia EA, Dominska M, Stefanovic L, et al. Isolation and characterization of point mutations in mismatch repair genes that destabilize microsatellites in yeast. *Mol Cell Biol* 2001;21:8157-67.
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. Kantelinen J, Kansikas M, Candelin S, et al. Mismatch repair analysis of inherited MSH2 and/or MSH6 variation pairs found in cancer patients. *Hum Mutat* 2012;33:1294-301.
46. Hinrichsen I, Schafer D, Langer D, et al. Functional testing strategy for coding genetic variants of unclear significance in MLH1 in Lynch syndrome diagnosis. *Carcinogenesis* 2015;36:202-11.

**TABLE 1 .** Results from methylation tolerance assay and protein expression performed with *MLH1* and *MSH2* VUS

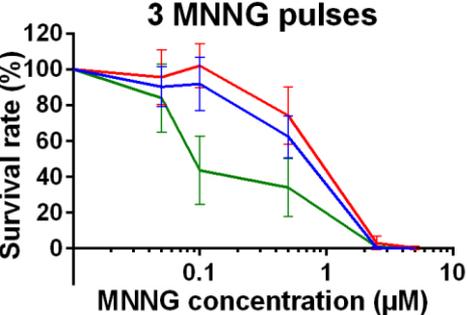
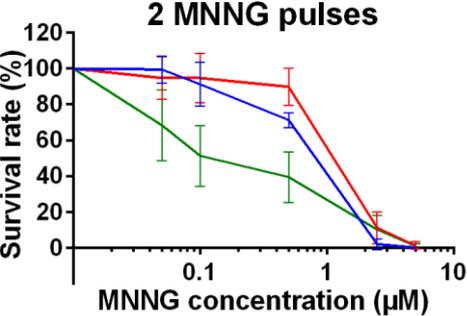
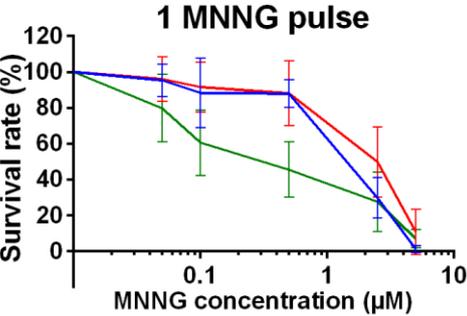
Variant	Mean survival score (%)	Protein expression (%)	Classification
<b><i>MLH1</i> VUS</b>			
L85P	77,12	114,4	Potentially damaging
S106R	84,96	103,99	Damaging
T116K	44,83	67,38	Neutral
Y126N	51,99	50,38	Neutral
A160V	27,69	142,09	Neutral
V179D	61,88	63,43	Potentially damaging
N215S	46,73	104,95	Neutral
G244V	101,24	35,58	Damaging
L260F	76,37	43,31	Potentially damaging
H264Y	63,65	32,24	Potentially neutral
E319K	32,28	84,01	Neutral
N338S	44,57	46,84	Neutral
G454R	38,42	51	Neutral
R474Q	35,77	74,43	Neutral
R474W	29,17	78,62	Neutral
R474G	28,28	101,33	Neutral
E523D	49,04	69,03	Neutral
A539D	98,51	18,29	Damaging
I565T	45,13	43,79	Neutral
L582F	75,67	23,5	Potentially damaging
A586D	87,77	37,66	Damaging
K618E	50	44,91	Competent
L622P	75,68	22,99	Potentially neutral
P640L	102,05	25,19	Damaging
W666R	88,21	39,54	Damaging
<b><i>MSH2</i> VUS</b>			
A2T	22,63	205,65	Neutral
T33P	21,11	81,75	Neutral
Y43C	15,53	98,41	Neutral
M152T	10,26	145,18	Neutral
R243Q	23,56	47,93	Neutral
A305T	30,54	110,85	Neutral
V342I	24,08	127,08	Neutral
P349A	31,54	135,87	Neutral
V470E	79,63	162,96	Damaging
A500P	66,01	139,42	Damaging
G548C	29,82	124,21	Neutral
E561K	8,70	124,93	Neutral
N583S	12,65	125,21	Neutral
L602P	78,64	154,89	Damaging
H610N	18,20	130,05	Neutral
P622T	71,03	92,05	Damaging
K627N	20,71	141,69	Neutral
R638G	59,66	108,43	Damaging
N671Y	67,36	82,90	Damaging
R711Q	53,08	90,61	Potentially damaging
S723F	21,71	243,90	Neutral
L736P	55,58	111,22	Potentially damaging
Y856C	21,46	61,11	Neutral
T905I	20,99	152,84	Neutral

**FIGURE 1**

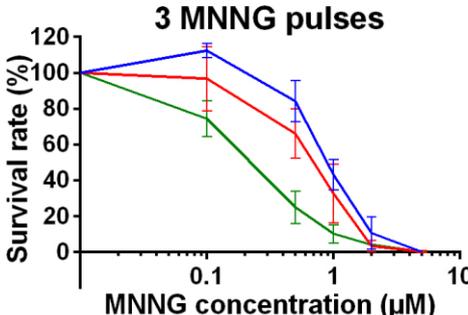
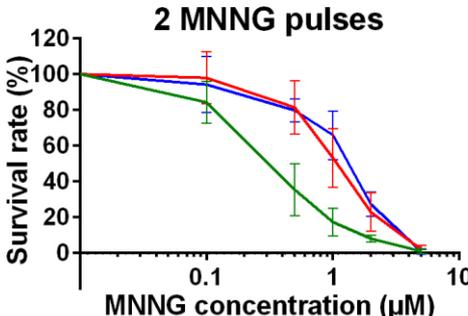
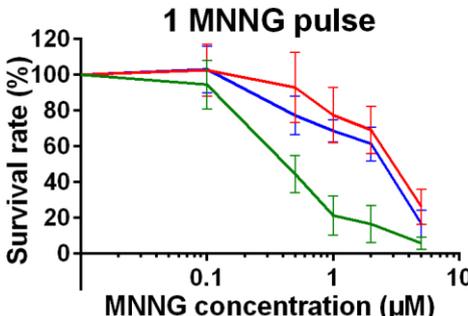


# FIGURE 2

**A**

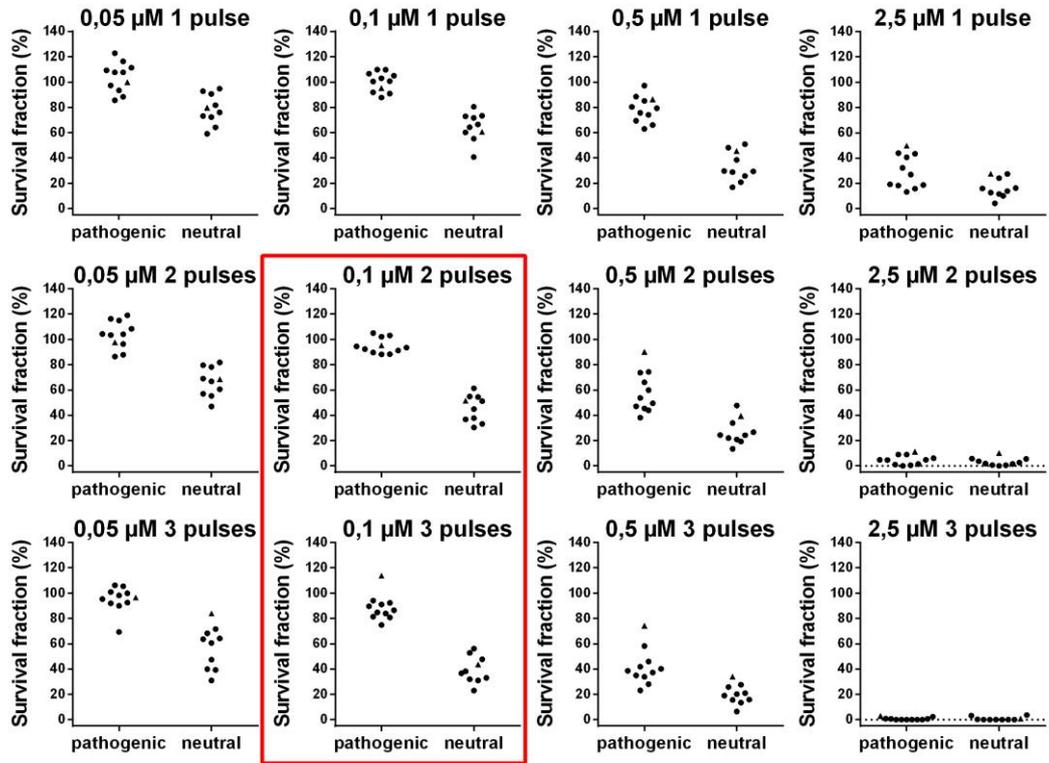


**B**



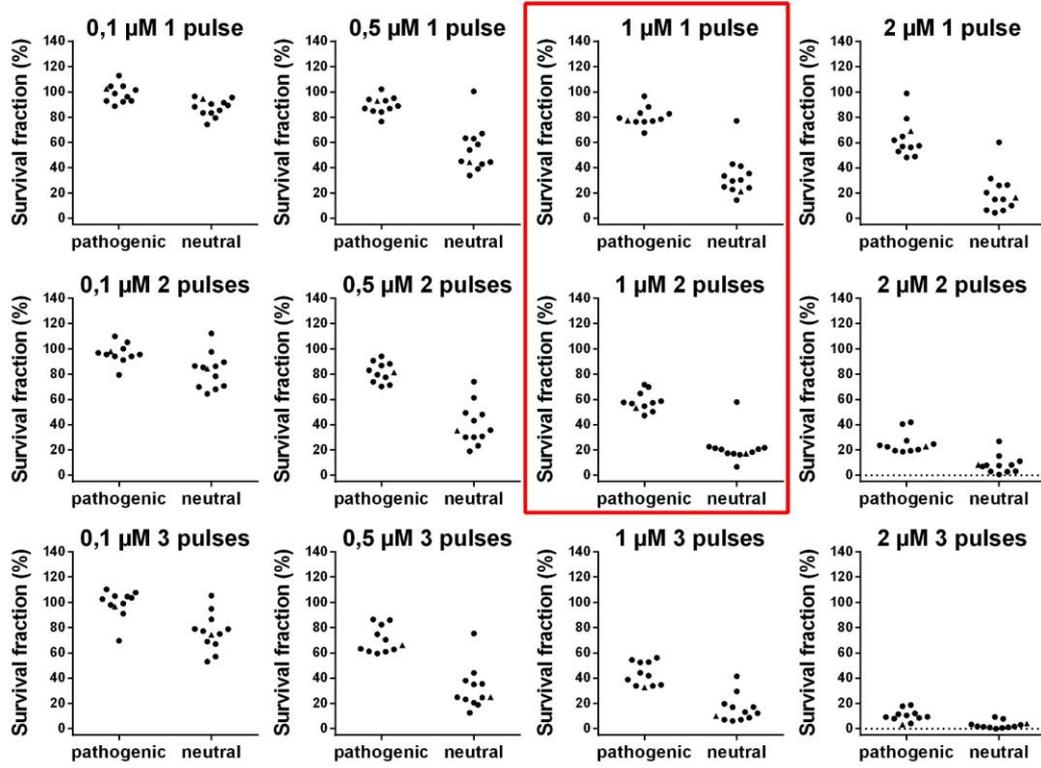
**FIGURE 3. A**

**A. *MLH1***



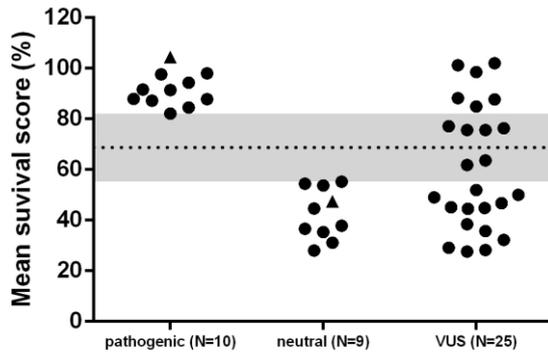
**FIGURE 3.B**

**B. MSH2**

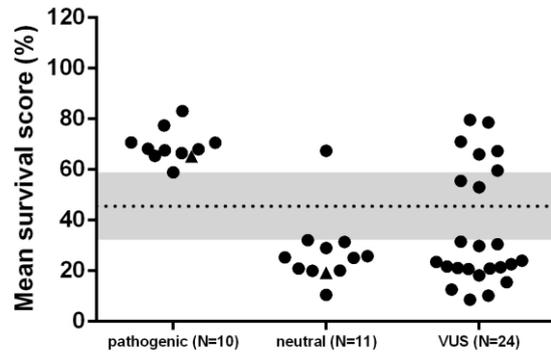


# FIGURE 4

## A. *MLH1*

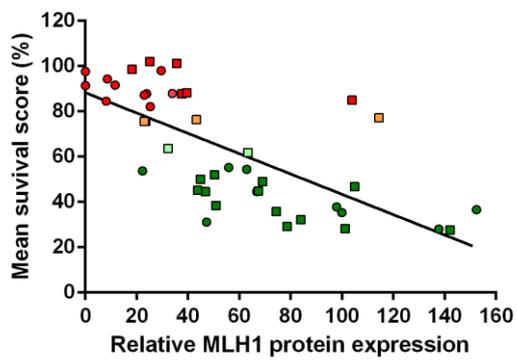


## B. *MSH2*

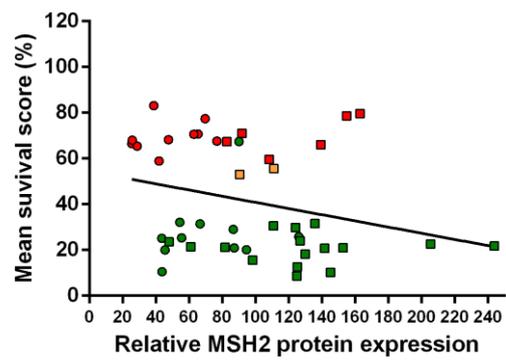


# FIGURE 5

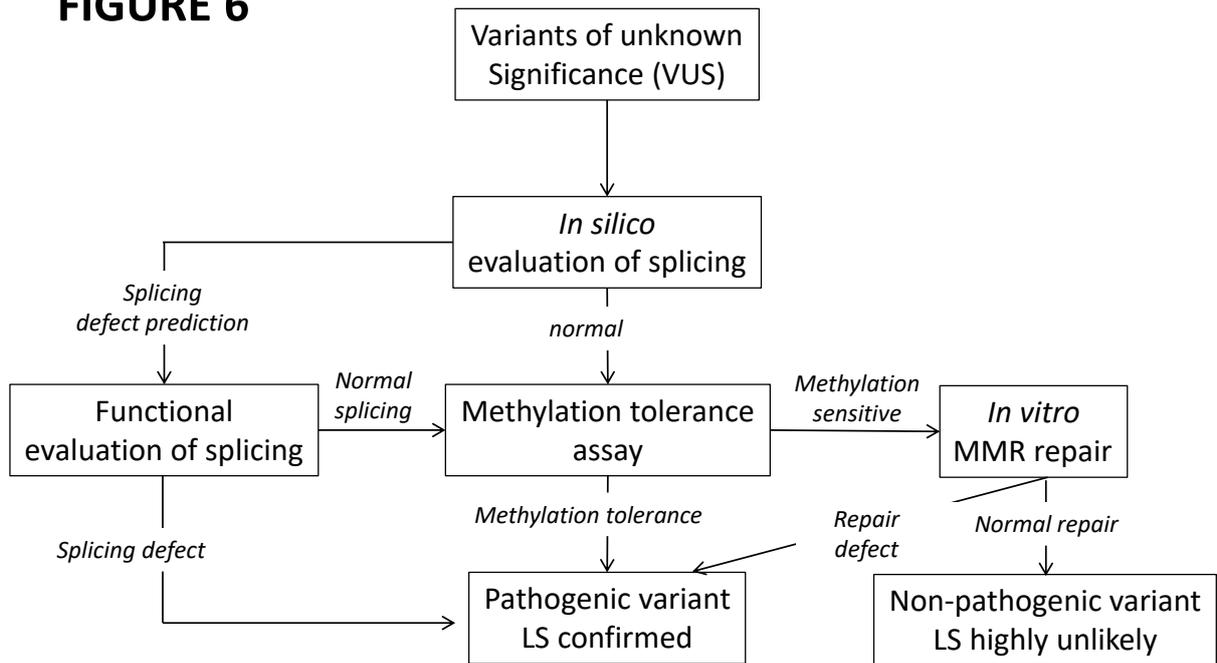
## A. *MLH1*



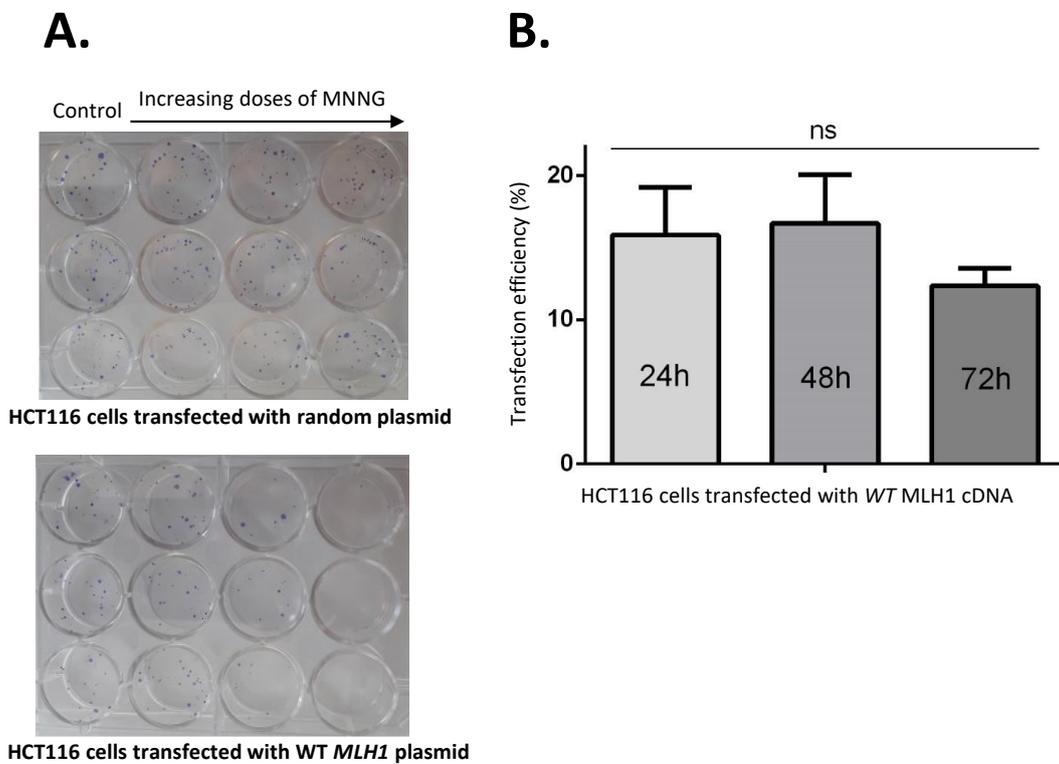
## B. *MSH2*



**FIGURE 6**

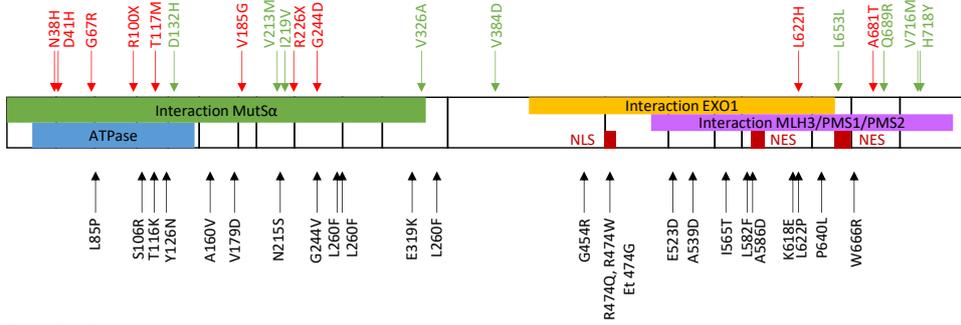


**SUPPL FIGURE 1**



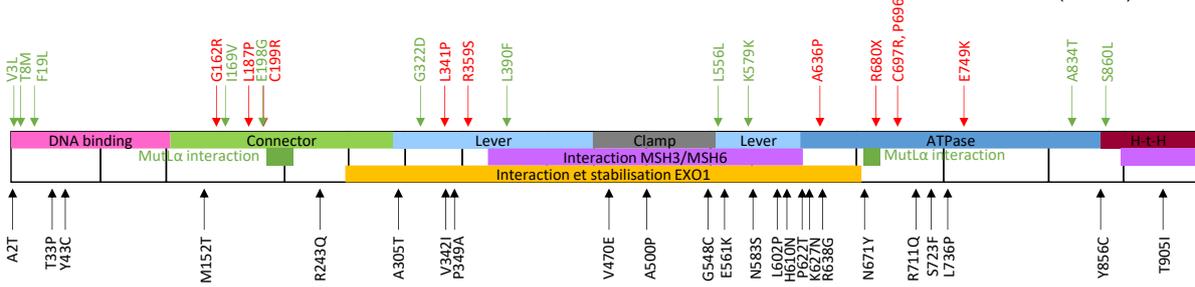
# SUPPL FIGURE 2

## A. MLH1



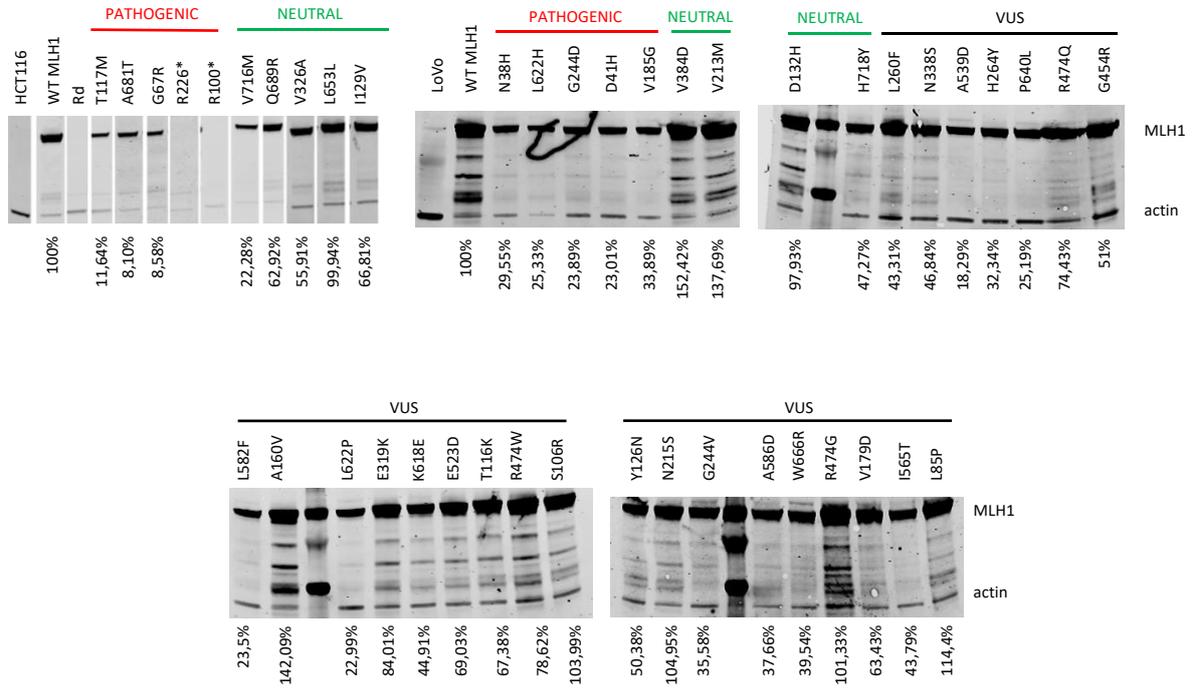
**MLH1 :**  
 neutral (N = 9)  
 pathogenic (N = 10)  
 VUS (N = 25)

## B. MSH2

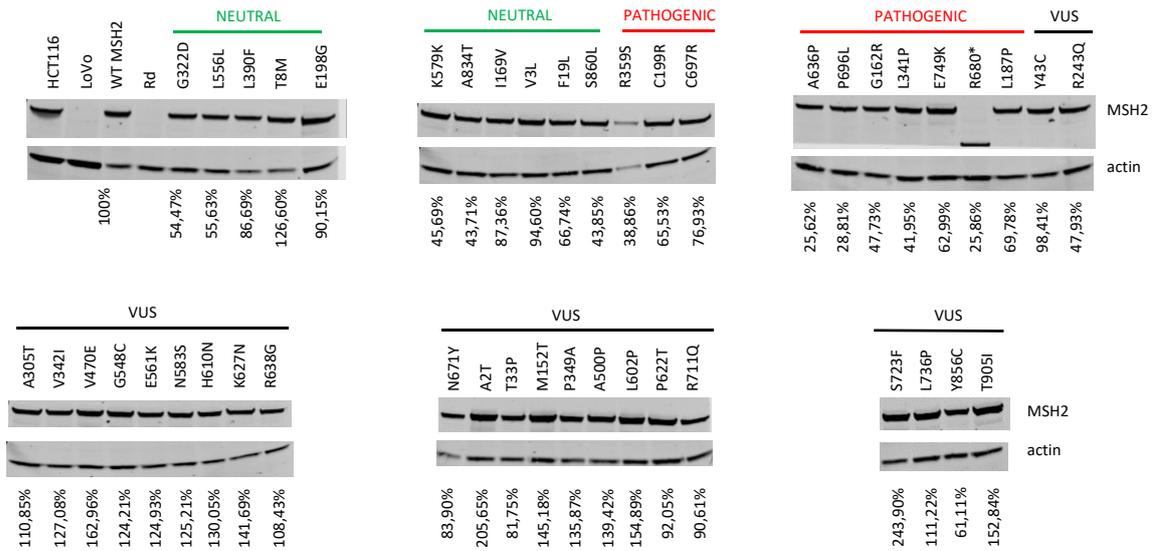


**MSH2 :**  
 neutral (N = 11)  
 pathogenic (N = 10)  
 VUS (N = 24)

# SUPPL FIGURE 3.A

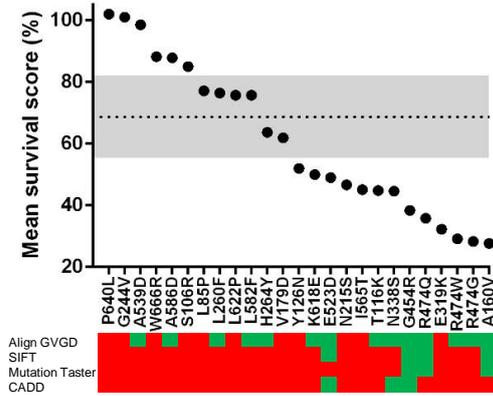


# SUPPL FIGURE 3.B

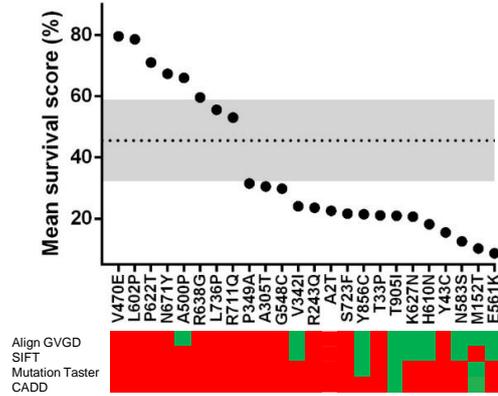


# SUPPL FIGURE 4

## A. *MLH1* VUS

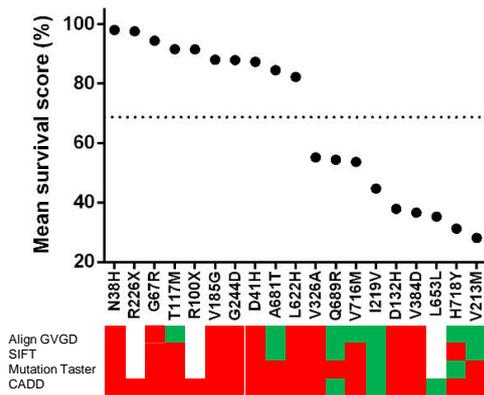


## B. *MSH2* VUS

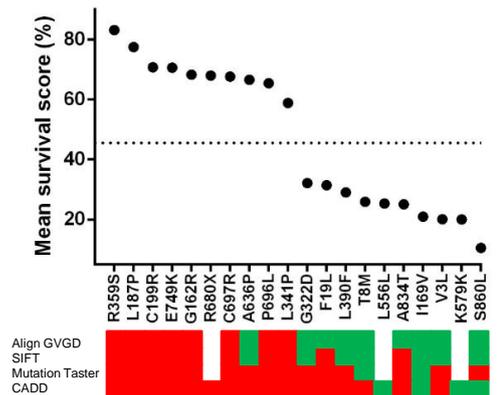


# SUPPL FIGURE 5

## A. *MLH1*



## B. *MSH2*



**Table S1. Proven pathogenic and non pathogenic *MLH1* and *MSH2* variants used in the validation study**

Variant		Classification	
Protein	DNA	InSiGHT	UMD
<b><i>MLH1</i> variants</b>			
N38H	112A>C	5	
D41H	121G>C	5	4
G67R	199G>A	5	5
R100X	298C>T	5	5
T117M	350C>T	5	5
V185G	554T>G	5	
R226X	676C>T	5	5
G244D	731G>A	5	5
L622H	1865T>A	5	
A681T	2041G>A	5	5
D132H	394G>C	1	1
V213M	637G>A	1	1
I219V	655A>G	1	1
V326A	977T>C	1	1
V384D	1151T>A	1	1
L653L	1959G>T	1	1
Q689R	2066A>G	1	1
V716M	2146G>A	1	1
H718Y	2152C>T	1	1
<b><i>MSH2</i> variants</b>			
G162R	484G>A	5	4
L187P	560T>C	5	
C199R	595T>C	5	5
L341P	1022T>C	4	5
R359S	1077A>T	5	5
A636P	1906G>C	5	5
R680X	2038C>T	5	5
P696L	2087C>T	5	4
C697R	2089T>C	5	5
E749K	2245G>A	4	5
V3L	7G>T		2
T8M	23C>T	1	2
F19L	55T>C		2
I169V	505A>G	2	
E198G	593A>G	1	
G322D	965G>A	1	1
L390F	1168C>T	1	1
L556L	1666T>C	1	1
K579K	1737A>G	1	
A834T	2500G>A	2	
S860L	2579C>T	2	

**Table S2. Clinical data available in InSiGHT/UMD databases for the 25 *MLH1* VUS that were selected for screening**

Variant		Classification		Tumor and age (years) <sup>b</sup>	MSI	IHC and/or other somatic data <sup>c</sup>	Amsterdam or Bethesda criteria	Cancer and age (years) in relatives <sup>b</sup>	Allele frequency (%)	Functional tests <sup>d</sup>	Screen result <sup>e</sup>
Protein	DNA	InSiGHT	UMD <sup>a</sup>								
L85P	254T>C	/	3*	CRC 49	YES	MLH1+ PMS2-	/	/	/		77.12% Potentially damaging
S106R	318C>A	3	3	/	YES	MLH1+ MSH2+ MSH6+ PMS2+	/	/	/		84.96% Damaging
T116K	347C>A	3	3	CRC 28	/	/	/	mat uncle: CRC 58	/		44.83% Neutral
Y126N	376T>A	3	3	LGD Ad 30 BC 73	/	/	/	non carrier sister: CRC 74; several BC in family	0.0074		51.99% Neutral
A160V	479C>T	3	/	CRC 68	YES	MLH1-	Amsterdam I	/	0.014	102% repair activity in vitro; dominant mutator effect; expression >75% in transfected HCT116 cells <sup>1</sup>	27.69% Neutral
				BC	/	/	/	No CRC			
V179D	536T>A	/	3*	CRC 59	/	/	/	Mother: CRC 65; mat aunt: GC 61; father: PrC PC 79; pat uncle: HCC; pat uncle's son: MSTO (asbestos)	/		61.88% Potentially neutral
N215S	644A>G	3	3	Spo CRC <50	YES	/	/	/	0.0016		46.73% Neutral
G244V	731G>T	3	3	/	/	MLH1-	Bethesda	/	/	24% repair activity in vitro; No dominant mutator effect; expression	101.24% Damaging
				Spo CRC	YES	/	/	/			
				/	YES	MLH1+ MSH2+ MSH6+ PMS2-	/	/			

					YES	MLH1-	Asmterdam I	Non-carrier: CRC 60; asymptomatic carrier 80		>75% in transfected HCT116 cells <sup>1</sup>	
					YES	MLH1+ MSH2+ MSH6+ PMS2-		Carrier paternal uncle; paternal aunt asymptomatic carrier 65			
L260F	778C>T	3	3	CRC 41	/	MLH1+ PMS2+	Amsterdam II	Tested individuals: CCR 36; PrC 68; Untested individuals: CCR 34; CCR 45; CCR 55; CCR 45; EC 49; GC 58	/	/	76.37%  Potentially damaging
				/	YES	MLH1-	/	/			
				Spo BC	YES	MLH1+ MSH2+  No <i>MLH1</i> methylation	/	/			
H264Y	790C>T	3	3	CRC 30	YES	/	/	/	0.00087	/	63.65%  Potentially neutral
E319K	955G>A	3	3	BC 45	/	MLH1+ MSH2+ MSH6+ PMS2+		Mother: BC 40; mat aunt: BC 40; mat aunt's daughter: BC 40	0.0082		32.28%  Neutral
N338S	1013A>G	3	3	CRC 38	/	/	Bethesda	Father: lymphoma ; uncle: CRC	0.0058	Repair activity in vitro and expression in transfected human cells similar to WT <sup>2</sup>	44.57%  Neutral
				CRC 33	YES	MLH1- PMS2-	Bethesda	Untested individuals: skin; CCR 80; PrC 80; skin; CCR 50			
G454R	1360G>C	3	1	/	NO	/	/	/	0.015	/	38.42%  Neutral
				/	/	MLH1+ MSH2-	/	/			
				Ad	/	/	Bethesda	One unaffected carrier; Untested individuals: CRC 39; EC 42			

R474Q	1421G>A	3	1	/	/	/	/	/	0.00082	102% repair activity in vitro ; dominant mutator effect; expression >75% in transfected HCT116 cells <sup>1</sup>	35.77% Neutral
R474W	1420C>T	3	/	/	/	/	/	/	0.0016	/	29.17% Neutral
R474G	1420C>G	/	3*	/	/	/	/	/	0.0082	/	28.28% Neutral
E523D	1569G>T	3	/	/	/	/	/	/	/	96% repair activity in vitro; dominant mutator effect; expression >75% in transfected HCT116 cells <sup>1</sup>	49.04% Neutral
A539D	1616C>A	3	3	CRC <50	/	/	/	2 affected carriers in the family	/	/	98.51% Damaging
I565T	1694T>C	/	3*	/	/	/	/	/	/	/	45.13% Neutral
L582F	1744C>T	3	3	CRC 37	YES	MLH1- MSH2+ MSH6+ PMS2- No <i>MLH1</i> methylation	Bethesda	/	/	30% repair activity in vitro <sup>3</sup> ; nuclear localization in murine NIH-3T3; defective PMS2 interaction <sup>4</sup>	75.67% Potentially damaging
				CRC 52	YES	MLH1-	/	/			
A586D	1757C>A	3	3	/	/	/	Amsterdam II	/	/	/	87.77% Damaging
K618E	1852A>G	3	3				Amsterdam II		0.34	/	50%
				CRC, extracolonic LS and non LS tumor	/	/	No Amsterdam	Father: extracolonic HNPCC tumor; aunt and cousin: CCR			Neutral

				CRC 77	YES	MLH1- MSH2+ MSH6+ PMS2+ <i>MLH1</i> methylation (region H not D)	No Bethesda no Amsterdam	no			
L622P	1865T>C	3	/	BC 38, CRC 48	YES	MLH1- PMS2-	Bethesda II	Untested individuals: GC 70; bladder; unknown cancer	/	75.68% Potentially damaging	
				CRC 43	YES	MLH1- PMS2-	Amsterdam II	Untested individuals with CRC: 31; 49; 69; 78; unknown			
				CRC 48, 52, melanoma 54, GC 63	YES	MLH1- PMS2-	/	Brother: leukemia 51; sister: melanoma 40; pat uncle: CRC <60; pat grand-father: CRC <49			
				CRC 37	YES	MLH1- PMS2- No <i>MLH1</i> methylation	/	Mother: CRC 45			
P640L	1919C>T	3	5		YES	MLH1- PMS2-	Amsterdam I	/	/	No dominant mutator effect, no interaction with PSM2, reduced <i>MLH1</i> expression (in human transfected cells and in yeast cells) <sup>5</sup>	102.05% Damaging
				EC 36, DC 48, CRC 56, UC 64	/	/	Amsterdam I	Father: CRC 38; pat uncle and grand-father: CRC			
				CRC 31	/	MLH1- PMS2-	Bethesda	Untested individuals: CRC 29; GC 55; CRC 81			
				CRC 48, 55	YES	MLH1-	Bethesda	Untested relatives: EC; CRC; unknown			
				/	YES	/	Amsterdam I	1 affected carrier relative			
W666R	1996T>C	3	3	/	/	/	/	/	/	88.21% Damaging	

<sup>a</sup> An asterisk indicates a private variant

<sup>b</sup> Ad: adenoma; BC: breast cancer; CRC: colorectal cancer; EC: endometrial cancer; GC: gastric cancer; HCC: hepatocellular carcinoma; LGD : low grade dysplasia; mat: maternal; LS: Lynch syndrome; MSTO: mesothelioma; pat: paternal; PC: pancreatic cancer; PrC: prostate cancer; Spo: sporadic

<sup>c</sup> IHC: immunohistochemistry; + indicates a staining of the corresponding protein; - indicates an absence of of the corresponding protein

<sup>d</sup> WT: wild type

<sup>e</sup> expressed as the mean survival score and the proposed interpretation

1. Takahashi M, Shimodaira H, Andreutti-Zaugg C, et al. Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. *Cancer Res* 2007;67:4595-604.
2. Koger N, Paulsen L, Lopez-Kostner F, et al. Evaluation of MLH1 variants of unclear significance. *Genes Chromosomes Cancer* 2018;57:350-358.
3. Drost M, Zonneveld J, van Dijk L, et al. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. *Hum Mutat* 2010;31:247-53.
4. 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.
5. Hardt K, Heick SB, Betz B, et al. Missense variants in hMLH1 identified in patients from the German HNPCC consortium and functional studies. *Fam Cancer* 2011;10:273-84.

**Table S3. Clinical data available in InSiGHT/UMD databases for 24 *MSH2* VUS that were selected for screening**

Variant		Classification		Tumor and age (years) <sup>b</sup>	MSI	IHC and/or other somatic data <sup>c</sup>	Amsterdam or Bethesda criteria	Cancer and age (years) in relatives <sup>b</sup>	Allele frequency (%)	Functional tests <sup>d</sup>	Screen result <sup>e</sup>
Protein	DNA	InSiGHT	UMD <sup>a</sup>								
A2T	4G>A	3	1	LS spectrum cancer < 50	/	/	Amsterdam I		0.039	/	22.63% Neutral
				/	/	MSH2-	Bethesda	6 families ; 1 family: pathogenic <i>MSH2</i> mutation c.1835C>G <i>in cis</i>			
				CRC 39	YES	MSH2-	Bethesda	Brother: CRC 46, sister: EC 47; sister: CRC 68; father: liver cancer 52; uncle: CRC 30; niece: CRC 45; niece: EC 50; nephew: CRC 37; cousin: CRC 60; cousin: GC 31; great-cousin: CRC, brain tumor 18; cousin: CRC 42, 47; segregation with the disease in the family (at least over 10 meiosis)			
				Cancer 28	/	MSH2-, <i>MSH2</i> methylation, germline <i>EPCAM</i> mutation, WT <i>KRAS</i> , WT <i>BRAF</i>	/	/			
				Cancer 45	/	MSH2-, <i>MSH2</i> methylation, germline <i>EPCAM</i> mutation, WT	/	/			

						<i>KRAS, WT BRAF</i>					
				CRC 50	YES	MSH2-	/	/			
				/	YES	MSH2-	/	/			
				/	YES	MSH2-, <i>EPCAM</i> deletion (last exon <i>in cis</i> )	/	3 families			
				/	NO	MSH2+	/	2 families			
T33P	97A>C	3	3	EC 45	YES	MLH1+ MSH2+ MSH6+	Bethesda	/	0.0064	52% repair activity in vitro <sup>1</sup> , proficient mismatch binding and reduced ATP-dependent release <sup>2</sup> , intact MSH6 interaction <sup>1</sup> , MSH2 expression similar to WT <sup>1,2</sup>	21.11% Neutral
Y43C	128A>G	3	3	Synchronous CRC 50	NO	MLH1+ MSH2+ MSH6-	/	father: CRC 50; sister: CRC 50; sister: CCR 52; daughter: EC 39	0.0028	/	15.53% Neutral
				CRC 49	/	/	/	mother: CRC 73			
				GC 65	/	/	/	sister: ovary 65; sister: EC, sarcoma 64; father and mother: brain tumor			
M152T	455T>C	/	3	Spo EC < 50	NO	MLH1+ MSH2+ MSH6+ PMS2+	/	/		/	10.26% Neutral
R243Q	728G>A	3	3	Spo CRC <50	NO	MSH2-MSH6-	/	/	0.0017	/	23.56% Neutral
A305T	913G>A	3	3	CRC < 50	NO	/	Amsterdam I	/	0.005	53% repair activity in vitro <sup>3</sup> ; proficient mismatch binding & ATP-dependent release; 2.2-fold	30.54% Neutral
				/	NO	MSH2+	/	/			
				CRC 47	YES	MLH1- PMS2-	/	brother: GC 41			

										induction of mutation rate in bacteria; nuclear localization in MEF; intact EXO1 interaction; expression similar to WT <sup>4</sup>	
V342I	1024G>A	3	3	CRC < 40	YES	MLH1+ MSH2-	Amsterdam I	/	/		24.08%
								3 recurrent CRC families			Neutral
P349A	1045C>G	3	3	Spo renal cell carcinoma 63	/	/	/	/	0.009	/	31.54% Neutral
				CRC 37	YES	/	/	/			
				/	/	MSH2+ MSH6+/-	Amsterdam II	daughter (not tested) : MSH2- MSH6-			
				/	/	/	Amsterdam I	/			
				BC 38, CRC	/	/	/	/			
				/	YES	MLH1+ MSH2- MSH6- PMS2+	/	/			
V470E	1409T>A	3	3	CRC 40	YES	MLH1+ MSH2+ MSH6+ PMS2+	/	pat grand-mother: CRC 50; pat grand-father: CRC 57; 2 2nd degree relatives LS tumors < 50	/	/	79.63% Damaging
						MLH1+ MSH2+	Amsterdam II				
				EC 46, CRC 58	YES	MLH1+ MSH2+ MSH6+ PMS2+	No Amsterdam	Variant carrier brother: CRC 40, 51, GC 51; sister: CRC 38, EC; sister: CRC 30, ovary 32; pat grand-father: CRC 55; half-sister (same father): BC 63, ad; half-sister's son: ad			

A500P	1498G>C	/	3	UC 68, CRC 74	YES	MLH1+ MSH2+ MSH6+ PMS2+ No <i>MLH1</i> methylation WT <i>BRAF</i>	No Amsterdam	no	/	/	66.01% Damaging
G548C	1642G>T	3	3	/	/	/	/	/	/	/	29.82% Neutral
E561K	1681G>A	3	3	/	/	/	/	/	0.0017	/	8.70% Neutral
N583S	1748A>G	3	3	CRC 35	/	/	/	/	0.01	/	12.65%
				CRC <50	/	/	Amsterdam I	/			Neutral
L602P	1805T>C	/	3	CRC 38	YES	/	Amsterdam	brother: CRC 40; father: UC 57, 58, 60, lymphoma	/	/	78.64% Damaging
H610N	1828C>A	3	3	/	/	/	/	/	/	/	18.20% Neutral
P622T	1864C>A	3	/	CRC 23	YES	/	Amsterdam I		/	/	71.03% Damaging
				Ad 40, CRC 51, HGD Ad 52	YES	MSH2- MSH6-	No Amsterdam	father: CRC 55; mother: EC 45, CRC 55			
K627N	1881A>C	3	3	/	/	/	/	/	/	/	20.71% Neutral
R638G	1912A>G	3	3	CRC 44	YES	MSH2-	/	/	/	/	59.66% Damaging
N671Y	2011A>T	3	3	CRC 42	/	/	Bethesda	/	/	/	67.36% Damaging
R711Q	2132G>A	/	3	/	YES	MLH1+ MSH2+	Amsterdam	/	0.00082	/	53.08% Potentially damaging

S723F	2168C>T	3	5	CRC 49	YES	MSH2-	Bethesda	/	/	No repair activity in different assays <sup>3,5,6</sup> ; complete loss of MMR function in yeast <sup>6</sup> ; intact Msh3 & Msh6 interaction in yeast <sup>5</sup> ; 59% expression in yeast <sup>5</sup>	21.71% Neutral
				/	/	/	/	Danish family			
				/	YES	/	/	/			
L736P	2207T>C	/	3*	Sebaceoma 66	YES	MSH6-	No Amsterdam	No family history	/	/	55.58% Potentially damaging
Y856C	2567A>G	3	3	/	/	/	Amsterdam	/	0.00082	/	21.46% Neutral
T905I	2714C>T	3	3	/	/	/	/	multiple late-onset LS cancers in the family	0.0076	/	20.99% Neutral

<sup>a</sup> An asterisk indicates a private variant

<sup>b</sup> Ad: adenoma; BC: breast cancer; CRC: colorectal cancer; GC: gastric cancer; EC: endometrial cancer; HGD: high grade dysplasia; Spo: sporadic

<sup>c</sup> IHC: immunohistochemistry; WT: wild type; + indicates a staining of the corresponding protein; - indicates an absence of of the corresponding protein

<sup>d</sup> WT: wild type

<sup>e</sup> expressed as the mean survival score and the proposed interpretation

- Ollila S, Sarantaus L, Kariola R, et al. Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. *Gastroenterology* 2006;131:1408-17.
- Ollila S, Dermadi Bebek D, Jiricny J, et al. Mechanisms of pathogenicity in human MSH2 missense mutants. *Hum Mutat* 2008;29:1355-63.
- Drost M, Zonneveld JB, van Hees S, et al. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. *Hum Mutat* 2012;33:488-94.
- Lutzen A, de Wind N, Georgijevic D, et al. Functional analysis of HNPCC-related missense mutations in MSH2. *Mutat Res* 2008;645:44-55.
- Gammie AE, Erdeniz N, Beaver J, et al. Functional characterization of pathogenic human MSH2 missense mutations in *Saccharomyces cerevisiae*. *Genetics* 2007;177:707-21.
- Sia EA, Dominska M, Stefanovic L, et al. Isolation and characterization of point mutations in mismatch repair genes that destabilize microsatellites in yeast. *Mol Cell Biol* 2001;21:8157-67.



# ANNEXES



# ANNEXE 1

Les critères de classification des VSI développés par l'ACMG  
(*American College of Medical Genetics and Genomics*)

Adapté de (Richards *et al.*, 2015)



**Arguments en faveur de la pathogénicité :**

Argument de pathogénicité	Catégorie
<b>Très fort</b>	<p><b>PVS1</b> Variant ayant un effet nul prédit (non-sens ; décalage du cadre de lecture ; site canonique d'épissage : -2, -1, +1, +2 ; codon d'initiation de la traduction ; délétion d'un ou plusieurs exons hors phase ; délétion d'un ou plusieurs exons en phase mais emportant un domaine fonctionnel) dans un gène où la perte de fonction est le mécanisme pathogène connu.</p>
<b>Fort</b>	<p><b>PS1</b> Variant à l'origine du même changement d'acide aminé qu'un variant pathogène connu.</p> <p><b>PS2</b> Variant <i>de novo</i> chez un patient sans antécédent familial, avec confirmation de la paternité ET de la maternité.</p> <p><b>PS3</b> Etudes fonctionnelles <i>in vivo</i> ou <i>in vitro</i> bien établies montrant un impact délétère du variant sur le gène ou son produit.</p> <p><b>PS4</b> Prévalence du variant chez les individus atteints significativement supérieure à celle des contrôles de la même origine ethnique que les patients.</p>
<b>Moyen</b>	<p><b>PM1</b> Variant non tronquant situé sur un hot spot mutationnel et/ou un domaine fonctionnel critique bien établi (ex : site actif d'une enzyme) exempt de variants bénins.</p> <p><b>PM2</b> Variant absent des bases de données de populations contrôles (ou très faible fréquence si récessif) : GnomAD, ExAC, 1000 génomes.</p> <p><b>PM3</b> Variant observé en <i>trans</i> avec un variant pathogène distinct, si la pathologie a une transmission récessive.</p> <p><b>PM4</b> Variant affectant la longueur de la protéine : insertions/délétions en phase dans une région non-répétée ou perte du codon stop.</p> <p><b>PM5</b> Variant à l'origine d'un changement d'acide aminé différent à la même position qu'un variant faux-sens pathogène connu.</p> <p><b>PM6</b> Variant <i>de novo</i> sans confirmation de la paternité et de la maternité.</p>
<b>Faible</b>	<p><b>PP1</b> Variant qui ségrége avec la pathologie chez de multiples apparentés atteints.</p> <p><b>PP2</b> Variant faux-sens dans un gène avec un faible taux de variants faux-sens bénins et dans lequel les variants faux-sens sont un mécanisme fréquemment responsable de la pathologie.</p> <p><b>PP3</b> Effet d'un variant faux-sens prédit délétère par l'ensemble des logiciels de prédiction de pathogénicité interrogés (conservation, effet sur épissage...).</p> <p><b>PP4</b> Phénotype ou histoire familiale spécifique en faveur de la pathogénicité associée au gène comprenant le variant.</p> <p><b>PP5</b> Source documentée classant ce variant comme pathogène, dans la situation où les arguments permettant de définir la pathogénicité ne sont pas disponibles pour une réévaluation indépendante.</p>

### Arguments en faveur de la neutralité :

Argument de bénignité	Catégorie
<b>Suffisant</b>	<b>BA1</b> Fréquence allélique >5% dans les bases de données de populations contrôles
<b>Fort</b>	<p><b>BS1</b> Fréquence allélique trop importante par rapport à la fréquence de la pathologie.</p> <p><b>BS2</b> Présence chez un individu sain d'un variant homozygote pour une maladie récessive, hétérozygote pour une maladie dominante ou hémizygotique pour une maladie liée à X, si une pénétrance complète est attendue en bas âge.</p> <p><b>BS3</b> Etudes fonctionnelles bien établies montrant un impact non délétère du variant sur la fonction de la protéine ou sur l'épissage.</p> <p><b>BS4</b> Variant ne ségrégeant pas avec la pathologie chez des apparentés atteints</p>
<b>Faible</b>	<p><b>BP1</b> Variant faux-sens dans un gène où seuls les variants tronquants sont décrits comme associés à la pathologie</p> <p><b>BP2</b> Variant observé en <i>trans</i> avec un variant pathogène pour une maladie dominante à pénétrance complète, variant observé en <i>cis</i> avec un variant pathogène.</p> <p><b>BP3</b> Délétion/insertion en phase dans une région répétée sans fonction connue.</p> <p><b>BP4</b> Effet d'un variant faux-sens prédit bénin par l'ensemble des logiciels de prédiction de pathogénicité interrogés (conservation, effet sur épissage...).</p> <p><b>BP5</b> Co-occurrence du variant avec un variant pathogène dans un autre gène impliqué dans la même pathologie</p> <p><b>BP6</b> Source documentée classant ce variant comme bénin, dans la situation où les arguments permettant de définir la pathogénicité ne sont pas disponibles pour une réévaluation indépendante.</p> <p><b>BP7</b> Variant synonyme sans impact prédit sur l'épissage ET pour lequel la séquence nucléotidique n'est pas très conservée.</p>

### Interprétation :

#### Classe 5 : variant pathogène

- 1 **PVS** ET  $\geq$  1 **PS** OU  $\geq$  2 **PM** OU 1 **PM** et 1 **PP** OU  $\geq$  2 **PP**
- $\geq$  2 **PS**
- 1 **PS** ET  $\geq$  3 **PM** OU 2 **PM** et  $\geq$  2 **PP** OU 1 **PM** et  $\geq$  4 **PP**

#### Classe 4 : variant probablement pathogène

- 1 **PVS** et 1 **PM**
- 1 **PS** et 1-2 **PM**
- 1 **PS** ET  $\geq$  2 **PP**
- $\geq$  3 **PM**
- 2 **PM** et  $\geq$  2 **PP**
- 1 **PM** et  $\geq$  4 **PP**

#### Classe 3 : variant de signification incertaine

Les arguments pondérés ne permettent pas de classer le variant dans une autre classe

#### Classe 2 : variant probablement bénin

- 1 **BS** et 1 **BP**
- $\geq$  2 **BP**

#### Classe 1 : variant bénin

- **BA** OU
- $\geq$  2 **BS**

## ANNEXE 2

Les critères de classification des VSI utilisés dans la base de données InSiGHT  
(*International Society for Gastrointestinal Hereditary Tumors*)

Adapté de (Thompson *et al.*, 2014)



## Classe 5 : altération pathogène

- Variants avec une probabilité de pathogénicité > 0,99 avec un modèle de probabilité multifactorielle (non basé sur la co-ségrégation dans une seule famille et sans données tumorales conflictuelles).
- Variants dans le codon d'initiation de *MLH1*.
- Variants qui résultent en un codon stop prématuré (mutation non-sens ou altération du cadre de lecture) prédit pour aboutir à l'interruption de domaines protéiques fonctionnels connus ou de séquences hautement conservées (ne concerne donc pas le dernier exon de *MSH2*, *MSH6* ou *PMS2* et les codons après le codon 743 de *MLH1*).
- Variants pour lesquels une étude de l'ARNm issu d'un échantillon biologique constitutionnel du patient montre que l'allèle variant ne produit pas de transcrit sauvage et provoque un épissage aberrant conduisant à l'apparition d'un codon stop prématuré ou une délétion en phase conduisant à une perte d'un domaine fonctionnel ou d'une conformation de la protéine. Les altérations d'épissage doivent être confirmées par un test minigène ou autre étude expérimental de l'ARN par un laboratoire indépendant si l'aberration n'était pas prédite.
- Variant détecté *de novo* chez un patient présentant une tumeur du spectre de Lynch classique (phénotype MSI/perte d'expression de la protéine correspondante en immunohistochimie, pas de méthylation de *MLH1*) ; maternité et paternité confirmées.
- Large délétion génomique.
- Large duplication génomique pour laquelle une analyse expérimentale définit les points de cassure et montre un décalage du cadre de lecture avant la dernière jonction exon-intron.
- Variants présentant toutes les caractéristiques suivantes sans résultats conflictuels (les arguments combinés atteignent LR > 100:1) :
  - fonction altérée dans un test expérimental réalisé sur la protéine ou sur l'ARNm  
**OU** pour les gènes *MLH1* et *MSH2*, co-occurrence d'un variant clairement pathogène sur l'autre allèle du même gène chez un patient présentant les caractéristiques cliniques du CMMRD et une déficience du MMR dans les cellules normales et des cancers du spectre Lynch chez le parent qui a transmis le variant (si le génotype parental est inconnu, les deux parents doivent montrer des cancers du spectre Lynch à un âge précoce).  
**OU** présence du variant dans différents haplotypes dans les familles présentant les caractéristiques cliniques du syndrome de Lynch (permettant de conclure que le phénotype n'est pas dû à une altération de séquence ségrégeant en *cis* avec le variant).
  - Co-ségrégation avec la maladie permettant de calculer un rapport de vraisemblance  $\geq 10:1$  ou plus de 5 méioses informatives  
**OU** au moins une famille qui remplit les critères d'Amsterdam révisés avec au moins 4 porteurs affectés  
**OU** au moins deux familles qui remplissent les critères d'Amsterdam et présentent une co-ségrégation  
**OU** au moins deux familles avec au moins 3 porteurs affectés (en plus du cas index) dans chaque famille.
  - Au moins deux tumeurs indépendantes montrant un phénotype MSI grâce à un panel standard de 5 à 10 marqueurs et/ou perte d'expression de la protéine MMR correspondante.
  - Vérification que le variant ne soit pas un polymorphisme présent à une fréquence allélique supérieure à 1% dans des populations contrôles appropriées (ExAC, GnomAD).

#### **Classe 4 : altération probablement pathogène**

- Variants avec une probabilité de pathogénicité entre 0,95 et 0,99 avec un modèle de probabilité multifactorielle (ou > 0,99 si la co-ségrégation n'a été réalisée que dans une famille ou en présence de données tumorales conflictuelles).
- Variants dans les codons d'initiation de *MSH6* ou *PMS2*.
- Variants qui conduisent à une perte complète de l'expression de la protéine à partir de cet allèle (la présence du transcrit normal doit être analysé en présence et en absence d'inhibiteur du NMD).
- Variant détecté *de novo* chez un patient présentant une tumeur du spectre de Lynch classique (phénotype MSI/perte d'expression de la protéine correspondante en immunohistochimie, pas de méthylation de MLH1) ; maternité et/ou paternité non confirmée.
- Variants présents en position  $\pm 1$  ou  $\pm 2$  de la jonction exon-intron ou substitution G>non-G à la dernière base d'un exon si les 6 premières bases de l'intron ne sont pas GTRRGT et qui n'ont pas été testé *in vitro* pour un potentiel épissage aberrant (indépendamment des prédictions bio-informatiques).
- Variants présentant ((les arguments combinés atteignent LR > 20:1) :
  - Fonction altérée dans un test expérimental réalisé sur la protéine ou sur l'ARNm  
**OU** pour les gènes *MLH1* et *MSH2*, co-occurrence d'un variant clairement pathogène sur l'autre allèle du même gène chez un patient présentant les caractéristiques cliniques du CMMRD et une déficience du MMR dans les cellules normales et des cancers du spectre Lynch chez le parent qui a transmis le variant (si le génotype parental est inconnu, les deux parents doivent montrer des cancers du spectre Lynch à un âge précoce)  
**OU** présence du variant dans différents haplotypes dans les familles présentant les caractéristiques cliniques du syndrome de Lynch (permettant de conclure que le phénotype n'est pas dû à une altération de séquence ségrégeant en *cis* avec le variant).
  - Co-ségrégation avec la maladie permettant de calculer un rapport de vraisemblance  $\geq 5:1$  ou plus de 3 méioses informatives  
**OU** au moins une famille qui remplit les critères d'Amsterdam révisés avec au moins 3 porteurs affectés  
**OU** au moins deux familles avec au moins 2 porteurs affectés (en plus du cas index)  
**OU** au moins deux tumeurs indépendantes montrant un phénotype MSI grâce à un panel standard de 5 à 10 marqueurs et/ou perte d'expression de la protéine MMR correspondante.
- Une substitution faux-sens qui provoque le même changement d'acide aminé qu'une autre substitution faux-sens classée classe 5 avec un ARN normal et absent dans des populations contrôles appropriées.

#### **Classe 3 : altération de signification incertaine :**

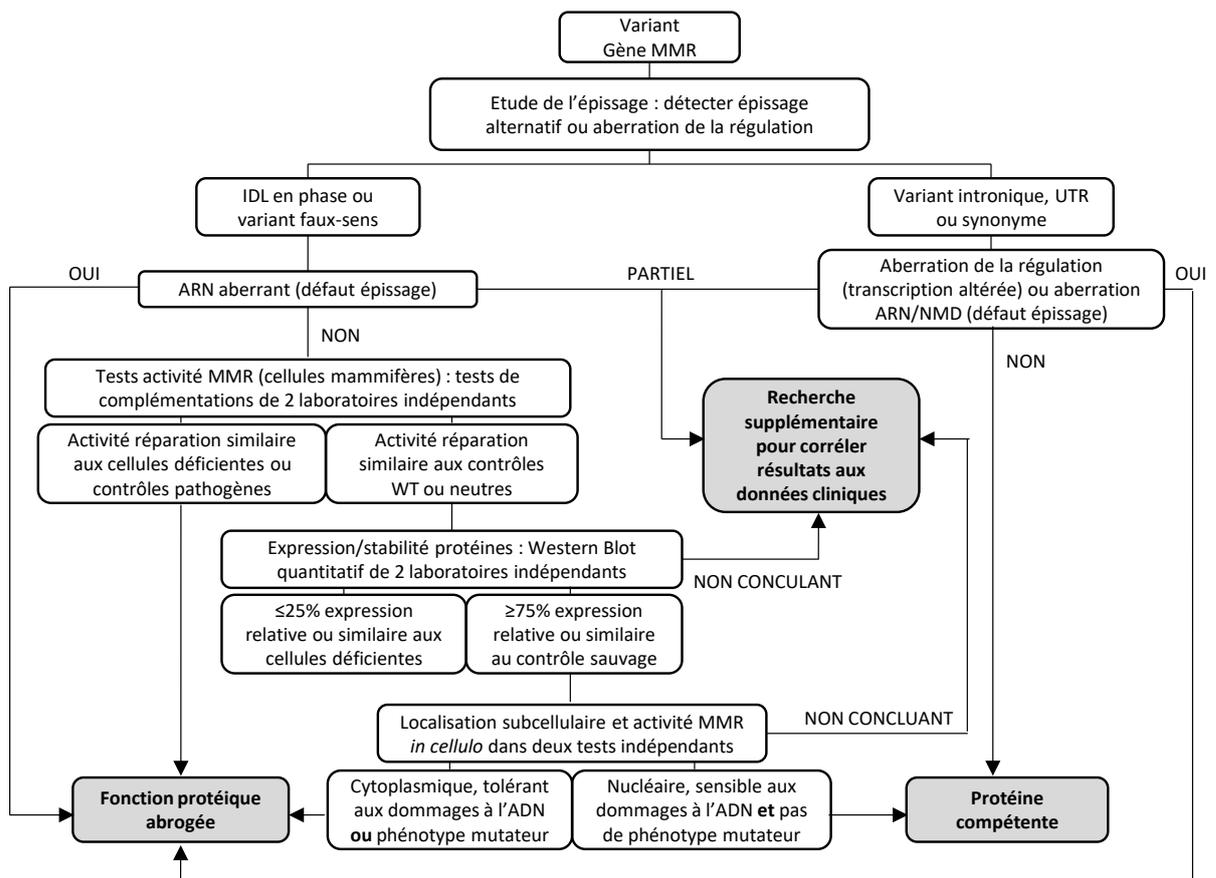
- Variants avec une probabilité de pathogénicité entre 0,05 et 0,949 avec un modèle de probabilité multi-factorielle.
- Variants dans le codon d'initiation de *MSH2*.
- Variants avec des arguments insuffisants pour classer (larges duplications génomiques pas encore testées pour résulter en un décalage du cadre de lecture avant la dernière jonction exon-intron ; altérations faux-sens ; petites délétion/insertions en phase ; variants synonymes ; variants introniques ; variants dans le promoteur ou dans les régions régulatrices).
- Variants présentant des données conflictuelles.

## **Classe 2 : altération probablement neutre**

- Variants avec une probabilité de pathogénicité entre 0,001 et 0,049 avec un modèle de probabilité multifactorielle.
- Substitutions synonymes et variants introniques sans aberrations de l'ARNm déterminées par des tests expérimentaux.
- Variants présents à une fréquence allélique  $\geq 1\%$  dans un groupe ethnique (testé sur  $\geq 160$  individus) et qui n'ont pas été exclus comme mutations fondatrices.
- Variants dans la région 5'-UTR ou du promoteur qui ont une expression bi-allélique.
- Variants présentant :
  - Fonction normale dans un test expérimental réalisé sur la protéine ou sur l'ARNm  
**OU** co-occurrence d'un variant clairement pathogène sur l'autre allèle du même gène chez un patient ayant un CCR après 45 ans (ou autre cancer du spectre de Lynch survenant à un âge plus tardif que dans le cadre du syndrome de Lynch) et qui n'a pas de manifestations de CMMRD.
  - Variant présent dans un groupe contrôle à une fréquence allélique 0,01-1% (testé sur  $\geq 160$  individus)  
**OU** absence de co-ségrégation de la maladie avec un variant conférant un risque dominant important ( $LR \leq 0,01$ )  
**OU** risque estimé  $< 4$  (avec intervalle de confiance de 95%), déterminé par une large étude cas-témoin bien conçue qui considère la taille, l'origine ethnique et géographique et la qualité des mesures de contrôle  
**OU** au moins 3 CCR présentant un phénotype MSS et/ou pas de perte d'expression d'une protéine MMR ; et/ou perte de protéine(s) MMR dans des tumeurs du spectre de Lynch ne correspondant pas au gène porteur du variant.
- Variant présentant trois points parmi les suivants :
  - présent dans un groupe contrôle de référence à une fréquence allélique entre 0,01 et 1% (testé sur  $\geq 160$  individus).
  - absence de co-ségrégation de la maladie avec un variant conférant un risque dominant important ( $LR \leq 0,01$ )
  - risque estimé  $< 4$  (avec intervalle de confiance de 95%), déterminé par une large étude cas-témoin bien conçue qui considère la taille, l'origine ethnique et géographique et la qualité des mesures de contrôle
  - au moins 3 CCR présentant un phénotype MSS et/ou pas de perte d'expression d'une protéine MMR ; et/ou perte de protéine(s) MMR dans des tumeurs du spectre de Lynch ne correspondant pas au gène porteur du variant.
- Un variant exonique qui provoque le même changement d'acide aminé qu'un variant faux-sens classés classe 1 avec un ARN normal (déterminés par résultats expérimentaux *in vitro*).

### **Classe 1 : altération neutre**

- Variants avec une probabilité de pathogénicité <0,001 avec un modèle de probabilité multifactorielle.
- Variants connus dans un groupe contrôle de référence à une fréquence allélique  $\geq 1\%$  (testé sur  $\geq 160$  individus) et qui ont été exclus comme mutations fondatrices.
- Variants présents dans la région 5'-UTR ou du promoteur présentant :
  - Expression bi-allélique.
  - Remplit les critères de fréquence de la classe 1 : présent dans les groupes contrôles de référence à une fréquence allélique  $\geq 0,001$  **OU** manque de co-ségrégation avec la maladie indiquant un variant dominant à haut risque d'être pathogène.
- Variant présentant :
  - fonction compétente dans des essais expérimentaux sur la protéine ou l'ARNm  
**OU** co-occurrence en *trans* avec une variation de séquence connue pour être pathogène dans le même gène chez un patient ayant un CCR après 45 ans (ou autre cancer du spectre de Lynch survenant à un âge plus tardif que dans le cadre du syndrome de Lynch) et qui n'a pas de manifestations de CMMRD
  - deux points parmi les suivants :
    - variant présent dans un groupe contrôle de référence à une fréquence allélique 0,01-1% (testé sur  $\geq 160$  individus)  
**OU** absence de co-ségrégation de la maladie avec un variant conférant un risque dominant important (LR  $\leq 0,01$ )  
**OU** risque estimé <4 (avec intervalle de confiance de 95%), déterminé par une large étude cas-témoin bien conçue qui considère la taille, l'origine ethnique et géographique et la qualité des mesures de contrôle  
**OU** au moins 3 CCR présentant un phénotype MSS et/ou pas de perte d'expression d'une protéine MMR ; et/ou perte de protéine(s) MMR dans des tumeurs du spectre de Lynch ne correspondant pas au gène porteur du variant.
- Variant présentant tous les points suivants :
  - variant présent dans un groupe contrôle de référence à une fréquence allélique 0,01-1% (testé sur  $\geq 160$  individus)  
**ET** absence de co-ségrégation de la maladie avec un variant conférant un risque dominant important (LR  $\leq 0,01$ )  
**ET** risque estimé <4 (avec intervalle de confiance de 95%), déterminé par une large étude cas-témoin bien conçue qui considère la taille, l'origine ethnique et géographique et la qualité des mesures de contrôle  
**ET** au moins 3 CCR présentant un phénotype MSS et/ou pas de perte d'expression d'une protéine MMR ; et/ou perte de protéine(s) MMR dans des tumeurs du spectre de Lynch ne correspondant pas au gène porteur du variant.



**FIGURE 11 | Interprétation des données fonctionnelles pour la caractérisation des VSI proposée par InSiGHT**

Pour les variants qui présentent une fonction MMR normale/non conclusive/intermédiaire dans deux essais indépendants mais qui présentent une fonction protéique déficiente dans deux essais indépendants, on conclut à une fonction abrogée. *Adapté de (Thompson et al., 2014).*



## CONCLUSION, DISCUSSION ET PERSPECTIVES



## Conclusion, discussion et perspectives

Mon laboratoire d'accueil travaille sur l'oncogenèse MSI, un mode de transformation caractérisé par la déficience d'un système de réparation de l'ADN : le système MMR. L'inactivation de ce système engendre une instabilité génétique, particulièrement marquée au niveau des microsatellites, qui a donné son nom à ces tumeurs. Si la réparation des mésappariements est la première fonction évoquée quand on parle du système MMR, celui-ci remplit également une autre fonction importante pour la tumorigenèse : la signalisation de certains dommages à l'ADN, notamment ceux provoqués par l'exposition à des agents méthylants. Les cellules tumorales MMR-déficientes présentent donc deux caractéristiques spécifiques : une instabilité des microsatellites et une tolérance aux agents méthylants.

Au cours de ma thèse, je me suis intéressée à l'exploitation du mécanisme de tolérance à la méthylation pour développer un test de caractérisation fonctionnelle des VSI des gènes MMR identifiés chez des patients suspectés d'être atteints du LS, syndrome génétique de prédisposition aux cancers MSI.

Les patients LS présentent un risque accru de développer de nombreux cancers (notamment des CCR) et se voient par conséquent proposer une surveillance. Ces mesures préventives permettent d'améliorer la durée et la qualité de vie du patient et il est par conséquent important de pouvoir poser un diagnostic. Comme tout syndrome génétique, le LS est dû à une mutation dans un gène, en l'occurrence l'un des 4 gènes du système MMR : *MLH1*, *MSH2*, *MSH6* ou *PMS2*. Le diagnostic repose sur l'identification d'une altération pathogène constitutionnelle sur l'un de ces gènes.

Des mutations d'autres gènes prédisposent au développement de CCR (par exemple *APC*, *MUTYH*, *STK1/LKB1*...). Néanmoins, l'héritabilité du CCR est encore loin d'être complètement élucidée. En effet, environ 15 à 20% des CCR sont observés chez des individus qui ont un apparenté au premier degré qui a également développé un CCR. Cependant, moins de 5% des patients atteints de CCR présentent un variant génétique dans un des gènes de susceptibilité au CCR connus. Bien que l'existence de gènes de prédisposition encore non identifiés puisse expliquer une partie de l'héritabilité manquante, des facteurs polygéniques et épigénétiques, environnementaux et comportementaux pourraient justifier ce risque familial inexpliqué (Boland *et al.*, 2018). Cependant, le fait que deux apparentés développent un CCR peut également être dû au hasard. En effet, le CCR est relativement fréquent dans la population générale, puisqu'il est estimé qu'un individu âgé de 50 à 74 ans a un risque de 3,5% de développer un cancer.

Actuellement, les techniques de séquençage de l'ADN évoluent très vite et permettent de réaliser plus rapidement et à moindre coût l'analyse de l'ADN constitutionnel des patients. Celle-ci se

fait de plus en plus grâce à un panel multi-gènes, qui permet d'analyser en parallèle les 4 gènes MMR, mais également d'autres gènes impliqués dans d'autres syndromes de prédisposition au cancer. Cette technique de recherche des mutations permet de répondre à de nombreuses interrogations, mais a également pour conséquence de flouter les frontières entre différents syndromes génétiques et de poser la question de l'inclusion de nouvelles localisations tumorales dans le spectre Lynch. Par exemple, l'utilisation d'un panel multigénique a permis de mettre en évidence des mutations dans les gènes *BRCA1/2* chez des patients suspectés d'être atteints d'un LS (Yurgelun *et al.*, 2015) mais il a également permis de détecter des patientes atteintes de cancer du sein porteuses de mutations dans les gènes MMR (alors que le cancer du sein n'est à l'heure actuelle pas considéré comme faisant partie du spectre Lynch) (Espenschied *et al.*, 2017).

Par ailleurs, l'analyse des résultats de séquençage se heurte à une autre difficulté : la détection de Variants de Signification Incertaine (VSI). Il s'agit généralement d'altérations faux-sens ou de petites insertions/délétions en phase dont il est difficile d'apprécier le caractère pathogène ou non. On ne sait pas si la variation de séquence conduit à l'expression d'une protéine fonctionnelle et on ignore donc si l'individu est atteint ou non du LS. Le nombre de VSI détectés, pour lesquels on ne peut pas rendre de conclusion définitive en l'état, ne peut qu'augmenter avec les années. Ainsi, non seulement l'utilisation du NGS soulève de plus en plus d'interrogations sur l'implication des différents gènes suppresseurs de tumeurs dans la tumorigenèse associée à différents syndromes génétiques, mais il permet également de mettre en évidence de plus en plus de VSI dans des gènes d'intérêt. A l'heure actuelle, la base de données InSiGHT recense 789 VSI sur un nombre total de 4033 variants uniques (tous gènes MMR confondus), soit 20%.

Dans le cadre du LS, des arguments permettent d'aider à la classification des variants détectés, tels que la présence du phénotype MSI dans la tumeur, la co-ségrégation du variant avec la maladie dans la famille, la fréquence du variant dans les bases de données de patients et de population, les études *in silico*, ou encore l'altération de l'épissage ou de la fonction de la protéine. Cependant, ces données ne sont pas toujours disponibles. L'absence d'arguments ou la présence de données contradictoires conduisent à l'impossibilité de classer ces variants, qui sont alors annotés « Variants de Signification Incertaine (VSI) » ou « variants de classe 3 ». Par ailleurs, beaucoup de VSI sont orphelins ou très rares et accumuler suffisamment d'arguments pour aider à leur interprétation peut prendre beaucoup de temps. Dans ce cas, les tests fonctionnels constituent une contribution importante. En effet, ils peuvent mettre en évidence une altération de l'activité MMR qui peut être utilisée pour prédire si le variant prédispose à la maladie ou s'il n'a pas d'influence significative sur le risque de cancer.

Un premier niveau de caractérisation fonctionnelle consiste à étudier l'épissage de l'ARN variant. En effet, un variant peut altérer un site consensus donneur ou accepteur d'épissage ou modifier un site exhausteur ou inhibiteur d'épissage qui peut être reconnu par des facteurs d'épissage. L'étude de l'épissage n'est pas une problématique restreinte à l'étude des variants des gènes MMR et nous ne développerons pas ce point. En revanche, le deuxième niveau de caractérisation fonctionnelle des VSI s'intéresse à l'effet du variant sur la protéine et nous allons nous intéresser plus en avant à cet aspect.

A l'heure actuelle, il n'existe pas de test « *gold standard* » pour l'analyse fonctionnelle de l'effet des VSI des gènes MMR sur la fonction de la protéine : aucun n'est considéré comme une référence et n'est réalisé de façon systématique en prospectif sur l'ensemble des VSI détectés. Pourtant, durant les 20 dernières années, une cinquantaine d'études ont été publiées qui traitent de l'étude fonctionnelle d'une série de variants MMR. Les tests fonctionnels décrits analysent soit un aspect restrictif de la protéine soit une fonction du système MMR dans sa globalité. Ils peuvent être réalisés dans différents modèles (levure, souris, cellules humaines) et être menés *in vivo* ou *in vitro*. Nous allons revenir sur l'ensemble de ces possibilités et présenter les avantages et les limitations de chaque méthode.

## A. Les tests fonctionnels disponibles : état de l'art

### 1. *Les tests restrictifs versus les tests globaux*

Les tests fonctionnels peuvent d'une part analyser un aspect restrictif de la protéine, tel que son expression et sa stabilité, son interaction avec ses protéines partenaires, sa localisation nucléaire, sa capacité à réaliser les différentes étapes de la fonction de réparation des mésappariements (liaison au mésappariement, activité ATPase...) ou encore son aptitude à empêcher la recombinaison homologue (Ou *et al.*, 2007 ; Rasmussen *et al.*, 2012). D'autre part, les tests peuvent également évaluer une des deux fonctions du système MMR (réparation des mésappariements et signalisation des dommages à l'ADN) dans sa globalité.

Dans les deux cas, des tests différents peuvent permettre d'analyser la même fonction. Par exemple, la capacité des protéines variantes à interagir avec leur partenaire (aspect restrictif) peut être étudiée par co-immunoprécipitation dans des cellules humaines (Perera et Bapat, 2008) ou peut être analysée par un test double hybride chez la levure (Kondo *et al.*, 2003). La capacité des protéines variantes à réparer les mésappariements (aspect global) peut être analysée par complémentation fonctionnelle chez la levure (Drotschmann *et al.*, 1999) ou grâce à un modèle « *cell free* » (Drost *et al.*, 2010).

L'inconvénient majeur des tests évaluant un aspect restrictif de la protéine réside dans la probabilité de faux-négatifs. En effet, il est possible qu'un résultat négatif puisse être dû au fait que le VSI analysé altère un domaine qui n'est pas impliqué dans la fonction étudiée alors que ce VSI montrerait effectivement une perte de fonction avec un test relatif à la fonction altérée (Couch *et al.*, 2008). Pour pallier à cette limitation, il est judicieux de tester une fonction du MMR dans sa globalité : réparation des mésappariements ou signalisation des dommages à l'ADN.

## **2. Test de la fonction de réparation des mésappariements versus signalisation apoptotique**

Environ 80% des études publiées sur la caractérisation fonctionnelle des VSI analysent l'activité de réparation d'une manière ou d'une autre. Trois grandes méthodes d'étude de cette fonction sont décrites à l'heure actuelle. Premièrement, chez la levure, grâce à la complémentation d'une lignée MMR-déficiente, à l'effet dominant mutateur ou à la modification du gène d'une lignée MMR-compétente. Ensuite, *in vitro* grâce à un système à 3 composants : les protéines variantes (produites en cellules d'insectes généralement), l'extrait nucléaire d'une lignée cellulaire MMR-déficiente et un substrat ADN contenant un mésappariement. Enfin, dans des cellules humaines : la protéine variante est exprimée dans une lignée originellement MMR-déficiente et la fonction de réparation est analysée grâce à un plasmide rapporteur de l'instabilité des microsatellites.

Au contraire, les tests qui étudient la fonction de signalisation sont beaucoup plus anecdotiques. Approximativement 20% des études (8 publications de 4 équipes différentes) portent sur l'étude de la fonction de signalisation (Arora *et al.*, 2017 ; Blasi *et al.*, 2006 ; Mastrocola et Heinen, 2010 ; Houlleberghs *et al.*, 2016 ; Houlleberghs *et al.*, 2017 ; Wiolders *et al.*, 2011 ; Wiolders *et al.*, 2013; Wiolders *et al.*, 2014).

La fonction de réparation des mésappariements est donc la plus souvent utilisée dans les tests de caractérisation des VSI qui ont été développés jusqu'alors. Cependant, il a été suggéré que l'inactivation de la signalisation des dommages à l'ADN pourrait être au moins aussi importante pour la tumorigenèse que la perte de la réparation des dommages en elle-même (Heinen *et al.*, 2002). Dans les cellules MMR<sup>+/-</sup> d'un individu Lynch, l'inactivation somatique du système MMR est provoqué par un événement mutationnel spontané et aléatoire sur l'allèle sauvage. On peut imaginer que c'est alors l'inactivation de la fonction de signalisation qui confère un avantage sélectif aux cellules MMR-déficientes, notamment si celles-ci sont exposées à des mutagènes environnementaux comme c'est le cas dans l'intestin. D'ailleurs, l'expérience de Reese montre que sans pression de sélection, les cellules MMR-déficientes sont contre-sélectionnées (Reese *et al.*, 2003). Par contre, c'est ensuite l'accumulation des mutations au cours des divisions qui finit par provoquer l'activation d'un oncogène

ou l'inactivation d'un gène suppresseur de tumeur. Ainsi, il y a une notion de temporalité : la pression de sélection provoquée par le mécanisme de tolérance aux dommages à l'ADN pourrait constituer l'événement initial de la tumorigenèse, notamment dans le colon. L'accumulation des mutations, via l'inactivation de la réparation des mésappariements, conduirait ensuite à l'accélération de la progression tumorale (Fishel, 2001). Il a été décrit que l'évolution via la sélection naturelle dépend de deux éléments essentiels : la pression de sélection et la possibilité de l'apparition de variation génétique. L'inactivation du système MMR permet effectivement ces deux aspects, via la perte de la fonction de signalisation et la perte de la fonction de réparation, respectivement (Breivik et Gaudernack, 1999).

### ***3. Les différents modèles expérimentaux : levure, cellules murines et humaines***

L'utilisation du modèle levure est tentant car les résultats sont obtenus rapidement et de nombreux VSI peuvent être testés en parallèle. Cependant, l'homologie entre l'homme et la levure reste la contrainte principale de ces expérimentations. D'une part, les protéines MMR de la levure ne présentent pas 100% d'homologie avec les protéines humaines et donc ne permettent pas l'analyse de l'ensemble des VSI détectés chez les individus. D'autre part, la substitution d'un acide aminé conservé au cours de l'évolution chez des espèces distantes n'a pas forcément les mêmes conséquences fonctionnelles (Wielders *et al.*, 2011). Même si on modifie le nucléotide homologue, le phénotype qui en découlera ne sera pas nécessairement directement transposable à l'homme. Les levures et l'homme présentent également des différences dans leurs modifications post-traductionnelles et leurs interactions protéines-protéines. Enfin, le fond génétique de la levure peut affecter la fonctionnalité de certains VSI : un même VSI étudié dans deux souches différentes peut donner des résultats différents, comme nous le verrons plus loin.

La même problématique d'homologie existe quand on travaille sur des cellules souches embryonnaires de souris : les protéines MSH2 et MSH6 de l'homme et de la souris présentent 93% et 86% d'homologie, respectivement (Houllberghs *et al.*, 2016 ; Houllberghs *et al.*, 2017). De plus, des différences fonctionnelles sont possibles entre les protéines de l'homme et de la souris.

De manière générale, les modèles levure et murin ne permettent de travailler que sur les altérations exoniques.

Les modèles de lignées cellulaires humaines présentent également des limites. En effet, les lignées utilisées pour l'analyse des VSI des gènes MMR sont principalement des lignées MMR-déficientes, qui ont accumulé et continuent à accumuler des mutations au cours des divisions et ne présentent donc pas un fond génétique stable. Par ailleurs, le gène *MRE11*, un composant de la

signalisation des cassures doubles brin régulée par ATM, est inactivé dans plusieurs lignées MMR-déficientes, notamment HCT116 et LoVo (Giannini *et al.*, 2002 ; Takemura *et al.*, 2006).

#### **4. Tests *in vitro* versus *in vivo***

Les tests *in vitro* permettent d'étudier un aspect restrictif de la fonction ou une fonction MMR dans sa globalité en dehors du contexte cellulaire. Cependant, les tests *in vitro* ne récapitulent pas exactement l'environnement dans lequel les protéines MMR fonctionnent (Heinen et Rasmussen, 2012). Pour remplir leur fonction *in vivo*, les protéines MMR doivent être adressées au noyau et potentiellement subir des modifications post-traductionnelles, ce qui ne peut pas être étudié *in vitro*. De plus, les protéines *in vivo* agissent dans le contexte de la chromatine de cellules en train de répliquer leur ADN.

Cette absence de contexte physiologique peut conduire à des résultats erronés. Par exemple, certains variants peuvent présenter une activité normale dans un test de réparation *in vitro* et seront donc considérés comme MMR-compétents. Cependant, ils peuvent avoir un défaut de localisation, de taux d'expression ou de stabilité qui n'est pas évalué par le test en question. Par conséquent, alors qu'une perte de fonction observée dans un test suggère une pathogénicité du VSI, l'inverse n'est pas forcément vrai car une fonction normale dans un test n'indique pas nécessairement une neutralité du VSI. Ainsi, il est conseillé que ces variants soient étudiés plus en détail en ce qui concerne la stabilité, l'interaction protéine-protéine ou la localisation dans la cellule par exemple (Couch *et al.*, 2008). Il est donc probable que la classification finale de neutralité requière l'intégration de multiples arguments.

#### **5. Le test d'expression protéique**

L'expression des protéines, estimée par Western Blot, indique la quantité de protéine stable à un instant donné. Le taux de protéines dans la cellule peut être modifié à différents niveaux : la transcription et la traduction peuvent être affectées, tout comme la stabilité de la protéine. De plus, de nombreux exemples, dans la littérature comme dans notre étude, ont montré qu'il n'y a pas nécessairement une corrélation entre le taux d'expression d'une protéine et sa fonctionnalité. Au sein de notre cohorte de variants *MLH1*, nous pouvons mentionner le variant V716M qui fait partie de la cohorte contrôle de variants neutres, et qui présente cependant un faible taux d'expression protéique, comme l'ont déjà observé d'autres auteurs (Hinrichsen *et al.*, 2013). Par ailleurs, deux VSI *MLH1* que nous avons classés comme déficient et potentiellement déficient dans notre test présentent une expression protéique comparable à la protéine sauvage : le variant S106R a été observé chez un patient dont la tumeur conserve un marquage des protéines *MLH1* et *PMS2* et le variant L85P est associé à

une perte somatique isolée de PMS2. Dans ces 2 cas, on observe donc l'expression d'une protéine MLH1 qui a perdu sa fonctionnalité. Le mécanisme de pathogénicité du variant L85P pourrait faire intervenir la perte d'interaction avec sa protéine partenaire PMS2. Ainsi, une protéine faiblement exprimée peut conserver sa fonctionnalité alors qu'une protéine exprimée à un taux semblable à celui de la protéine sauvage peut être non fonctionnelle. En ce qui concerne la cohorte de variants *MSH2*, nous n'avons observé aucune corrélation entre le taux d'expression protéique et le taux de survie après exposition à un agent méthylant. Deux autres équipes étudiant des variants *MSH2* ont déjà remarqué une absence de corrélation entre les résultats de leur test fonctionnel et l'expression protéique par Western Blot. L'équipe d'Ollila note qu'aucun des variants n'altère l'expression des protéines, bien que certains variants altèrent la fonction de réparation (Ollila *et al.*, 2006) et l'équipe de Lützen remarque que les protéines *MSH2* variantes ont un taux d'expression comparable à la protéine *MSH2* sauvage, alors qu'elles présentent des réponses différentes dans divers tests fonctionnels (Lützen *et al.*, 2008).

Par ailleurs, la limitation de la quasi-totalité des tests fonctionnels disponibles est l'expression exogène des protéines. En effet, beaucoup des approches développées sont basées sur la transfection d'un plasmide permettant la surexpression de la protéine à partir d'un promoteur fort. Cette forte expression peut être toxique et perturber l'activité de la protéine ou au contraire elle est susceptible de masquer une perte d'activité partielle d'un variant hypomorphe qui en conditions physiologiques ne serait pas capable de remplir sa fonction mais qui le devient en cas de surexpression. Ainsi, il est possible que la forte expression des protéines provoque des faux-négatifs (variant pathogène qui apparaissent faussement neutres) (Sijmons *et al.*, 2013).

## 6. *Les résultats discordants*

Il est possible qu'un même VSI puisse donner des résultats différents selon les tests utilisés, que ce soient des tests pour étudier le même aspect fonctionnel de la protéine ou pour en étudier différents aspects.

Par exemple, une équipe a étudié plusieurs variants *MLH1* grâce à deux tests d'étude d'interaction des protéines : co-immunoprécipitation sur extraits cellulaires totaux dans des cellules humaines et test double hybride chez la levure (Fan *et al.*, 2007). Une discordance de résultats a été observée pour certains variants qui présentaient une interaction correcte en co-IP mais une mauvaise interaction en test double hybride. L'hypothèse avancée par les auteurs est que le complexe MutL $\alpha$  variant est capable de se former dans le cytoplasme et est donc détecté en co-immunoprécipitation mais n'est pas capable d'être correctement localisé au noyau, d'où un résultat négatif dans le test double hybride.

Les résultats sont susceptibles d'être discordants quand plusieurs équipes étudient le même variant pour des fonctions différentes. Nous prendrons l'exemple des variants *MLH1* R265C et K618A (**TABLE 4**). Les différences peuvent être dues au fond génétique différent entre les différentes lignées cellulaires utilisées ou encore au changement d'autre(s) variable(s) d'un laboratoire à l'autre. Ainsi, on se rend compte de l'importance de développer un test avec lequel on pourrait tester tous les variants afin de pouvoir comparer les résultats. Cela implique d'étudier une fonction globale du système MMR, étant donné que tous les acides aminés ne sont pas impliqués dans l'ensemble des différentes fonctions de la protéine.

<b>MLH1 R265C</b>	Expression/stabilité	Fonction réparation
(Trojan <i>et al.</i> , 2002)	Transfection cellules 293T : <b>expression normale</b>	Transfection cellules 293T → test MMR <i>in vitro</i> : <b>réparation normale</b>
(Perera et Bapat, 2008)	Transfection cellules HCT116 : <b>réduction de la stabilité protéique</b>	
(Takahashi <i>et al.</i> , 2007)		Transfection cellules HCT116 → test MMR <i>in vitro</i> : <b>réparation intermédiaire</b>
(Plotz <i>et al.</i> , 2006)		Transfection cellules 293T → test MMR <i>in vitro</i> : <b>pas de réparation</b>

<b>MLH1 K618A</b>	Interaction protéiques	Expresion/stabilité	Localisation	Fonction réparation
(Guerrette <i>et al.</i> , 1999)	GST pull down : <b>pas d'interaction PMS2</b>			
(Raevaara <i>et al.</i> , 2005)	Expression MLH1 K618A et PMS2 WT dans cellules insectes → Western Blot et co-immunoprécipitation : <b>expression normale MLH1 et interaction normale PMS2</b>		Transfection cellules 293T : <b>localisation normale</b>	Test MMR <i>in vitro</i> : <b>réparation normale</b>
(Kondo <i>et al.</i> , 2003)	Test levure double hybride : <b>interaction normale PMS2, interaction réduite EXO1</b>			
(Belvederesi <i>et al.</i> , 2006)	GST pull down : <b>interaction PMS2 réduite</b>	Transfection cellules HCT116 K618A et PMS2 WT : <b>expression légèrement diminuée</b>		

**TABLE 4 | Exemple de résultats discordants dans l'interprétation des VSI par différentes équipes**

Enfin, une équipe a étudié le taux de mutations sur un microsatellite rapporteur de levure exprimant des variants *MLH1*. Le phénotype mutateur induit par chaque variant a été attribué à une des classes suivantes : « *Non Mutator* », « *Low Mutator* », « *Low-Intermediate mutator* », « *Intermediate Mutator* », « *Mutator* » et « *High Mutator* ». L'étude a été conduite en parallèle dans deux souches de levures différentes. Il est intéressant de noter que tous les variants n'ont pas été classés de la même manière dans les deux souches. La plus grande différence de classement observé est le passage de la catégorie « *Low Mutator* » à la catégorie « *High Mutator* ». Les auteurs observent que plus le taux de mutations augmente, moins il y a de changement de catégorie. Cela suggère selon eux qu'il y a un seuil au-dessus duquel le fond génétique n'a plus d'impact sur la pathogénicité de

l'allèle. Ils proposent donc que les allèles conférant des phénotypes moyennement mutateurs sont plus susceptibles de montrer des différences selon le fond génétique (Wanat *et al.*, 2007).

## **7. Le paramétrage et la détermination des performances**

La limite de la presque totalité des tests fonctionnels de caractérisation des VSI est qu'ils n'ont pas été paramétrés en amont sur une cohorte cas-contrôle bien définie constituée de variants connus pour être clairement neutres et clairement pathogènes. Par conséquent, la sensibilité et la spécificité ainsi que l'éventuelle valeur seuil qui permettrait de discriminer les variants pathogènes des variants neutres n'ont pas pu être déterminés. Les résultats ne sont donc que descriptifs et expriment par exemple un taux d'activité de la protéine variante en fonction du taux d'activité de la protéine sauvage (fixée par défaut à 100%), sans pour autant conclure quant à la conséquence de ce taux sur la fonctionnalité de la protéine et donc sans en déduire le caractère pathogène ou neutre du VSI étudié. C'est pourquoi les tests déjà développés appartiennent au monde de la recherche plutôt qu'à l'univers du diagnostic de routine.

Les cliniciens sont donc dans l'attente d'un test standardisé, validé sur une cohorte cas-contrôle, dont la sensibilité et la spécificité ont été déterminés et sur lequel ils pourraient s'appuyer avec confiance. Un tel test devrait permettre l'analyse de tous les VSI d'un même gène, de façon simple et rapide et à grande échelle, afin de pouvoir être appliqué dans un laboratoire de diagnostic de routine.

## **B. Notre test de caractérisation fonctionnelle des VSI des gènes MMR dans le contexte du syndrome de Lynch**

### **1. Test d'aide au diagnostic du CMMRD**

Dans notre laboratoire, un test d'aide au diagnostic du CMMRD a été développé récemment (Bodo *et al.*, 2015). Le CMMRD, syndrome de déficience constitutionnelle du système MMR, est dû à une mutation bi-allélique d'un gène MMR et prédispose au développement de tumeurs MSI dès l'enfance. Le diagnostic est basé sur la détection d'altérations génétiques des gènes MMR dans l'ADN germlinal du patient. Le diagnostic peut être compliqué en cas de détection d'un ou de deux VSI ou si l'analyse permet de mettre en évidence aucune ou seulement une altération.

Brièvement, cette étude montre que l'utilisation des deux caractéristiques spécifiques des cellules tumorales MMR-déficientes (le phénotype MSI et le phénotype de tolérance aux agents méthylants) peut aider au diagnostic du CMMRD. Pour ce faire, des cellules lymphoblastoïdes du patient sont immortalisées par infection par le virus EBV. Le phénotype MSI et la réponse cellulaire à

un agent méthylant sont étudiées dans la lignée lymphoblastoïde. Si la lignée présente un statut MSI et un phénotype de tolérance à la méthylation, cela indique que les cellules lymphoblastoïdes sont MMR-déficientes et le diagnostic de CMMRD est confirmé.

## ***2. Développement de notre test***

Nous avons souhaité adapter ce test de réponse à la méthylation pour aider cette fois-ci au diagnostic du LS, dans le contexte de la détection de VSI. La même approche ne peut pas être utilisée en l'état, car les patients LS présentent une mutation mono-allélique d'un gène MMR et leurs cellules sont donc MMR-compétentes. La stratégie que nous avons choisie est celle de la complémentation fonctionnelle : un plasmide porteur de la séquence du VSI *MLH1* (ou *MSH2*) à tester est transfecté dans la lignée HCT116 *MLH1*-déficiente (ou LoVo *MSH2*-déficiente). Le plasmide porte également la séquence du gène rapporteur *mCherry* (ou *eGFP*). Les cellules transfectées sont triées par cytométrie de flux puis remises en culture. Elles sont ensuite exposées à un agent méthylant, et un test de clonogénicité permet d'estimer leur taux de survie par rapport aux mêmes cellules non traitées. Nous avons commencé ce projet par une étude de faisabilité, au cours de laquelle nous avons testé la capacité de l'ADNc du gène *MLH1* (ou *MSH2*) sauvage à restaurer la sensibilité à l'agent méthylant des cellules HCT116 (ou LoVo), initialement MMR-déficientes et donc tolérantes à la méthylation. Nous avons poursuivi par une étude cas-contrôle au cours de laquelle nous avons analysé pour chaque gène une vingtaine de variants connus pour être pathogènes ou neutres, afin de déterminer les conditions optimales de l'expérimentation, la sensibilité et la spécificité mais également la valeur seuil séparant la réponse des variants pathogènes et neutres. Une fois les conditions de l'expérimentation établies, nous avons testé 25 VSI du gène *MLH1* et 24 VSI *MSH2* afin de les caractériser. Parmi l'ensemble des 49 VSI, 12 ont été classés comme déficients et 29 comme neutres ; 8 présentent une restauration intermédiaire de la fonction de signalisation des dommages à l'ADN, empêchant leur classification.

Nous avons également étudié l'expression des protéines variantes par Western Blot sur cellules triées. Enfin, nous avons comparé nos résultats aux autres données disponibles : études *in silico*, données cliniques des patients et conclusions des autres tests fonctionnels publiés.

Le but de notre test fonctionnel de caractérisation des VSI des gènes MMR n'est pas de remplacer l'intégralité de la classification des variants mais de proposer une plus-value pour la classification des VSI, variants pour lesquels seuls les tests fonctionnels peuvent apporter des arguments permettant leur classification finale (cf. **ARTICLE**).

### 3. *Les avantages et les limites de notre test*

Une principale limite de notre méthode est la forte surexpression des protéines due au fait que l'expression se fait à partir d'un promoteur fort. En cas d'expression ectopique, le taux d'expression, le traitement des ARNm ou encore les régulations épigénétiques peuvent être altérés (Wiolders *et al.*, 2011). Une potentielle perte d'activité (totale ou partielle) de la protéine MMR en conditions physiologiques peut être masquée par la surexpression de la protéine. De même, un haut taux de protéines peut être toxique pour la cellule ou peut perturber le fonctionnement normal de la protéine (Rasmussen *et al.*, 2012). En effet, il a été montré qu'à la fois une expression trop faible mais aussi trop forte des protéines MMR affectent les capacités du système MMR.

Une autre limite est liée au fait que nous étudions un variant en dehors du contexte génétique du patient chez qui il a été détecté. Or, il n'est pas rare qu'un individu soit porteur d'un autre variant (neutre, pathogène ou de signification incertaine) sur le même allèle ou sur l'autre allèle d'un même gène ou encore sur un autre gène MMR. Il est possible qu'un VSI donné n'ait pas d'effet individuellement mais puisse avoir un effet en synergie avec un autre variant (Martinez et Kolodner, 2010). On peut citer l'exemple du variant *MSH2* c.380A>G qui est considéré comme neutre car compétent dans plusieurs tests fonctionnels et présent chez de nombreux individus sains. Il est cependant très fréquemment détecté chez des individus atteints de CRC (Kantelinen *et al.*, 2012). C'est par exemple la quatrième substitution faux-sens classée neutre en terme de fréquence, sur un total de 29, sur la base de données UMD. Ce variant pourrait présenter un défaut d'activité extrêmement subtil, qui serait potentialisé en présence d'un autre variant. Par ailleurs, d'autres altérations dans d'autres gènes peuvent également influencer le risque de cancer. Notre technique ne permet en l'état d'analyser qu'un seul variant à la fois et n'autorise donc pas d'étudier l'interaction entre plusieurs variants.

Ensuite, notre test est dans l'incapacité de détecter des variants *MSH2* dont la pathogénicité serait due à un défaut d'interaction avec sa protéine partenaire MSH3. En effet, le test de réponse à la méthylation fait intervenir la reconnaissance d'un mésappariement, réalisée par le complexe MutS $\alpha$  (MSH2-MSH6). Le complexe MutS $\beta$  (MSH2-MSH3), qui reconnaît les boucles d'insertion/délétion n'est pas impliqué dans la fonction de signalisation des dommages induits par un agent méthylant. Ainsi, notre test ne permet pas de mettre en évidence les variants *MSH2* qui seraient déficients pour l'interaction entre MSH2 et MSH3 (Houllberghs *et al.*, 2016).

Enfin, il est nécessaire de s'assurer en amont que l'épissage n'est pas affecté par le variant. En effet, certains variants pathogènes sont connus pour affecter non pas la protéine mais l'épissage. Or, les plasmides que nous utilisons ne portent que la séquence codante du gène d'intérêt. Ainsi, un variant d'épissage ne sera pas détecté par notre méthode. Nous nous sommes donc assurés de ne

sélectionner que des variants pour lesquels l'épissage n'était pas altéré. Dans le cas d'un variant d'épissage dont la séquence ARN est connue, la séquence du gène peut être modélisée en conséquence sur le plasmide, afin de pouvoir quand même tester ce variant.

Dès le début du projet, nous avons pensé notre test dans l'optique de pouvoir le transposer par la suite dans un laboratoire de diagnostic de routine. Nous avons donc cherché à développer une technique la moins lourde et la plus rapide possible, et qui permet d'analyser un certain nombre de variants en parallèle. Notre test a l'avantage d'être réalisé dans un modèle cellulaire humain et d'analyser une fonction globale du système MMR *in vivo*. En outre, l'atout de notre méthode est incontestablement d'avoir réalisé une étude cas-contrôle, qui nous a permis de déterminer la sensibilité et la spécificité de notre test, ainsi que de définir les valeurs seuil au-dessus desquelles un variant *MLH1* ou *MSH2* devrait être considéré pathogène. Enfin, la comparaison des résultats de notre test fonctionnel aux autres données disponibles (données cliniques, *in silico* et expression des protéines variantes) permet d'asseoir la validité de nos résultats. Par ailleurs, notre technique ne nécessite pas de matériel biologique du patient.

Il s'agit d'une méthode qui ne requiert pas une connaissance technique très développée. Cependant, elle nécessite de la culture cellulaire et du tri cellulaire par cytométrie en flux que tous les laboratoires ne sont pas en mesure de mettre en œuvre. A l'avenir, nous espérons que notre technique pourra être mise en place dans un petit nombre de laboratoires référents qui pourront étudier, en rétrospectif et en prospectif, l'intégralité des VSI connus.

## C. Les questions ouvertes

### ***1. Découplage des deux fonctions du système MMR***

Nous avons vu que de nombreuses études se concentrent sur la fonction de réparation des mésappariements du système MMR. Cependant, nous considérons qu'il est hasardeux de n'analyser que la fonction de réparation, au détriment de la fonction de signalisation des dommages. En effet, il est possible qu'un variant conduise à l'expression d'une protéine qui soit capable de remplir une fonction et pas l'autre et extrapoler sur le caractère pathogène ou non du variant à partir de l'étude d'une seule des deux fonctions pourrait donc conduire à des conclusions erronées.

i. Variants compétents pour la réparation mais déficients pour la signalisation

Dans plusieurs modèles *in vitro* ou *in vivo*, il a été montré que des cellules avec une expression réduite de MLH1, MSH2 ou MSH6 ont une activité de réparation des mésappariements normale, mais sont plus tolérantes aux composés endommageant l'ADN (indiquant une activité de signalisation des dommages à l'ADN défaillante) (Cejka *et al.*, 2003 ; Claij *et al.*, 2003 ; Evensen *et al.*, 2018 ; Lettieri *et al.*, 1999 ; O'Brien et Brown, 2006). Il a été suggéré que le taux de protéine nécessaire à la fonction de signalisation était plus élevé que le taux de protéine nécessaire à la fonction de réparation (Fedier et Fink, 2004). Par ailleurs, il a été décrit que les cellules lymphoblastoïdes Msh2<sup>+/-</sup> issues de patients Lynch sont 4 fois plus résistantes au témozolomide que les cellules Msh2<sup>+/+</sup> (Marra *et al.*, 2001). Une résistance augmentée au MNNG des lignées lymphoblastoïdes MSH2<sup>+/-</sup> et encore plus des lignées MSH6<sup>+/-</sup> a également été observée au laboratoire (Bodo *et al.*, 2015). Par contre, il n'y a pas d'haplo-insuffisance de ces cellules vis-à-vis de la fonction de réparation des mésappariements car les patients porteurs d'une mutation délétère hétérozygote MSH2 ou MSH6, ne présentent pas de phénotype MSI.

Ainsi, en fonction du taux de protéine produit, il est théoriquement envisageable d'observer un variant qui présente une bonne réparation des mésappariements mais une mauvaise signalisation des dommages ce qui peut conduire à des résultats discordants dans les deux tests fonctionnels. Cette situation est très peu probable dans notre étude, car dans notre protocole expérimental, les protéines sont largement surexprimées par rapport au taux d'expression endogène.

ii. Variants déficients pour la réparation et compétents pour la signalisation

Un tel découplage des fonctions du système MMR a été mis en évidence pour quelques variants MSH2 et MSH6 chez la levure (Drotschmann *et al.*, 2004) ainsi que pour le variant MSH2 G674A et le variant MSH6 T1219D chez la souris (Lin *et al.*, 2004 ; Yang *et al.*, 2004). Ces deux derniers variants ont ensuite été étudiés dans un modèle cellulaire humain, dans lequel les résultats ont été confirmés (Geng *et al.*, 2012). Ces protéines variantes sont déficientes pour la réparation mais compétentes pour la signalisation des dommages. Il a été montré que les complexes MSH2-MSH6 comprenant l'une de ces protéines variantes sont capables de reconnaître un mésappariement ou un dommage à l'ADN mais qu'elles ont perdu l'activité ATPasique nécessaire au changement de conformation de MutS $\alpha$  qui lui permet de glisser le long de l'ADN. En conséquence, MutS $\alpha$  reste fixé au mésappariement et est incapable d'activer le processus de réparation (Geng *et al.*, 2012).

L'existence de telles altérations a des conséquences théoriques à différents niveaux. Tout d'abord, dans la compréhension du mécanisme de tolérance à la méthylation même, puisque le fait que ces protéines variantes ne soient pas en mesure d'entreprendre les cycles futiles d'essais de

réparation donne un argument en faveur du modèle de signalisation directe, bien que les deux modèles ne soient pas mutuellement exclusifs (cf. **FIGURE 3** chapitre 1). Ensuite, sur le mécanisme de tumorigenèse dans le LS puisque les souris qui portent ces mutations développent également des cancers mais plus tardifs que ceux développés par les souris porteuses de mutations qui abolissent les deux fonctions MMR. Ainsi, l'inactivation de la réparation des mésappariements est suffisante pour développer un cancer mais l'inactivation concomitante de la signalisation des dommages accélère le processus.

De telles altérations n'ont à ma connaissance pas été décrites chez des patients potentiellement atteints du LS. Cependant, on peut supposer qu'elles existent. Si on ne teste qu'une des deux fonctions du système MMR, on pourrait conclure à tort sur la pathogénicité ou la neutralité du variant. Nous proposons donc qu'un VSI déficient pour la fonction de signalisation dans notre test est probablement pathogène, alors que l'inverse n'est pas vrai. C'est la raison pour laquelle nous proposons de tester la fonction de réparation de ces variants dans cette situation (cf. **FIGURE 6** dans l'**ARTICLE**).

## ***2. L'existence potentielle de variants « intermédiaires »***

Certains VSI de notre cohorte présentent un taux de survie après exposition à un agent méthylant intermédiaire entre celle des variants connus pathogènes et les variants connus neutres. De tels variants ont également déjà été mis en évidence dans d'autres études, et les auteurs parlent d'« *efficiency polymorphism* » (Ellison *et al.*, 2001), de variants « *partly loss of function* » (Fan *et al.*, 2007), « *low mutator* » (Wanat *et al.*, 2007) ou encore « *moderately functional* » (Arora *et al.*, 2017).

La question est de savoir si ces réponses intermédiaires sont induites par des biais de design de l'expérimentation ou si elles reflètent une réalité biologique.

En ce qui concerne le biais d'expérimentation, nous pouvons par exemple imaginer, dans notre étude, le cas d'un variant hypomorphe, dont le faible taux d'expression en contexte physiologique expliquerait le caractère pathogène, mais qui permettrait une restauration intermédiaire de la fonction dans notre test, due à l'expression très forte des protéines induites par le promoteur du CMV.

Par ailleurs, les conséquences fonctionnelles d'un changement d'acide aminé peuvent en théorie conduire à tous les stades intermédiaires entre une activité normale et une perte totale d'activité de la protéine. Ainsi, une fonctionnalité « intermédiaire » pourrait corrélérer avec un phénotype clinique LS « intermédiaire » caractérisé par moins de cancers au cours de la vie (faible pénétrance) et/ou un âge plus tardif au développement d'une tumeur par exemple (Cyr *et al.*, 2012). Ce concept de « phénotype intermédiaire » est de plus en plus pris en considération mais n'a pas été clairement démontré pour l'instant pour les gènes MMR dans le contexte du LS. Il a par contre été mis

en évidence pour les altérations du gène *TP53* (responsables du syndrome Li-Fraumeni) (Zerdoumi *et al.*, 2013), des gènes *BRCA1/2* (Shimelis *et al.*, 2017) ainsi que du gène *PMS2* dans le contexte du CMMRD (Li *et al.*, 2015).

Cette hypothèse de l'existence de variants avec une fonctionnalité résiduelle remettrait en cause la classification actuelle des variants en 5 classes. En effet, cette classification est basée sur l'hypothèse que les altérations génétiques sont censées être soit pathogènes (classes 5 et 4 : variants pathogènes ou probablement pathogènes) soit neutres (classes 1 et 2 : variants neutres ou probablement neutres). La seule autre possibilité est que la signification du variant soit inconnue (classe 3). Ainsi, il n'y a actuellement pas de notion de variant « intermédiaire ». L'existence possible de tels variants participe à la complexité de la définition clinique et moléculaire du LS.

## D. Perspectives

Nous avons vu que nous mettons en évidence des variants qui présentent une réponse intermédiaire entre celle des variants connus pathogènes et celle des variants connus neutres. En augmentant le nombre de variants dans la série cas-contrôle, il est possible que les résultats obtenus permettent de diminuer la zone d'incertitude entre les variants connus neutres et les variants connus pathogènes et ainsi de diminuer la probabilité d'obtenir des VSI intermédiaires pour lesquels on a du mal à conclure.

L'étude a été réalisée avec les variants *MLH1* analysés dans la lignée de cancer colorectal HCT116 *MLH1*-déficiente et avec les variants *MSH2* dans la lignée LoVo *MSH2*-déficiente. En théorie, l'étude pourrait être complétée par une analyse des variants *MSH6* dans la lignée de CCR HCT15 *MSH6*-déficiente.

Nous avons à plusieurs reprises insisté sur l'expression exogène des protéines MMR à partir d'un plasmide transfecté. Cette limitation n'est de loin pas spécifique de notre test, puisqu'à ma connaissance tous les tests développés à l'heure actuelle et qui font intervenir de la transfection dans des cellules utilisent un promoteur fort. Il serait envisageable de réaliser les mêmes expériences avec un autre promoteur qui permettrait l'expression de la protéine à un taux qui s'approcherait du taux physiologique et ainsi s'affranchir des biais induits par une surexpression. Il serait intéressant de comparer ces résultats avec les résultats déjà acquis pour constater de potentielles différences : il est probable qu'un variant classé compétent avec notre test actuel mais déficient avec ce nouveau test soit en fait un variant hypomorphe.

Une autre limitation de notre test est de ne pas pouvoir analyser le VSI dans le contexte génétique du patient chez qui il a été mis en évidence. Pour pallier à ces deux limitations (expression

exogène et hors contexte génétique du patient), il serait envisageable de réaliser des lignées lymphoblastoïdes à partir des cellules sanguines d'un patient porteur d'un VSI. L'inactivation de l'allèle sauvage par technologie CRISPR-Cas9 permettrait théoriquement d'obtenir des cellules n'exprimant la protéine d'intérêt qu'à partir de l'allèle variant. Le test de méthylation peut ensuite être réalisé directement sur ces cellules, sans nécessité de transfection et de tri cellulaire. L'inconvénient de cette approche est qu'elle permet de poser le diagnostic pour l'individu en question et non pour ses apparentés, car s'ils sont susceptibles d'être porteurs du variant, ils ne partagent pas l'intégralité du contexte génétique.

Par ailleurs, nous avons évoqué l'importance d'étudier à la fois la fonction de signalisation et la fonction de réparation, afin de mettre en évidence toutes les combinaisons possibles (déficience ou compétence pour la signalisation et la réparation ; ou déficience pour une fonction et compétence pour l'autre). Il serait intéressant de mettre en place un tel test, qui pourrait être réalisé en parallèle de notre test de signalisation, comme cela est déjà le cas pour le test diagnostique du syndrome CMMRD.

Les critères de classification des VSI sont en constante évolution et de nouvelles contributions peuvent les compléter à tout moment. On peut citer l'exemple du travail de Shirts, qui montre que les données somatiques peuvent aider à la caractérisation des VSI et qui propose donc de rajouter cet argument dans les critères ACMG (Shirts *et al.*, 2018). De la même façon, il est envisageable de proposer l'ajout de notre test de tolérance à la méthylation dans les critères d'interprétation, afin de les améliorer et d'être en mesure de donner des résultats diagnostiques définitifs et claires à ces individus qui sont actuellement dans l'incertitude de leur condition.

# BIBLIOGRAPHIE

- 1000 Genomes Project Consortium, A. Auton, L. D. Brooks, *et al.* (2015) A Global Reference for Human Genetic Variation. *Nature* 526(7571): 68–74.
- Aaltonen, L. A., P. Peltomäki, F. S. Leach, *et al.* (1993) Clues to the Pathogenesis of Familial Colorectal Cancer. *Science (New York, N.Y.)* 260(5109): 812–816.
- Abbott, P. J., and R. Saffhill (1979) DNA Synthesis with Methylated Poly(DC-DG) Templates. Evidence for a Competitive Nature to Miscoding by O(6)-Methylguanine. *Biochimica Et Biophysica Acta* 562(1): 51–61.
- Adamson, A. W., D. I. Beardsley, W.-J. Kim, *et al.* (2005) Methylator-Induced, Mismatch Repair-Dependent G<sub>2</sub> Arrest Is Activated through Chk1 and Chk2. *Molecular Biology of the Cell* 16(3): 1513–1526.
- Adzhubei, I. A., S. Schmidt, L. Peshkin, *et al.* (2010) A Method and Server for Predicting Damaging Missense Mutations. *Nature Methods* 7(4): 248–249.
- Allan, J. M., and L. B. Travis (2005) Mechanisms of Therapy-Related Carcinogenesis. *Nature Reviews. Cancer* 5(12): 943–955.
- Allay, E., M. Veigl, and S. L. Gerson (1999) Mice Over-Expressing Human O6alkylguanine-DNA Alkyltransferase Selectively Reduce O6methylguanine Mediated Carcinogenic Mutations to Threshold Levels after N-Methyl-N-Nitrosourea. *Oncogene* 18(25): 3783–3787.
- Andersen, Sofie Dabros, Sascha Emilie Liberti, Anne Lützen, *et al.* (2012) Functional Characterization of MLH1 Missense Variants Identified in Lynch Syndrome Patients. *Human Mutation* 33(12): 1647–1655.
- Aquilina, G., and M. Bignami (2001) Mismatch Repair in Correction of Replication Errors and Processing of DNA Damage. *Journal of Cellular Physiology* 187(2): 145–154.
- Aquilina, G., S. Ceccotti, S. Martinelli, R. Hampson, and M. Bignami (1998) N-(2-Chloroethyl)-N'-Cyclohexyl-N-Nitrosourea Sensitivity in Mismatch Repair-Defective Human Cells. *Cancer Research* 58(1): 135–141.
- Arlow, T., K. Scott, A. Wagenseller, and A. Gammie (2013) Proteasome Inhibition Rescues Clinically Significant Unstable Variants of the Mismatch Repair Protein Msh2. *Proceedings of the National Academy of Sciences of the United States of America* 110(1): 246–251.
- Arnold, C. N., A. Goel, and C. R. Boland (2003) Role of HMLH1 Promoter Hypermethylation in Drug Resistance to 5-Fluorouracil in Colorectal Cancer Cell Lines. *International Journal of Cancer* 106(1): 66–73.
- Aronson, M., S. Gallinger, Z. Cohen, *et al.* (2016) Gastrointestinal Findings in the Largest Series of Patients With Hereditary Biallelic Mismatch Repair Deficiency Syndrome: Report from the International Consortium. *The American Journal of Gastroenterology* 111(2): 275–284.
- Arora, S., P. J. Huwe, R. Sikder, *et al.* (2017) Functional Analysis of Rare Variants in Mismatch Repair Proteins Augments Results from Computation-Based Predictive Methods. *Cancer Biology & Therapy* 18(7): 519–533.
- Auclair, J., M.-P. Busine, C. Navarro, *et al.* (2006) Systematic MRNA Analysis for the Effect of MLH1 and MSH2 Missense and Silent Mutations on Aberrant Splicing. *Human Mutation* 27(2): 145–154.
- Auclair, J., D. Leroux, F. Desseigne, *et al.* (2007) Novel Biallelic Mutations in MSH6 and PMS2 Genes: Gene Conversion as a Likely Cause of PMS2 Gene Inactivation. *Human Mutation* 28(11): 1084–1090.
- Aust, D. E., M. Haase, L. Dobryden, *et al.* (2005) Mutations of the BRAF Gene in Ulcerative Colitis-Related Colorectal Carcinoma. *International Journal of Cancer* 115(5): 673–677.

- Avdievich, E., C. Reiss, S. J. Scherer, *et al.* (2008) Distinct Effects of the Recurrent Mlh1G67R Mutation on MMR Functions, Cancer, and Meiosis. *Proceedings of the National Academy of Sciences of the United States of America* 105(11): 4247–4252.
- Bakry, D., M. Aronson, C. Durno, *et al.* (2014) Genetic and Clinical Determinants of Constitutional Mismatch Repair Deficiency Syndrome: Report from the Constitutional Mismatch Repair Deficiency Consortium. *European Journal of Cancer (Oxford, England: 1990)* 50(5): 987–996.
- Bali, V., and Z. Bebok (2015) Decoding Mechanisms by Which Silent Codon Changes Influence Protein Biogenesis and Function. *The International Journal of Biochemistry & Cell Biology* 64: 58–74.
- Balmaña, J., F. Balaguer, A. Cervantes, D. Arnold, and ESMO Guidelines Working Group (2013) Familial Risk-Colorectal Cancer: ESMO Clinical Practice Guidelines. *Annals of Oncology: Official Journal of the European Society for Medical Oncology* 24 Suppl 6: vi73-80.
- Baretti, M., and D. T. Le (2018) DNA Mismatch Repair in Cancer. *Pharmacology & Therapeutics* 189: 45–62.
- Becker, K., J. Dosch, C. M. Gregel, B. A. Martin, and B. Kaina (1996) Targeted Expression of Human O(6)-Methylguanine-DNA Methyltransferase (MGMT) in Transgenic Mice Protects against Tumor Initiation in Two-Stage Skin Carcinogenesis. *Cancer Research* 56(14): 3244–3249.
- Belvederesi, L., F. Bianchi, E. Galizia, *et al.* (2008) MSH2 Missense Mutations and HNPCC Syndrome: Pathogenicity Assessment in a Human Expression System. *Human Mutation* 29(11): E296-309.
- Belvederesi, L., F. Bianchi, C. Loretelli, *et al.* (2006) Assessing the Pathogenicity of MLH1 Missense Mutations in Patients with Suspected Hereditary Nonpolyposis Colorectal Cancer: Correlation with Clinical, Genetic and Functional Features. *European Journal of Human Genetics: EJHG* 14(7): 853–859.
- Béroud, C., G. Collod-Bérout, C. Boileau, T. Soussi, and C. Junien (2000) UMD (Universal Mutation Database): A Generic Software to Build and Analyze Locus-Specific Databases. *Human Mutation* 15(1): 86–94.
- Bignami, M., M. O’Driscoll, G. Aquilina, and P. Karran (2000) Unmasking a Killer: DNA O6-Methylguanine and the Cytotoxicity of Methylating Agents. *Mutation Research/Reviews in Mutation Research* 462(2–3): 71–82.
- Biswas, S., P. Trobridge, J. Romero-Gallo, *et al.* (2008) Mutational Inactivation of TGFBR2 in Microsatellite Unstable Colon Cancer Arises from the Cooperation of Genomic Instability and the Clonal Outgrowth of Transforming Growth Factor Beta Resistant Cells. *Genes, Chromosomes & Cancer* 47(2): 95–106.
- Blasi, Monica Francesca, Ilenia Ventura, Gabriele Aquilina, *et al.* (2006) A Human Cell-Based Assay to Evaluate the Effects of Alterations in the MLH1 Mismatch Repair Gene. *Cancer Research* 66(18): 9036–9044.
- Bodo, S., C. Colas, O. Buhard, *et al.* (2015) Diagnosis of Constitutional Mismatch Repair-Deficiency Syndrome Based on Microsatellite Instability and Lymphocyte Tolerance to Methylating Agents. *Gastroenterology* 149(4): 1017-1029.e3.
- Bodo, S., M. Svrcek, I. Sourrouille, *et al.* (2015) Azathioprine Induction of Tumors with Microsatellite Instability: Risk Evaluation Using a Mouse Model. *Oncotarget* 6(28): 24969–24977.
- Boland, C. R. (2005) Evolution of the Nomenclature for the Hereditary Colorectal Cancer Syndromes. *Familial Cancer* 4(3): 211–218.
- Boland, C. R., and A. Goel (2010) Microsatellite Instability in Colorectal Cancer. *Gastroenterology* 138(6): 2073-2087.e3.

- Boland, C. R., and H. T. Lynch (2013) The History of Lynch Syndrome. *Familial Cancer* 12(2): 145–157.
- Boland, C. R., and M. Shike (2010) Report from the Jerusalem Workshop on Lynch Syndrome-Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology* 138(7): 2197.e1–7.
- Boland, C. R., S. N. Thibodeau, S. R. Hamilton, *et al.* (1998) 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 Research* 58(22): 5248–5257.
- Boland, P. M., M. B. Yurgelun, and C. R. Boland (2018) Recent Progress in Lynch Syndrome and Other Familial Colorectal Cancer Syndromes. *CA: A Cancer Journal for Clinicians* 68(3): 217–231.
- Bonadona, V., B. Bonaïti, S. Olschwang, *et al.* (2011) Cancer Risks Associated with Germline Mutations in MLH1, MSH2, and MSH6 Genes in Lynch Syndrome. *JAMA* 305(22): 2304–2310.
- Borie, C., C. Colas, P. Dartigues, *et al.* (2009) The Mechanisms Underlying MMR Deficiency in Immunodeficiency-Related Non-Hodgkin Lymphomas Are Different from Those in Other Sporadic Microsatellite Instable Neoplasms. *International Journal of Cancer* 125(10): 2360–2366.
- Borràs, E., M. Pineda, A. Brieger, *et al.* (2012) Comprehensive Functional Assessment of MLH1 Variants of Unknown Significance. *Human Mutation* 33(11): 1576–1588.
- Borràs, E., M. Pineda, J. Cadiñanos, *et al.* (2013) Refining the Role of PMS2 in Lynch Syndrome: Germline Mutational Analysis Improved by Comprehensive Assessment of Variants. *Journal of Medical Genetics* 50(8): 552–563.
- Bourdais, R., B. Rousseau, A. Pujals, *et al.* (2017) Polymerase Proofreading Domain Mutations: New Opportunities for Immunotherapy in Hypermutated Colorectal Cancer beyond MMR Deficiency. *Critical Reviews in Oncology/Hematology* 113: 242–248.
- Branch, P., G. Aquilina, M. Bignami, and P. Karran (1993) Defective Mismatch Binding and a Mutator Phenotype in Cells Tolerant to DNA Damage. *Nature* 362(6421): 652–654.
- Branch, P., R. Hampson, and P. Karran (1995) DNA Mismatch Binding Defects, DNA Damage Tolerance, and Mutator Phenotypes in Human Colorectal Carcinoma Cell Lines. *Cancer Research* 55(11): 2304–2309.
- Breivik, J., and G. Gaudernack (1999) Genomic Instability, DNA Methylation, and Natural Selection in Colorectal Carcinogenesis. *Seminars in Cancer Biology* 9(4): 245–254.
- Brentnall, T. A., D. A. Crispin, M. P. Bronner, *et al.* (1996) Microsatellite Instability in Nonneoplastic Mucosa from Patients with Chronic Ulcerative Colitis. *Cancer Research* 56(6): 1237–1240.
- Brieger, A., J. Trojan, J. Raedle, G. Plotz, and S. Zeuzem (2002) Transient Mismatch Repair Gene Transfection for Functional Analysis of Genetic HMLH1 and HMSH2 Variants. *Gut* 51(5): 677–684.
- Brown, K. D., A. Rathi, R. Kamath, *et al.* (2003) The Mismatch Repair System Is Required for S-Phase Checkpoint Activation. *Nature Genetics* 33(1): 80–84.
- Buermeyer, A. B., C. W.-V. Patten, S. M. Baker, and R. M. Liskay (1999) The Human MLH1 CDNA Complements DNA Mismatch Repair Defects in Mlh1-Deficient Mouse Embryonic Fibroblasts: 5.
- Buhard, O., A. Lagrange, A. Guilloux, *et al.* (2016) HSP110 T17 Simplifies and Improves the Microsatellite Instability Testing in Patients with Colorectal Cancer. *Journal of Medical Genetics* 53(6): 377–384.
- Buhard, O., N. Suraweera, A. Lectard, A. Duval, and R. Hamelin (2004) Quasimonomorphic Mononucleotide Repeats for High-Level Microsatellite Instability Analysis. *Disease Markers* 20(4–5): 251–257.

- Burn, J., A.-M. Gerdes, F. Macrae, *et al.* (2011) Long-Term Effect of Aspirin on Cancer Risk in Carriers of Hereditary Colorectal Cancer: An Analysis from the CAPP2 Randomised Controlled Trial. *Lancet* (London, England) 378(9809): 2081–2087.
- Burn, J., J. C. Mathers, and D. T. Bishop (2013) Chemoprevention in Lynch Syndrome. *Familial Cancer* 12(4): 707–718.
- Cahill, D. P., K. K. Levine, R. A. Betensky, *et al.* (2007) Loss of the Mismatch Repair Protein MSH6 in Human Glioblastomas Is Associated with Tumor Progression during Temozolomide Treatment. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 13(7): 2038–2045.
- Capecchi, M. R., N. E. Capecchi, S. H. Hughes, and G. M. Wahl (1974) Selective Degradation of Abnormal Proteins in Mammalian Tissue Culture Cells. *Proceedings of the National Academy of Sciences of the United States of America* 71(12): 4732–4736.
- Capper, D., A. Voigt, G. Bozukova, *et al.* (2013) BRAF V600E-Specific Immunohistochemistry for the Exclusion of Lynch Syndrome in MSI-H Colorectal Cancer. *International Journal of Cancer* 133(7): 1624–1630.
- Carethers, J. M., D. P. Chauhan, D. Fink, *et al.* (1999) Mismatch Repair Proficiency and in Vitro Response to 5-Fluorouracil. *Gastroenterology* 117(1): 123–131.
- Carethers, J M, M T Hawn, D P Chauhan, *et al.* (1996) Competency in Mismatch Repair Prohibits Clonal Expansion of Cancer Cells Treated with N-Methyl-N'-Nitro-N-Nitrosoguanidine. *Journal of Clinical Investigation* 98(1): 199–206.
- Carethers, J. M., and E. M. Stoffel (2015) Lynch Syndrome and Lynch Syndrome Mimics: The Growing Complex Landscape of Hereditary Colon Cancer. *World Journal of Gastroenterology* 21(31): 9253–9261.
- Cawkwell, L., F. Sutherland, H. Murgatroyd, *et al.* (2000) Defective HMSH2/HMLH1 Protein Expression Is Seen Infrequently in Ulcerative Colitis Associated Colorectal Cancers. *Gut* 46(3): 367–369.
- Cejka, P., L. Stojic, N. Mojas, *et al.* (2003) Methylation-Induced G(2)/M Arrest Requires a Full Complement of the Mismatch Repair Protein HMLH1. *The EMBO Journal* 22(9): 2245–2254.
- Chalastanis, A., V. Penard-Lacronique, M. Svrcek, *et al.* (2010) Azathioprine-Induced Carcinogenesis in Mice According to Msh2 Genotype. *Journal of the National Cancer Institute* 102(22): 1731–1740.
- Chan, T. L., S. T. Yuen, C. K. Kong, *et al.* (2006) Heritable Germline Epimutation of MSH2 in a Family with Hereditary Nonpolyposis Colorectal Cancer. *Nature Genetics* 38(10): 1178–1183.
- Chang, D. K., D. Metzgar, C. Wills, and C. R. Boland (2001) Microsatellites in the Eukaryotic DNA Mismatch Repair Genes as Modulators of Evolutionary Mutation Rate. *Genome Research* 11(7): 1145–1146.
- Chang, L., M. Chang, H. M. Chang, and F. Chang (2017) Expanding Role of Microsatellite Instability in Diagnosis and Treatment of Colorectal Cancers. *Journal of Gastrointestinal Cancer* 48(4): 305–313.
- Chao, Elizabeth C., Jonathan L. Velasquez, Mavee S. L. Witherspoon, *et al.* (2008) Accurate Classification of MLH1/MSH2 Missense Variants with Multivariate Analysis of Protein Polymorphisms-Mismatch Repair (MAPP-MMR). *Human Mutation* 29(6): 852–860.
- Chatterjee, S., and J. K. Pal (2009) Role of 5'- and 3'-Untranslated Regions of MRNAs in Human Diseases. *Biology of the Cell* 101(5): 251–262.
- Choi, Y., G. E. Sims, S. Murphy, J. R. Miller, and A. P. Chan (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* 7(10): e46688.

- Christensen, L. L., R. Kariola, M. K. Korhonen, *et al.* (2009) Functional Characterization of Rare Missense Mutations in MLH1 and MSH2 Identified in Danish Colorectal Cancer Patients. *Familial Cancer* 8(4): 489–500.
- Claes, K., J. Vandesompele, B. Poppe, *et al.* (2002) Pathological Splice Mutations Outside the Invariant AG/GT Splice Sites of BRCA1 Exon 5 Increase Alternative Transcript Levels in the 5' End of the BRCA1 Gene. *Oncogene* 21(26): 4171–4175.
- Claij, N., A. van der Wal, M. Dekker, L. Jansen, and H. te Riele (2003) DNA Mismatch Repair Deficiency Stimulates N-Ethyl-N-Nitrosourea-Induced Mutagenesis and Lymphomagenesis. *Cancer Research* 63(9): 2062–2066.
- Colle, R., R. Cohen, D. Cochereau, *et al.* (2017) Immunotherapy and Patients Treated for Cancer with Microsatellite Instability. *Bulletin Du Cancer* 104(1): 42–51.
- Cortes-Ciriano, I., S. Lee, W.-Y. Park, T.-M. Kim, and P. J. Park (2017) A Molecular Portrait of Microsatellite Instability across Multiple Cancers. *Nature Communications* 8: 15180.
- Couch, F. J., L. J. Rasmussen, R. Hofstra, *et al.* (2008) Assessment of Functional Effects of Unclassified Genetic Variants. *Human Mutation* 29(11): 1314–1326.
- Cunningham, D., . Humblet, S. Siena, *et al.* (2004) Cetuximab Monotherapy and Cetuximab plus Irinotecan in Irinotecan-Refractory Metastatic Colorectal Cancer. *The New England Journal of Medicine* 351(4): 337–345.
- Cunningham, J. M., E. R. Christensen, D. J. Tester, *et al.* (1998) Hypermethylation of the HMLH1 Promoter in Colon Cancer with Microsatellite Instability. *Cancer Research* 58(15): 3455–3460.
- Cyr, J. L., G. D. Brown, J. Stroop, and C. D. Heinen (2012) The Predicted Truncation from a Cancer-Associated Variant of the MSH2 Initiation Codon Alters Activity of the MSH2-MSH6 Mismatch Repair Complex. *Molecular Carcinogenesis* 51(8): 647–658.
- D'Atri, S., L. Tentori, P. M. Lacal, *et al.* (1998) Involvement of the Mismatch Repair System in Temozolomide-Induced Apoptosis. *Molecular Pharmacology* 54(2): 334–341.
- De Carlo, E., L. Gerratana, G. De Maglio, *et al.* (2018) Defining a Prognostic Score Based on O6-Methylguanine-DNA Methyltransferase Cut-off Methylation Level Determined by Pyrosequencing in Patients with Glioblastoma Multiforme. *Journal of Neuro-Oncology*. <http://link.springer.com/10.1007/s11060-018-2981-7>, accessed September 3, 2018.
- De Jong, A. E., H. Morreau, M. Van Puijenbroek, *et al.* (2004) The Role of Mismatch Repair Gene Defects in the Development of Adenomas in Patients with HNPCC. *Gastroenterology* 126(1): 42–48.
- De Vos, M., B. E. Hayward, S. Picton, E. Sheridan, and D. T. Bonthron (2004) Novel PMS2 Pseudogenes Can Conceal Recessive Mutations Causing a Distinctive Childhood Cancer Syndrome. *American Journal of Human Genetics* 74(5): 954–964.
- Dilruba, S., and G. V. Kalayda (2016) Platinum-Based Drugs: Past, Present and Future. *Cancer Chemotherapy and Pharmacology* 77(6): 1103–1124.
- Dohet, C., R. Wagner, and M. Radman (1985) Repair of Defined Single Base-Pair Mismatches in *Escherichia Coli*. *Proceedings of the National Academy of Sciences of the United States of America* 82(2): 503–505.
- Dolan, M, and E Pegg (1997) And Its Role in Chemotherapy. *Clinical Cancer Research*: 12.
- Domingo, E., L. Freeman-Mills, E. Rayner, *et al.* (2016) Somatic POLE Proofreading Domain Mutation, Immune Response, and Prognosis in Colorectal Cancer: A Retrospective, Pooled Biomarker Study. *The Lancet. Gastroenterology & Hepatology* 1(3): 207–216.

- Drost, M., A. Lützen, S. van Hees, *et al.* (2013) Genetic Screens to Identify Pathogenic Gene Variants in the Common Cancer Predisposition Lynch Syndrome. *Proceedings of the National Academy of Sciences of the United States of America* 110(23): 9403–9408.
- Drost, M., J. B. M. Zonneveld, L. van Dijk, *et al.* (2010) A Cell-Free Assay for the Functional Analysis of Variants of the Mismatch Repair Protein MLH1. *Human Mutation* 31(3): 247–253.
- Drotschmann, K., A. B. Clark, and T. A. Kunkel (1999) Mutator Phenotypes of Common Polymorphisms and Missense Mutations in MSH2. *Current Biology: CB* 9(16): 907–910.
- Drotschmann, K., A. B. Clark, H. T. Tran, *et al.* (1999) Mutator Phenotypes of Yeast Strains Heterozygous for Mutations in the MSH2 Gene. *Proceedings of the National Academy of Sciences of the United States of America* 96(6): 2970–2975.
- Drotschmann, K., R. P. Topping, J. E. Clodfelter, and F. R. Salsbury (2004) Mutations in the Nucleotide-Binding Domain of MutS Homologs Uncouple Cell Death from Cell Survival. *DNA Repair* 3(7): 729–742.
- Drummond, J. T., A. Anthoney, R. Brown, and P. Modrich (1996) Cisplatin and Adriamycin Resistance Are Associated with MutL<sub>α</sub> and Mismatch Repair Deficiency in an Ovarian Tumor Cell Line: 5.
- Drummond, J. T., J. Genschel, E. Wolf, and P. Modrich (1997) DHFR/MSH3 Amplification in Methotrexate-Resistant Cells Alters the HMutS<sub>α</sub>/HMutS<sub>β</sub> Ratio and Reduces the Efficiency of Base-Base Mismatch Repair. *Proceedings of the National Academy of Sciences of the United States of America* 94(19): 10144–10149.
- Duckett, D. R., S. M. Bronstein, Y. Taya, and P. Modrich (1999) HMutS<sub>α</sub>- and HMutL<sub>α</sub>-Dependent Phosphorylation of P53 in Response to DNA Methylator Damage: 5.
- Dumenco, L. L., E. Allay, K. Norton, and S. L. Gersont (1993) The Prevention of Thymic Lymphomas in Transgenic Mice by Human O6-Alkylguanine-DNA Alkyltransferase 259: 5.
- Durno, C. A., P. M. Sherman, M. Aronson, *et al.* (2015) Phenotypic and Genotypic Characterisation of Biallelic Mismatch Repair Deficiency (BMMR-D) Syndrome. *European Journal of Cancer (Oxford, England: 1990)* 51(8): 977–983.
- Durno, C., C. R. Boland, S. Cohen, *et al.* (2017) Recommendations on Surveillance and Management of Biallelic Mismatch Repair Deficiency (BMMRD) Syndrome: A Consensus Statement by the US Multi-Society Task Force on Colorectal Cancer. *Gastroenterology* 152(6): 1605–1614.
- Duval, A., and R. Hamelin (2002a) Mutations at Coding Repeat Sequences in Mismatch Repair-Deficient Human Cancers: Toward a New Concept of Target Genes for Instability. *Cancer Research* 62(9): 2447–2454.
- Duval, A., and R. Hamelin (2002b) Genetic Instability in Human Mismatch Repair Deficient Cancers. *Annales De Genetique* 45(2): 71–75.
- Duval, A., M. Raphael, C. Brennetot, *et al.* (2004) The Mutator Pathway Is a Feature of Immunodeficiency-Related Lymphomas. *Proceedings of the National Academy of Sciences of the United States of America* 101(14): 5002–5007.
- El-Bchiri, J., A. Guilloux, P. Dartigues, *et al.* (2008) Nonsense-Mediated mRNA Decay Impacts MSI-Driven Carcinogenesis and Anti-Tumor Immunity in Colorectal Cancers. *PLoS One* 3(7): e2583.
- Ellegren, H. (2004) Microsatellites: Simple Sequences with Complex Evolution. *Nature Reviews. Genetics* 5(6): 435–445.
- Ellison, A. R., J. Lofing, and G. A. Bitter (2001) Functional Analysis of Human MLH1 and MSH2 Missense Variants and Hybrid Human-Yeast MLH1 Proteins in *Saccharomyces Cerevisiae*. *Human Molecular Genetics* 10(18): 1889–1900.

- Ellison, A. R., J. Lofing, and G. A. Bitter (2004) Human MutL Homolog (MLH1) Function in DNA Mismatch Repair: A Prospective Screen for Missense Mutations in the ATPase Domain. *Nucleic Acids Research* 32(18): 5321–5338.
- Espenschied, C. R., H. LaDuca, S. Li, *et al.* (2017) Multigene Panel Testing Provides a New Perspective on Lynch Syndrome. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 35(22): 2568–2575.
- Esteller, M., S. N. Goodman, and J. G. Herman (2000) Inactivation of the DNA-Repair Gene MGMT and the Clinical Response of Gliomas to Alkylating Agents. *The New England Journal of Medicine*: 5.
- Evensen, N. A., P. P. Madhusoodhan, J. Meyer, *et al.* (2018) MSH6 Haploinsufficiency at Relapse Contributes to the Development of Thiopurine Resistance in Pediatric B-Lymphoblastic Leukemia. *Haematologica* 103(5): 830–839.
- Fabrizio, D. A., T. J. George, R. F. Dunne, *et al.* (2018) Beyond Microsatellite Testing: Assessment of Tumor Mutational Burden Identifies Subsets of Colorectal Cancer Who May Respond to Immune Checkpoint Inhibition. *Journal of Gastrointestinal Oncology* 9(4): 610–617.
- Fahrer, J., J. Frisch, G. Nagel, *et al.* (2015) DNA Repair by MGMT, but Not AAG, Causes a Threshold in Alkylation-Induced Colorectal Carcinogenesis. *Carcinogenesis* 36(10): 1235–1244.
- Fan, Y., W. Wang, M. Zhu, *et al.* (2007) Analysis of HMLH1 Missense Mutations in East Asian Patients with Suspected Hereditary Nonpolyposis Colorectal Cancer. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 13(24): 7515–7521.
- Fang, W. H., and P. Modrich (1993) Human Strand-Specific Mismatch Repair Occurs by a Bidirectional Mechanism Similar to That of the Bacterial Reaction. *The Journal of Biological Chemistry* 268(16): 11838–11844.
- Fearon, E. R. (2011) Molecular Genetics of Colorectal Cancer. *Annual Review of Pathology* 6: 479–507.
- Fearon, E. R., and B. Vogelstein (1990) A Genetic Model for Colorectal Tumorigenesis. *Cell* 61(5): 759–767.
- Fedier, A., and D. Fink (2004) Mutations in DNA Mismatch Repair Genes: Implications for DNA Damage Signaling and Drug Sensitivity (Review). *International Journal of Oncology* 24(4): 1039–1047.
- Fink, D., S. Aebi, and S. B. Howell (1998) The Role of DNA Mismatch Repair in Drug Resistance. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 4(1): 1–6.
- Fishel, R. (2001) The Selection for Mismatch Repair Defects in Hereditary Nonpolyposis Colorectal Cancer: Revising the Mutator Hypothesis. *Cancer Research* 61(20): 7369–7374.
- Fishel, R., M. K. Lescoe, M. R. Rao, *et al.* (1993) The Human Mutator Gene Homolog MSH2 and Its Association with Hereditary Nonpolyposis Colon Cancer. *Cell* 75(5): 1027–1038.
- Fleisher, A. S., M. Esteller, N. Harpaz, *et al.* (2000) Microsatellite Instability in Inflammatory Bowel Disease-Associated Neoplastic Lesions Is Associated with Hypermethylation and Diminished Expression of the DNA Mismatch Repair Gene, HMLH1. *Cancer Research* 60(17): 4864–4868.
- Fokkema, Ivo F. A. C., Peter E. M. Taschner, Gerard C. P. Schaafsma, *et al.* (2011) LOVD v.2.0: The next Generation in Gene Variant Databases. *Human Mutation* 32(5): 557–563.
- Fortune, J. M., Y. I. Pavlov, C. M. Welch, *et al.* (2005) *Saccharomyces Cerevisiae* DNA Polymerase Delta: High Fidelity for Base Substitutions but Lower Fidelity for Single- and Multi-Base Deletions. *The Journal of Biological Chemistry* 280(33): 29980–29987.
- Friedman, H. S., T. Kerby, and H. Calvert (2000) Temozolomide and Treatment of Malignant Glioma: 14.

- Fu, D., J. A. Calvo, and L. D. Samson (2012) Balancing Repair and Tolerance of DNA Damage Caused by Alkylating Agents. *Nature Reviews Cancer* 12(2): 104–120.
- Gammie, A. E., N. Erdeniz, J. Beaver, *et al.* (2007) Functional Characterization of Pathogenic Human MSH2 Missense Mutations in *Saccharomyces Cerevisiae*. *Genetics* 177(2): 707–721.
- Ganai, R. A., and E. Johansson (2016) DNA Replication-A Matter of Fidelity. *Molecular Cell* 62(5): 745–755.
- Gazzoli, I., M. Loda, J. Garber, S. Syngal, and R. D. Kolodner (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 Research* 62(14): 3925–3928.
- Geng, H., M. Sakato, V. DeRocco, *et al.* (2012) Biochemical Analysis of the Human Mismatch Repair Proteins HMutS $\alpha$  MSH2(G674A)-MSH6 and MSH2-MSH6(T1219D). *The Journal of Biological Chemistry* 287(13): 9777–9791.
- Genschel, J., L. R. Bazemore, and P. Modrich (2002) Human Exonuclease I Is Required for 5' and 3' Mismatch Repair. *The Journal of Biological Chemistry* 277(15): 13302–13311.
- Geurts-Giele, W. R. R., C. H. M. Leenen, H. J. Dubbink, *et al.* (2014) Somatic Aberrations of Mismatch Repair Genes as a Cause of Microsatellite-Unstable Cancers. *The Journal of Pathology* 234(4): 548–559.
- Ghodgaonkar, M. M., F. Lazzaro, M. Olivera-Pimentel, *et al.* (2013) Ribonucleotides Misincorporated into DNA Act as Strand-Discrimination Signals in Eukaryotic Mismatch Repair. *Molecular Cell* 50(3): 323–332.
- Giannini, G., E. Ristori, F. Cerignoli, *et al.* (2002) Human MRE11 Is Inactivated in Mismatch Repair-Deficient Cancers. *EMBO Reports* 3(3): 248–254.
- Giardiello, F. M., J. I. Allen, J. E. Axilbund, *et al.* (2014) Guidelines on Genetic Evaluation and Management of Lynch Syndrome: A Consensus Statement by the US Multi-Society Task Force on Colorectal Cancer. *The American Journal of Gastroenterology* 109(8): 1159–1179.
- Glassner, B. J., G. Weeda, J. M. Allan, *et al.* (1999) DNA Repair Methyltransferase (Mgmt) Knockout Mice Are Sensitive to the Lethal Effects of Chemotherapeutic Alkylating Agents. *Mutagenesis* 14(3): 339–347.
- Goldgar, D. E., D. F. Easton, G. B. Byrnes, *et al.* (2008) Genetic Evidence and Integration of Various Data Sources for Classifying Uncertain Variants into a Single Model. *Human Mutation* 29(11): 1265–1272.
- Goldstein, J., B. Tran, J. Ensor, *et al.* (2014) Multicenter Retrospective Analysis of Metastatic Colorectal Cancer (CRC) with High-Level Microsatellite Instability (MSI-H). *Annals of Oncology: Official Journal of the European Society for Medical Oncology* 25(5): 1032–1038.
- Greenon, J. K., J. D. Bonner, O. Ben-Yzhak, *et al.* (2003) Phenotype of Microsatellite Unstable Colorectal Carcinomas: Well-Differentiated and Focally Mucinous Tumors and the Absence of Dirty Necrosis Correlate with Microsatellite Instability. *The American Journal of Surgical Pathology* 27(5): 563–570.
- Gryfe, R., H. Kim, E. T. Hsieh, *et al.* (2000) Tumor Microsatellite Instability and Clinical Outcome in Young Patients with Colorectal Cancer. *The New England Journal of Medicine* 342(2): 69–77.
- Guerrette, S., S. Acharya, and R. Fishel (1999) The Interaction of the Human MutL Homologues in Hereditary Nonpolyposis Colon Cancer. *The Journal of Biological Chemistry* 274(10): 6336–6341.

- Guinney, J., R. Dienstmann, X. Wang, *et al.* (2015) The Consensus Molecular Subtypes of Colorectal Cancer. *Nature Medicine* 21(11): 1350–1356.
- Hamilton, S. R., B. Liu, R. E. Parsons, *et al.* (1995) The Molecular Basis of Turcot's Syndrome. *The New England Journal of Medicine* 332(13): 839–847.
- Hammond, W. A., A. Swaika, and K. Mody (2016) Pharmacologic Resistance in Colorectal Cancer: A Review. *Therapeutic Advances in Medical Oncology* 8(1): 57–84.
- Hampel, H., W. L. Frankel, E. Martin, *et al.* (2005) Screening for the Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer). *The New England Journal of Medicine* 352(18): 1851–1860.
- Hampel, H., W. Frankel, J. Panescu, *et al.* (2006) Screening for Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) among Endometrial Cancer Patients. *Cancer Research* 66(15): 7810–7817.
- Hanahan, D., and R. A. Weinberg (2011) Hallmarks of Cancer: The next Generation. *Cell* 144(5): 646–674.
- Haraldsdottir, S., H. Hampel, J. Tomsic, *et al.* (2014) Colon and Endometrial Cancers with Mismatch Repair Deficiency Can Arise from Somatic, Rather than Germline, Mutations. *Gastroenterology* 147(6): 1308-1316.e1.
- Hardt, K., S. B. Heick, B. Betz, *et al.* (2011) Missense Variants in HMLH1 Identified in Patients from the German HNPCC Consortium and Functional Studies. *Familial Cancer* 10(2): 273–284.
- Hause, R. J., C. C. Pritchard, J. Shendure, and S. J. Salipante (2016) Classification and Characterization of Microsatellite Instability across 18 Cancer Types. *Nature Medicine* 22(11): 1342–1350.
- Hawn, M. T., A. Umar, J. M. Carethers, *et al.* (1995) Evidence for a Connection between the Mismatch Repair System and the G2 Cell Cycle Checkpoint: 6.
- Hegi, M. E., M.-F. Hamou, N. de Tribolet, *et al.* (2005) MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. *The New England Journal of Medicine*: 7.
- Heidelberger, C., N. K. Chaudhuri, P. Danneberg, *et al.* (1957) Fluorinated Pyrimidines, a New Class of Tumour-Inhibitory Compounds. *Nature* 179(4561): 663–666.
- Heinen, C. D., and L. Juel Rasmussen (2012) Determining the Functional Significance of Mismatch Repair Gene Missense Variants Using Biochemical and Cellular Assays. *Hereditary Cancer in Clinical Practice* 10(1): 9.
- Heinen, C. D., C. Schmutte, and Richard Fishel (2002) DNA Repair and Tumorigenesis: Lessons from Hereditary Cancer Syndromes. *Cancer Biology & Therapy* 1(5): 477–485.
- Heinen, C. D., T. Wilson, A. Mazurek, *et al.* (2002) HNPCC Mutations in HMSH2 Result in Reduced HMSH2-HMSH6 Molecular Switch Functions. *Cancer Cell* 1(5): 469–478.
- Hemminki, A., P. Peltomäki, J. P. Mecklin, *et al.* (1994) Loss of the Wild Type MLH1 Gene Is a Feature of Hereditary Nonpolyposis Colorectal Cancer. *Nature Genetics* 8(4): 405–410.
- Herkert, J. C., R. C. Niessen, M. J. W. Olderde-Berends, *et al.* (2011) Paediatric Intestinal Cancer and Polyposis Due to Bi-Allelic PMS2 Mutations: Case Series, Review and Follow-up Guidelines. *European Journal of Cancer (Oxford, England: 1990)* 47(7): 965–982.
- Hickman, M. J., and L. D. Samson (1999) Role of DNA Mismatch Repair and P53 in Signaling Induction of Apoptosis by Alkylating Agents. *Proceedings of the National Academy of Sciences* 96(19): 10764–10769.
- Hinrichsen, I., A. Brieger, J. Trojan, *et al.* (2013) Expression Defect Size among Unclassified MLH1 Variants Determines Pathogenicity in Lynch Syndrome Diagnosis. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 19(9): 2432–2441.

- Hinrichsen, I., D. Schäfer, D. Langer, *et al.* (2015) Functional Testing Strategy for Coding Genetic Variants of Unclear Significance in MLH1 in Lynch Syndrome Diagnosis. *Carcinogenesis* 36(2): 202–211.
- Hitchins, M. P., and R. L. Ward (2009) Constitutional (Germline) MLH1 Epimutation as an Aetiological Mechanism for Hereditary Non-Polyposis Colorectal Cancer. *Journal of Medical Genetics* 46(12): 793–802.
- Houdayer, C., V. Caux-Moncoutier, S. Krieger, *et al.* (2012) Guidelines for Splicing Analysis in Molecular Diagnosis Derived from a Set of 327 Combined *In Silico/In Vitro* Studies on BRCA1 and BRCA2 Variants. *Human Mutation* 33(8): 1228–1238.
- Houllberg, H., M. Dekker, H. Lantermans, *et al.* (2016) Oligonucleotide-Directed Mutagenesis Screen to Identify Pathogenic Lynch Syndrome-Associated MSH2 DNA Mismatch Repair Gene Variants. *Proceedings of the National Academy of Sciences of the United States of America* 113(15): 4128–4133.
- Houllberg, H., A. Goverde, J. Lusseveld, *et al.* (2017) Suspected Lynch Syndrome Associated MSH6 Variants: A Functional Assay to Determine Their Pathogenicity. *PLoS Genetics* 13(5): e1006765.
- Hsieh, P., and K. Yamane (2008) DNA Mismatch Repair: Molecular Mechanism, Cancer, and Ageing. *Mechanisms of Ageing and Development* 129(7–8): 391–407.
- Hunter, C., R. Smith, D. P. Cahill, *et al.* (2006) A Hypermutation Phenotype and Somatic MSH6 Mutations in Recurrent Human Malignant Gliomas after Alkylator Chemotherapy. *Cancer Research* 66(8): 3987–3991.
- Ingham, D., C. P. Diggle, I. Berry, *et al.* (2013) Simple Detection of Germline Microsatellite Instability for Diagnosis of Constitutional Mismatch Repair Cancer Syndrome. *Human Mutation* 34(6): 847–852.
- Ionov, Y., M. A. Peinado, S. Malkhosyan, D. Shibata, and M. Perucho (1993) Ubiquitous Somatic Mutations in Simple Repeated Sequences Reveal a New Mechanism for Colonic Carcinogenesis. *Nature* 363(6429): 558–561.
- Ishizaki, K., T. Tsujimura, H. Yawata, *et al.* (1986) Transfer of the E. Coli O6-Methylguanine Methyltransferase Gene into Repair-Deficient Human Cells and Restoration of Cellular Resistance to N-Methyl-N'-Nitro-N-Nitrosoguanidine. *Mutation Research* 166(2): 135–141.
- Jacob, S., and F. Praz (2002) DNA Mismatch Repair Defects: Role in Colorectal Carcinogenesis. *Biochimie* 84(1): 27–47.
- Järvinen, H. J., M. Aarnio, H. Mustonen, *et al.* (2000) Controlled 15-Year Trial on Screening for Colorectal Cancer in Families with Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology* 118(5): 829–834.
- Jasperson, K. W., T. M. Tuohy, D. W. Neklason, and R. Burt (2010) Hereditary and Familial Colon Cancer. *Gastroenterology* 138(6): 2044–2058.
- Jiricny, J. (2006) The Multifaceted Mismatch-Repair System. *Nature Reviews. Molecular Cell Biology* 7(5): 335–346.
- Jonchere, V., L. Marisa, M. Greene, *et al.* (2018) Identification of Positively and Negatively Selected Driver Gene Mutations Associated With Colorectal Cancer With Microsatellite Instability. *Cellular and Molecular Gastroenterology and Hepatology* 6(3): 277–300.
- Jun, S.-H., T. G. Kim, and C. Ban (2006) DNA Mismatch Repair System. Classical and Fresh Roles. *The FEBS Journal* 273(8): 1609–1619.

- Kabbinavar, F., H. I. Hurwitz, L. Fehrenbacher, *et al.* (2003) Phase II, Randomized Trial Comparing Bevacizumab plus Fluorouracil (FU)/Leucovorin (LV) with FU/LV Alone in Patients with Metastatic Colorectal Cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 21(1): 60–65.
- Kadyrov, F. A., L. Dzantiev, N. Constantin, and P. Modrich (2006) Endonucleolytic Function of MutL $\alpha$  in Human Mismatch Repair. *Cell* 126(2): 297–308.
- Kaina, B., G. Fritz, S. Mitra, and T. Coquerelle (1991) Transfection and Expression of Human O6-Methylguanine-DNA Methyltransferase (MGMT) cDNA in Chinese Hamster Cells: The Role of MGMT in Protection against the Genotoxic Effects of Alkylating Agents. *Carcinogenesis* 12(10): 1857–1867.
- Kaina, B., A. A. van Zeeland, A. de Groot, and A. T. Natarajan (1990) DNA Repair and Chromosomal Stability in the Alkylating Agent-Hypersensitive Chinese Hamster Cell Line 27-1. *Mutation Research* 243(3): 219–224.
- Kaina, B., A. Ziouta, K. Ochs, and T. Coquerelle (1997) Chromosomal Instability, Reproductive Cell Death and Apoptosis Induced by O6-Methylguanine in Mex<sup>-</sup>, Mex<sup>+</sup> and Methylation-Tolerant Mismatch Repair Compromised Cells: Facts and Models. *Mutation Research* 381(2): 227–241.
- Kakar, S., L. J. Burgart, S. N. Thibodeau, *et al.* (2003) Frequency of Loss of HMLH1 Expression in Colorectal Carcinoma Increases with Advancing Age. *Cancer* 97(6): 1421–1427.
- Kane, M. F., M. Loda, G. M. Gaida, *et al.* (1997) Methylation of the HMLH1 Promoter Correlates with Lack of Expression of HMLH1 in Sporadic Colon Tumors and Mismatch Repair-Defective Human Tumor Cell Lines. *Cancer Research* 57(5): 808–811.
- Kantelinen, J., T. V. O. Hansen, M. Kansikas, *et al.* (2011) A Putative Lynch Syndrome Family Carrying MSH2 and MSH6 Variants of Uncertain Significance-Functional Analysis Reveals the Pathogenic One. *Familial Cancer* 10(3): 515–520.
- Kantelinen, J., M. Kansikas, S. Candelin, *et al.* (2012) Mismatch Repair Analysis of Inherited MSH2 and/or MSH6 Variation Pairs Found in Cancer Patients. *Human Mutation* 33(8): 1294–1301.
- Kariola, R., H. Hampel, W. L. Frankel, *et al.* (2004) MSH6 Missense Mutations Are Often Associated with No or Low Cancer Susceptibility. *British Journal of Cancer* 91(7): 1287–1292.
- Kariola, R., R. Otway, K. E. Lönnqvist, *et al.* (2003) Two Mismatch Repair Gene Mutations Found in a Colon Cancer Patient--Which One Is Pathogenic? *Human Genetics* 112(2): 105–109.
- Kariola, R., T. Raevaara, K. E. Lönnqvist, and M. Nyström-Lahti (2002) Functional Analysis of MSH6 Mutations Linked to Kindreds with Putative Hereditary Non-Polyposis Colorectal Cancer Syndrome. *Human Molecular Genetics* 11(11): 1303–1310.
- Karran, P., and N. Attard (2008) Thiopurines in Current Medical Practice: Molecular Mechanisms and Contributions to Therapy-Related Cancer. *Nature Reviews. Cancer* 8(1): 24–36.
- Karran, P., and M. Bignami (1992) Self-Destruction and Tolerance in Resistance of Mammalian Cells to Alkylation Damage. *Nucleic Acids Research* 20(12): 2933–2940.
- Karran, P., and M. G. Marinus (1982) Mismatch Correction at O6-Methylguanine Residues in *E. Coli* DNA. *Nature* 296(5860): 868–869.
- Karran, P., Offman, J., and M. Bignami (2003) Human Mismatch Repair, Drug-Induced DNA Damage, and Secondary Cancer. *Biochimie* 85(11): 1149–1160.
- Karran, P., and C. Stephenson (1990) Mismatch Binding Proteins and Tolerance to Alkylating Agents in Human Cells. *Mutation Research/DNA Repair* 236(2–3): 269–275.

- Kat, A., W. G. Thilly, W. H. Fang, *et al.* (1993) An Alkylation-Tolerant, Mutator Human Cell Line Is Deficient in Strand-Specific Mismatch Repair. *Proceedings of the National Academy of Sciences* 90(14): 6424–6428.
- Kataoka, H, J Hall, and P Karran (1986) Complementation of Sensitivity to Alkylating Agents in *Escherichia Coli* and Chinese Hamster Ovary Cells by Expression of a Cloned Bacterial DNA Repair Gene: 6.
- Killock, D. (2017) Immunotherapy: PD-1 Blockade Exploiting MMR Deficiency. *Nature Reviews. Clinical Oncology* 14(8): 459.
- Knudson, A. G. (1971) Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* 68(4): 820–823.
- Koi, M., A. Umar, D. P. Chauhan, *et al.* (1994) Human Chromosome 3 Corrects Mismatch Repair Deficiency and Microsatellite Instability and Reduces N-Methyl-N'-Nitro-N-Nitrosoguanidine Tolerance in Colon Tumor Cells with Homozygous HMLH1 Mutation. *Cancer Research* 54(16): 4308–4312.
- Kondo, E., H. Suzuki, A. Horii, and S. Fukushige (2003) A Yeast Two-Hybrid Assay Provides a Simple Way to Evaluate the Vast Majority of HMLH1 Germ-Line Mutations. *Cancer Research* 63(12): 3302–3308.
- Kosinski, J., I. Hinrichsen, J. M. Bujnicki, P. Friedhoff, and G. Plotz (2010) Identification of Lynch Syndrome Mutations in the MLH1-PMS2 Interface That Disturb Dimerization and Mismatch Repair. *Human Mutation* 31(8): 975–982.
- Kroutil, L. C., K. Register, K. Bebenek, and T. A. Kunkel (1996) Exonucleolytic Proofreading during Replication of Repetitive DNA. *Biochemistry* 35(3): 1046–1053.
- Kuismanen, S. A., M. T. Holmberg, R. Salovaara, A. de la Chapelle, and P. Peltomäki (2000) Genetic and Epigenetic Modification of MLH1 Accounts for a Major Share of Microsatellite-Unstable Colorectal Cancers. *The American Journal of Pathology* 156(5): 1773–1779.
- Kuismanen, S. A., A.-L. Moisio, P. Schweizer, *et al.* (2002) Endometrial and Colorectal Tumors from Patients with Hereditary Nonpolyposis Colon Cancer Display Different Patterns of Microsatellite Instability. *The American Journal of Pathology* 160(6): 1953–1958.
- Kunkel, T. A., and D. A. Erie (2015) Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annual Review of Genetics* 49: 291–313.
- Lahue, R., K. Au, and P. Modrich (1989) DNA Mismatch Correction in a Defined System. *Science* 245(4914): 160–164.
- Landrum, M. J., J. M. Lee, M. Benson, *et al.* (2016) ClinVar: Public Archive of Interpretations of Clinically Relevant Variants. *Nucleic Acids Research* 44(D1): D862–868.
- Lavoine, N., C. Colas, M. Muleris, *et al.* (2015) Constitutional Mismatch Repair Deficiency Syndrome: Clinical Description in a French Cohort. *Journal of Medical Genetics* 52(11): 770–778.
- Le, D. T., J. N. Uram, H. Wang, *et al.* (2015) PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *The New England Journal of Medicine* 372(26): 2509–2520.
- Leach, F. S., N. C. Nicolaides, N. Papadopoulos, *et al.* (1993) Mutations of a MutS Homolog in Hereditary Nonpolyposis Colorectal Cancer. *Cell* 75(6): 1215–1225.
- Lee, J.-B., W.-K. Cho, J. Park, *et al.* (2014) Single-Molecule Views of MutS on Mismatched DNA. *DNA Repair* 20: 82–93.
- Lek, M., K. J. Karczewski, E. V. Minikel, *et al.* (2016) Analysis of Protein-Coding Genetic Variation in 60,706 Humans. *Nature* 536(7616): 285–291.

- Lettieri, T., G. Marra, G. Aquilina, *et al.* (1999) Effect of HSMH6 CDNA Expression on the Phenotype of Mismatch Repair-Deficient Colon Cancer Cell Line HCT15. *Carcinogenesis* 20(3): 373–382.
- Lewandowska, M. A. (2013) The Missing Puzzle Piece: Splicing Mutations. *International Journal of Clinical and Experimental Pathology* 6(12): 2675–2682.
- Li, B., V. G. Krishnan, M. E. Mort, *et al.* (2009) Automated Inference of Molecular Mechanisms of Disease from Amino Acid Substitutions. *Bioinformatics (Oxford, England)* 25(21): 2744–2750.
- Li, F., G. Mao, D. Tong, *et al.* (2013) The Histone Mark H3K36me3 Regulates Human DNA Mismatch Repair through Its Interaction with MutS $\alpha$ . *Cell* 153(3): 590–600.
- Li, G.-M. (2008) Mechanisms and Functions of DNA Mismatch Repair. *Cell Research* 18(1): 85–98.
- Li, L., N. Hamel, K. Baker, *et al.* (2015) A Homozygous PMS2 Founder Mutation with an Attenuated Constitutional Mismatch Repair Deficiency Phenotype. *Journal of Medical Genetics* 52(5): 348–352.
- Li, S. K. H., and A. Martin (2016) Mismatch Repair and Colon Cancer: Mechanisms and Therapies Explored. *Trends in Molecular Medicine* 22(4): 274–289.
- Ligtenberg, M. J. L., R. P. Kuiper, T. L. Chan, *et al.* (2009) Heritable Somatic Methylation and Inactivation of MSH2 in Families with Lynch Syndrome Due to Deletion of the 3' Exons of TACSTD1. *Nature Genetics* 41(1): 112–117.
- Lin, D. P., Y. Wang, S. J. Scherer, *et al.* (2004) An Msh2 Point Mutation Uncouples DNA Mismatch Repair and Apoptosis. *Cancer Research* 64(2): 517–522.
- Lin, K. M., M. Shashidharan, A. G. Thorson, *et al.* (1998) Cumulative Incidence of Colorectal and Extracolonic Cancers in MLH1 and MSH2 Mutation Carriers of Hereditary Nonpolyposis Colorectal Cancer. *Journal of Gastrointestinal Surgery: Official Journal of the Society for Surgery of the Alimentary Tract* 2(1): 67–71.
- Lindblom, A., P. Tannergård, B. Werelius, and M. Nordenskjöld (1993) Genetic Mapping of a Second Locus Predisposing to Hereditary Non-Polyposis Colon Cancer. *Nature Genetics* 5(3): 279–282.
- Lindor, N. M., L. Guidugli, X. Wang, *et al.* (2012) A Review of a Multifactorial Probability-Based Model for Classification of BRCA1 and BRCA2 Variants of Uncertain Significance (VUS). *Human Mutation* 33(1): 8–21.
- Linnebacher, M., J. Gebert, W. Rudy, *et al.* (2001) Frameshift Peptide-Derived T-Cell Epitopes: A Source of Novel Tumor-Specific Antigens. *International Journal of Cancer* 93(1): 6–11.
- Longley, D. B., D. P. Harkin, and P. G. Johnston (2003) 5-Fluorouracil: Mechanisms of Action and Clinical Strategies. *Nature Reviews. Cancer* 3(5): 330–338.
- Longley, M. J., A. J. Pierce, and P. Modrich (1997) DNA Polymerase Delta Is Required for Human Mismatch Repair in Vitro. *The Journal of Biological Chemistry* 272(16): 10917–10921.
- López de Saro, F. J., and M. O'Donnell (2001) Interaction of the Beta Sliding Clamp with MutS, Ligase, and DNA Polymerase I. *Proceedings of the National Academy of Sciences of the United States of America* 98(15): 8376–8380.
- Lujan, S. A., J. S. Williams, Z. F. Pursell, *et al.* (2012) Mismatch Repair Balances Leading and Lagging Strand DNA Replication Fidelity. Christopher E. Pearson, ed. *PLoS Genetics* 8(10): e1003016.
- Lukash, L. L., J. Boldt, A. E. Pegg, *et al.* (1991) Effect of O6-Alkylguanine-DNA Alkyltransferase on the Frequency and Spectrum of Mutations Induced by N-Methyl-N'-Nitro-N-Nitrosoguanidine in the HPRT Gene of Diploid Human Fibroblasts. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 250(1–2): 397–409.

- Lützen, A., N. de Wind, D. Georgijevic, F. C. Nielsen, and L. J. Rasmussen (2008) Functional Analysis of HNPCC-Related Missense Mutations in MSH2. *Mutation Research* 645(1–2): 44–55.
- Lyda, M. H., A. Noffsinger, J. Belli, and C. M. Fenoglio-Preiser (2000) Microsatellite Instability and K-Ras Mutations in Patients with Ulcerative Colitis. *Human Pathology* 31(6): 665–671.
- Lynch, H. T., S. Lanspa, T. Shaw, *et al.* (2018) Phenotypic and Genotypic Heterogeneity of Lynch Syndrome: A Complex Diagnostic Challenge. *Familial Cancer* 17(3): 403–414.
- Lynch, H. T., M. W. Shaw, C. W. Magnuson, A. L. Larsen, and A. J. Krush (1966) Hereditary Factors in Cancer. Study of Two Large Midwestern Kindreds. *Archives of Internal Medicine* 117(2): 206–212.
- Lynch, Henry T., Carrie L. Snyder, Trudy G. Shaw, Christopher D. Heinen, and Megan P. Hitchins (2015) Milestones of Lynch Syndrome: 1895-2015. *Nature Reviews. Cancer* 15(3): 181–194.
- Maby, P., J. Galon, and J.-B. Latouche (2016) Frameshift Mutations, Neoantigens and Tumor-Specific CD8(+) T Cells in Microsatellite Unstable Colorectal Cancers. *Oncoimmunology* 5(5): e1115943.
- Malkhosyan, S., N. Rampino, H. Yamamoto, and M. Perucho (1996) Frameshift Mutator Mutations. *Nature* 382(6591): 499–500.
- Marisa, L., A. de Reyniès, A. Duval, *et al.* (2013) Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value. *PLoS Medicine* 10(5): e1001453.
- Marra, G., S. D’Atri, C. Corti, *et al.* (2001) Tolerance of Human MSH2+/- Lymphoblastoid Cells to the Methylating Agent Temozolomide. *Proceedings of the National Academy of Sciences of the United States of America* 98(13): 7164–7169.
- Martinez, S. L., and R. D. Kolodner (2010) Functional Analysis of Human Mismatch Repair Gene Mutations Identifies Weak Alleles and Polymorphisms Capable of Polygenic Interactions. *Proceedings of the National Academy of Sciences of the United States of America* 107(11): 5070–5075.
- Mastrocola, A. S., and C. D. Heinen (2010) Lynch Syndrome-Associated Mutations in MSH2 Alter DNA Repair and Checkpoint Response Functions in Vivo. *Human Mutation* 31(10): E1699-1708.
- Mazurek, A., C. N. Johnson, M. W. Germann, and R. Fishel (2009) Sequence Context Effect for HSMH2-HSMH6 Mismatch-Dependent Activation. *Proceedings of the National Academy of Sciences* 106(11): 4177–4182.
- Meikrantz, W., M. A. Bergom, A. Memisoglu, and L. Samson (1998) O6-Alkylguanine DNA Lesions Trigger Apoptosis. *Carcinogenesis* 19(2): 369–372.
- Mensenkamp, A. R., I. P. Vogelaar, W. A. G. van Zelst-Stams, *et al.* (2014) Somatic Mutations in MLH1 and MSH2 Are a Frequent Cause of Mismatch-Repair Deficiency in Lynch Syndrome-like Tumors. *Gastroenterology* 146(3): 643-646.e8.
- Metzker, M. L. (2010) Sequencing Technologies - the next Generation. *Nature Reviews. Genetics* 11(1): 31–46.
- Meyers, M., M. W. Wagner, H. S. Hwang, T. J. Kinsella, and D. A. Boothman (2001) Role of the HMLH1 DNA Mismatch Repair Protein in Fluoropyrimidine-Mediated Cell Death and Cell Cycle Responses. *Cancer Research* 61(13): 5193–5201.
- Miyaki, M., M. Konishi, K. Tanaka, *et al.* (1997) Germline Mutation of MSH6 as the Cause of Hereditary Nonpolyposis Colorectal Cancer. *Nature Genetics* 17(3): 271–272.
- Modrich, Paul (1991) Mechanisms and Biological Effects of Mismatch Repair: 25.
- Modrich, Paul (2006) Mechanisms in Eukaryotic Mismatch Repair. *The Journal of Biological Chemistry* 281(41): 30305–30309.

- Mojas, N., M. Lopes, and J. Jiricny (2007) Mismatch Repair-Dependent Processing of Methylation Damage Gives Rise to Persistent Single-Stranded Gaps in Newly Replicated DNA. *Genes & Development* 21(24): 3342–3355.
- Mojtahed, A., I. Schrijver, J. M. Ford, T. A. Longacre, and R. K. Pai (2011) A Two-Antibody Mismatch Repair Protein Immunohistochemistry Screening Approach for Colorectal Carcinomas, Skin Sebaceous Tumors, and Gynecologic Tract Carcinomas. *Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc* 24(7): 1004–1014.
- Moreira, L., F. Balaguer, N. Lindor, *et al.* (2012) Identification of Lynch Syndrome among Patients with Colorectal Cancer. *JAMA* 308(15): 1555–1565.
- Muleris, M., A. Chalastanis, N. Meyer, *et al.* (2008) Chromosomal Instability in Near-Diploid Colorectal Cancer: A Link between Numbers and Structure. *PloS One* 3(2): e1632.
- Nakagawa, H., J. C. Lockman, W. L. Frankel, *et al.* (2004) Mismatch Repair Gene PMS2: Disease-Causing Germline Mutations Are Frequent in Patients Whose Tumors Stain Negative for PMS2 Protein, but Paralogous Genes Obscure Mutation Detection and Interpretation. *Cancer Research* 64(14): 4721–4727.
- Nakatsuru, Y., S. Matsukuma, N. Nemoto, *et al.* (1993) O6-Methylguanine-DNA Methyltransferase Protects against Nitrosamine-Induced Hepatocarcinogenesis. *Medical Sciences*: 5.
- Ng, P. C., and S. Henikoff (2002) Accounting for Human Polymorphisms Predicted to Affect Protein Function. *Genome Research* 12(3): 436–446.
- Ng, P. C., and S. Henikoff (2003) SIFT: Predicting Amino Acid Changes That Affect Protein Function. *Nucleic Acids Research* 31(13): 3812–3814.
- Nicolaides, N. C., N. Papadopoulos, B. Liu, *et al.* (1994) Mutations of Two PMS Homologues in Hereditary Nonpolyposis Colon Cancer. *Nature* 371(6492): 75–80.
- Nielsen, S. V., A. Stein, A. B. Dinitzen, *et al.* (2017) Predicting the Impact of Lynch Syndrome-Causing Missense Mutations from Structural Calculations. *PLoS Genetics* 13(4): e1006739.
- Niessen, R. C., R. M. W. Hofstra, H. Westers, *et al.* (2009) Germline Hypermethylation of MLH1 and EPCAM Deletions Are a Frequent Cause of Lynch Syndrome. *Genes, Chromosomes & Cancer* 48(8): 737–744.
- Nojadeh, J. N., S. Behrouz Sharif, and E. Sakhinia (2018) Microsatellite Instability in Colorectal Cancer. *EXCLI Journal* 17: 159–168.
- Nyström-Lahti, M., C. Perrera, M. Räschele, *et al.* (2002) Functional Analysis of MLH1 Mutations Linked to Hereditary Nonpolyposis Colon Cancer. *Genes, Chromosomes & Cancer* 33(2): 160–167.
- O'Brien, V., and R. Brown (2006) Signalling Cell Cycle Arrest and Cell Death through the MMR System. *Carcinogenesis* 27(4): 682–692.
- Offman, J., G. Opelz, B. Doehler, *et al.* (2004) Defective DNA Mismatch Repair in Acute Myeloid Leukemia/Myelodysplastic Syndrome after Organ Transplantation. *Blood* 104(3): 822–828.
- Ollila, S., D. Dermadi Bebek, M. Greenblatt, and M. Nyström (2008) Uncertain Pathogenicity of MSH2 Variants N127S and G322D Challenges Their Classification. *International Journal of Cancer* 123(3): 720–724.
- Ollila, S., D. Dermadi Bebek, J. Jiricny, and M. Nyström (2008) Mechanisms of Pathogenicity in Human MSH2 Missense Mutants. *Human Mutation* 29(11): 1355–1363.
- Ollila, S., L. Sarantaus, R. Kariola, *et al.* (2006) Pathogenicity of MSH2 Missense Mutations Is Typically Associated with Impaired Repair Capability of the Mutated Protein. *Gastroenterology* 131(5): 1408–1417.

- Ou, J., R. C. Niessen, A. Lützen, *et al.* (2007) Functional Analysis Helps to Clarify the Clinical Importance of Unclassified Variants in DNA Mismatch Repair Genes. *Human Mutation* 28(11): 1047–1054.
- Papadopoulos, N., N. C. Nicolaides, Y. F. Wei, *et al.* (1994) Mutation of a MutL Homolog in Hereditary Colon Cancer. *Science (New York, N.Y.)* 263(5153): 1625–1629.
- Park, J. G., Y. J. Park, J. T. Wijnen, and H. F. Vasen (1999) Gene-Environment Interaction in Hereditary Nonpolyposis Colorectal Cancer with Implications for Diagnosis and Genetic Testing. *International Journal of Cancer* 82(4): 516–519.
- Parsons, M. T., D. D. Buchanan, B. Thompson, J. P. Young, and A. B. Spurdle (2012) Correlation of Tumour BRAF Mutations and MLH1 Methylation with Germline Mismatch Repair (MMR) Gene Mutation Status: A Literature Review Assessing Utility of Tumour Features for MMR Variant Classification. *Journal of Medical Genetics* 49(3): 151–157.
- Parsons, R., G. M. Li, M. Longley, *et al.* (1995) Mismatch Repair Deficiency in Phenotypically Normal Human Cells. *Science (New York, N.Y.)* 268(5211): 738–740.
- Pavlov, Y. I., I. M. Mian, and T. A. Kunkel (2003) Evidence for Preferential Mismatch Repair of Lagging Strand DNA Replication Errors in Yeast. *Current Biology: CB* 13(9): 744–748.
- Pegg, A. E. (2000) Repair of O6-Alkylguanine by Alkyltransferases. *Mutation Research/Reviews in Mutation Research* 462(2–3): 83–100.
- Pegg, A. E., and T. L. Byers (1992) Repair of DNA Containing O6-Alkylguanine. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 6(6): 2302–2310.
- Peltomäki, P. (2016) Update on Lynch Syndrome Genomics. *Familial Cancer* 15(3): 385–393.
- Peltomäki, P., R. A. Lothe, L. A. Aaltonen, *et al.* (1993) Microsatellite Instability Is Associated with Tumors That Characterize the Hereditary Non-Polyposis Colorectal Carcinoma Syndrome. *Cancer Research* 53(24): 5853–5855.
- Peltomäki, P., P. Sistonen, J. P. Mecklin, *et al.* (1992) Evidence That the MCC-APC Gene Region in 5q21 Is Not the Site for Susceptibility to Hereditary Nonpolyposis Colorectal Carcinoma. *Cancer Research* 52(16): 4530–4533.
- Peltomäki, P., and H. Vasen (2004) Mutations Associated with HNPCC Predisposition -- Update of ICG-HNPCC/INSiGHT Mutation Database. *Disease Markers* 20(4–5): 269–276.
- Perera, S., and B. Bapat (2008) The MLH1 Variants p.Arg265Cys and p.Lys618Ala Affect Protein Stability While p.Leu749Gln Affects Heterodimer Formation. *Human Mutation* 29(2): 332.
- Plon, S. E., D. M. Eccles, D. Easton, *et al.* (2008) Sequence Variant Classification and Reporting: Recommendations for Improving the Interpretation of Cancer Susceptibility Genetic Test Results. *Human Mutation* 29(11): 1282–1291.
- Plotz, G., C. Welsch, L. Giron-Monzon, *et al.* (2006) Mutations in the MutS $\alpha$  Interaction Interface of MLH1 Can Abolish DNA Mismatch Repair. *Nucleic Acids Research* 34(22): 6574–6586.
- Pluciennik, A., L. Dzantiev, R. R. Iyer, *et al.* (2010) PCNA Function in the Activation and Strand Direction of MutL $\alpha$  Endonuclease in Mismatch Repair. *Proceedings of the National Academy of Sciences of the United States of America* 107(37): 16066–16071.
- Polaczek, P., A. P. Putzke, K. Leong, and G. A. Bitter (1998) Functional Genetic Tests of DNA Mismatch Repair Protein Activity in *Saccharomyces Cerevisiae*. *Gene* 213(1–2): 159–167.
- Ponti, G., E. Castellsagué, C. Ruini, A. Percesepe, and A. Tomasi (2015) Mismatch Repair Genes Founder Mutations and Cancer Susceptibility in Lynch Syndrome. *Clinical Genetics* 87(6): 507–516.

- Popat, S., R. Hubner, and R. S. Houlston (2005) Systematic Review of Microsatellite Instability and Colorectal Cancer Prognosis. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 23(3): 609–618.
- Porkka, N., S. Valo, T. T. Nieminen, *et al.* (2017) Sequencing of Lynch Syndrome Tumors Reveals the Importance of Epigenetic Alterations. *Oncotarget* 8(64): 108020–108030.
- Potocnik, U., D. Glavac, R. Golouh, and M. Ravnik-Glavac (2001) Causes of Microsatellite Instability in Colorectal Tumors: Implications for Hereditary Non-Polyposis Colorectal Cancer Screening. *Cancer Genetics and Cytogenetics* 126(2): 85–96.
- Poulogiannis, G., I. M. Frayling, and M. J. Arends (2010) DNA Mismatch Repair Deficiency in Sporadic Colorectal Cancer and Lynch Syndrome. *Histopathology* 56(2): 167–179.
- Raevaara, T. E., M. K. Korhonen, H. Lohi, *et al.* (2005) Functional Significance and Clinical Phenotype of Nontruncating Mismatch Repair Variants of MLH1. *Gastroenterology* 129(2): 537–549.
- Rana, H. Q., and S. Syngal (2017) Biallelic Mismatch Repair Deficiency: Management and Prevention of a Devastating Manifestation of the Lynch Syndrome. *Gastroenterology* 152(6): 1254–1257.
- Rasmussen, L. J., C. D. Heinen, B. Royer-Pokora, *et al.* (2012) Pathological Assessment of Mismatch Repair Gene Variants in Lynch Syndrome: Past, Present, and Future. *Human Mutation* 33(12): 1617–1625.
- Reese, J. S., L. Liu, and S. L. Gerson (2003) Repopulating Defect of Mismatch Repair-Deficient Hematopoietic Stem Cells. *Blood* 102(5): 1626–1633.
- Reha-Krantz, L. J. (2010) DNA Polymerase Proofreading: Multiple Roles Maintain Genome Stability. *Biochimica Et Biophysica Acta* 1804(5): 1049–1063.
- Reyes, G. X., T. T. Schmidt, R. D. Kolodner, and H. Hombauer (2015) New Insights into the Mechanism of DNA Mismatch Repair. *Chromosoma* 124(4): 443–462.
- Ribic, C. M., D. J. Sargent, M. J. Moore, *et al.* (2003) Tumor Microsatellite-Instability Status as a Predictor of Benefit from Fluorouracil-Based Adjuvant Chemotherapy for Colon Cancer. *The New England Journal of Medicine* 349(3): 247–257.
- Ricciardone, M. D., T. Ozçelik, B. Cevher, *et al.* (1999) Human MLH1 Deficiency Predisposes to Hematological Malignancy and Neurofibromatosis Type 1. *Cancer Research* 59(2): 290–293.
- Richards, S., N. Aziz, S. Bale, *et al.* (2015) 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. *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 17(5): 405–424.
- Richman, S. (2015) Deficient Mismatch Repair: Read All about It (Review). *International Journal of Oncology* 47(4): 1189–1202.
- Ripperger, T., and B. Schlegelberger (2016) Acute Lymphoblastic Leukemia and Lymphoma in the Context of Constitutional Mismatch Repair Deficiency Syndrome. *European Journal of Medical Genetics* 59(3): 133–142.
- Ryan, E., K. Sheahan, B. Creavin, H. M. Mohan, and D. C. Winter (2017) The Current Value of Determining the Mismatch Repair Status of Colorectal Cancer: A Rationale for Routine Testing. *Critical Reviews in Oncology/Hematology* 116: 38–57.
- Saffhill, R., G. P. Margison, and P. J. O'Connor (1985) Mechanisms of Carcinogenesis Induced by Alkylating Agents. *Biochimica Et Biophysica Acta* 823(2): 111–145.

- Samson, L., B. Derfler, and E. A. Waldstein (1986) Suppression of Human DNA Alkylation-Repair Defects by *Escherichia Coli* DNA-Repair Genes. *Proceedings of the National Academy of Sciences* 83(15): 5607–5610.
- Sansom, O. J., S. M. Bishop, H. Court, *et al.* (2003) Apoptosis and Mutation in the Murine Small Intestine: Loss of Mlh1- and Pms2-Dependent Apoptosis Leads to Increased Mutation *In Vivo*. *DNA Repair* 2(9): 1029–1039.
- Sansom, O. J., N. J. Toft, D. J. Winton, and A. R. Clarke (2001) Msh-2 Suppresses *In Vivo* Mutation in a Gene Dose and Lesion Dependent Manner. *Oncogene* 20(27): 3580–3584.
- Sargent, D. J., S. Marsoni, G. Monges, *et al.* (2010) Defective Mismatch Repair as a Predictive Marker for Lack of Efficacy of Fluorouracil-Based Adjuvant Therapy in Colon Cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 28(20): 3219–3226.
- Sarkaria, J. N., G. J. Kitange, C. D. James, *et al.* (2008) Mechanisms of Chemoresistance in Malignant Glioma: 17.
- Schmeler, K. M., H. T. Lynch, L.-M. Chen, *et al.* (2006) Prophylactic Surgery to Reduce the Risk of Gynecologic Cancers in the Lynch Syndrome. *The New England Journal of Medicine* 354(3): 261–269.
- Schmidt, T. T., and H. Hombauer (2016) Visualization of Mismatch Repair Complexes Using Fluorescence Microscopy. *DNA Repair* 38: 58–67.
- Schulmann, K., Y. Mori, V. Croog, *et al.* (2005) Molecular Phenotype of Inflammatory Bowel Disease-Associated Neoplasms with Microsatellite Instability. *Gastroenterology* 129(1): 74–85.
- Shcherbakova, P. V., and T. A. Kunkel (1999) Mutator Phenotypes Conferred by MLH1 Overexpression and by Heterozygosity for Mlh1 Mutations. *Molecular and Cellular Biology* 19(4): 3177–3183.
- Shcherbakova, P. V., Y. I. Pavlov, O. Chilkova, *et al.* (2003) Unique Error Signature of the Four-Subunit Yeast DNA Polymerase Epsilon. *The Journal of Biological Chemistry* 278(44): 43770–43780.
- Shia, J. (2008) Immunohistochemistry versus Microsatellite Instability Testing for Screening Colorectal Cancer Patients at Risk for Hereditary Nonpolyposis Colorectal Cancer Syndrome. Part I. The Utility of Immunohistochemistry. *The Journal of Molecular Diagnostics: JMD* 10(4): 293–300.
- Shimelis, H., R. L. S. Mesman, C. Von Nicolai, *et al.* (2017) BRCA2 Hypomorphic Missense Variants Confer Moderate Risks of Breast Cancer. *Cancer Research* 77(11): 2789–2799.
- Shimodaira, H., N. Filosi, H. Shibata, *et al.* (1998) Functional Analysis of Human MLH1 Mutations in *Saccharomyces Cerevisiae*. *Nature Genetics* 19(4): 384–389.
- Shirts, B. H., E. Q. Konnick, S. Upham, *et al.* (2018) Using Somatic Mutations from Tumors to Classify Variants in Mismatch Repair Genes. *American Journal of Human Genetics* 103(1): 19–29.
- Shlien, A., B. B. Campbell, R. de Borja, *et al.* (2015) Combined Hereditary and Somatic Mutations of Replication Error Repair Genes Result in Rapid Onset of Ultra-Hypermutated Cancers. *Nature Genetics* 47(3): 257–262.
- Sijmons, R. H., M. S. Greenblatt, and M. Genuardi (2013) Gene Variants of Unknown Clinical Significance in Lynch Syndrome. An Introduction for Clinicians. *Familial Cancer* 12(2): 181–187.
- Singh, P. P., P. K. Sharma, G. Krishnan, and A. C. Lockhart (2015) Immune Checkpoints and Immunotherapy for Colorectal Cancer. *Gastroenterology Report* 3(4): 289–297.
- Sinicrope, F. A., N. R. Foster, S. N. Thibodeau, *et al.* (2011) DNA Mismatch Repair Status and Colon Cancer Recurrence and Survival in Clinical Trials of 5-Fluorouracil-Based Adjuvant Therapy. *Journal of the National Cancer Institute* 103(11): 863–875.

- Smith, C. G., D. Fisher, B. Claes, *et al.* (2013) Somatic Profiling of the Epidermal Growth Factor Receptor Pathway in Tumors from Patients with Advanced Colorectal Cancer Treated with Chemotherapy ± Cetuximab. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 19(15): 4104–4113.
- Solomon, I., E. Harrington, G. Hooker, *et al.* (2017) Lynch Syndrome Limbo: Patient Understanding of Variants of Uncertain Significance. *Journal of Genetic Counseling* 26(4): 866–877.
- Sourrouille, I., F. Coulet, J. H. Lefevre, *et al.* (2013) Somatic Mosaicism and Double Somatic Hits Can Lead to MSI Colorectal Tumors. *Familial Cancer* 12(1): 27–33.
- Speevak, M. D., S. S. Young, H. Feilotter, and P. Ainsworth (2003) Alternatively Spliced, Truncated Human BRCA2 Isoforms Contain a Novel Coding Exon. *European Journal of Human Genetics: EJHG* 11(12): 951–954.
- Spurdle, A. B., F. J. Couch, F. B. L. Hogervorst, *et al.* (2008) Prediction and Assessment of Splicing Alterations: Implications for Clinical Testing. *Human Mutation* 29(11): 1304–1313.
- St Charles, J. A., S. E. Liberti, J. S. Williams, S. A. Lujan, and T. A. Kunkel (2015) Quantifying the Contributions of Base Selectivity, Proofreading and Mismatch Repair to Nuclear DNA Replication in *Saccharomyces Cerevisiae*. *DNA Repair* 31: 41–51.
- Stone, E. A., and A. Sidow (2005) Physicochemical Constraint Violation by Missense Substitutions Mediates Impairment of Protein Function and Disease Severity. *Genome Research* 15(7): 978–986.
- Suraweera, N., A. Duval, M. Reperant, *et al.* (2002) Evaluation of Tumor Microsatellite Instability Using Five Quasimonomorphic Mononucleotide Repeats and Pentaplex PCR. *Gastroenterology* 123(6): 1804–1811.
- Suzuki, H., N. Harpaz, L. Tarmin, *et al.* (1994) Microsatellite Instability in Ulcerative Colitis-Associated Colorectal Dysplasias and Cancers. *Cancer Research* 54(18): 4841–4844.
- Svrcek, M., L. Beaugerie, J. Kirchgessner, *et al.* (2016) Azathioprine May Not Be Associated With Development of Colorectal Tumors With Microsatellite Instability in Patients With Inflammatory Bowel Disease. *The American Journal of Gastroenterology* 111(1): 149–151.
- Svrcek, M., N. El-Murr, K. Wanherdrick, *et al.* (2013) Overexpression of MicroRNAs-155 and 21 Targeting Mismatch Repair Proteins in Inflammatory Bowel Diseases. *Carcinogenesis* 34(4): 828–834.
- Swann, P. F., T. R. Waters, D. C. Moulton, *et al.* (1996) Role of Postreplicative DNA Mismatch Repair in the Cytotoxic Action of Thioguanine. *Science* 273(5278): 1109–1111.
- Tajima, A., M. T. Hess, B. L. Cabrera, R. D. Kolodner, and J. M. Carethers (2004) The Mismatch Repair Complex HMutS Alpha Recognizes 5-Fluorouracil-Modified DNA: Implications for Chemosensitivity and Resistance. *Gastroenterology* 127(6): 1678–1684.
- Takahashi, M., H. Shimodaira, C. Andreutti-Zaugg, *et al.* (2007) Functional Analysis of Human MLH1 Variants Using Yeast and in Vitro Mismatch Repair Assays. *Cancer Research* 67(10): 4595–4604.
- Takemura, H., V. A. Rao, O. Sordet, *et al.* (2006) Defective Mre11-Dependent Activation of Chk2 by Ataxia Telangiectasia Mutated in Colorectal Carcinoma Cells in Response to Replication-Dependent DNA Double Strand Breaks. *The Journal of Biological Chemistry* 281(41): 30814–30823.
- Tanakaya, K. (2018) Current Clinical Topics of Lynch Syndrome. *International Journal of Clinical Oncology*.

- Thibodeau, S. N., G. Bren, and D. Schaid (1993) Microsatellite Instability in Cancer of the Proximal Colon. *Science (New York, N.Y.)* 260(5109): 816–819.
- Thompson, B. A., A. B. Spurdle, J.-P. Plazzer, *et al.* (2014) Application of a 5-Tiered Scheme for Standardized Classification of 2,360 Unique Mismatch Repair Gene Variants in the InSiGHT Locus-Specific Database. *Nature Genetics* 46(2): 107–115.
- Thusberg, J., A. Olatubosun, and M. Vihinen (2011) Performance of Mutation Pathogenicity Prediction Methods on Missense Variants. *Human Mutation* 32(4): 358–368.
- Tindall, K. R., W. E. Glaab, A. Umar, *et al.* (1998) Complementation of Mismatch Repair Gene Defects by Chromosome Transfer. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 402(1–2): 15–22.
- Toft, N. J., D. J. Winton, J. Kelly, *et al.* (1999) Msh2 Status Modulates Both Apoptosis and Mutation Frequency in the Murine Small Intestine. *Proceedings of the National Academy of Sciences of the United States of America* 96(7): 3911–3915.
- Tokuriki, N., F. Stricher, J. Schymkowitz, L. Serrano, and D. S. Tawfik (2007) The Stability Effects of Protein Mutations Appear to Be Universally Distributed. *Journal of Molecular Biology* 369(5): 1318–1332.
- Tominaga, Y., T. Tsuzuki, A. Shiraishi, H. Kawate, and M. Sekiguchi (1997) Alkylation-Induced Apoptosis of Embryonic Stem Cells in Which the Gene for DNA-Repair, Methyltransferase, Had Been Disrupted by Gene Targeting. *Carcinogenesis* 18(5): 889–896.
- Tomita-Mitchell, A., A. G. Kat, L. A. Marcelino, *et al.* (2000) Mismatch Repair Deficient Human Cells: Spontaneous and MNNG-Induced Mutational Spectra in the HPRT Gene. *Mutation Research* 450(1–2): 125–138.
- Tomlinson, I. P., M. Ilyas, and W. F. Bodmer (1996) Allele Loss Occurs Frequently at HMLH1, but Rarely at HMSH2, in Sporadic Colorectal Cancers with Microsatellite Instability. *British Journal of Cancer* 74(10): 1514–1517.
- Tournier, I., M. Vezain, A. Martins, *et al.* (2008) A Large Fraction of Unclassified Variants of the Mismatch Repair Genes MLH1 and MSH2 Is Associated with Splicing Defects. *Human Mutation* 29(12): 1412–1424.
- Tricarico, R., M. Kasela, C. Marenzi, *et al.* (2017) Assessment of the InSiGHT Interpretation Criteria for the Clinical Classification of 24 MLH1 and MSH2 Gene Variants. *Human Mutation* 38(1): 64–77.
- Trojan, Joerg, Stefan Zeuzem, Ann Randolph, *et al.* (2002) Functional Analysis of HMLH1 Variants and HNPCC-Related Mutations Using a Human Expression System. *Gastroenterology* 122(1): 211–219.
- Truninger, K., M. Menigatti, J. Luz, *et al.* (2005) Immunohistochemical Analysis Reveals High Frequency of PMS2 Defects in Colorectal Cancer. *Gastroenterology* 128(5): 1160–1171.
- Umar, A., C. R. Boland, J. P. Terdiman, *et al.* (2004) Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability. *Journal of the National Cancer Institute* 96(4): 261–268.
- Umar, A., J. I. Risinger, E. Glaab, E. Tindall, and D. Kolodner (1997) Correction of Hypermutability, N-Methyl-N'-Nitro-N-Nitrosoguanidine Resistance, and Defective DNA Mismatch Repair by Introducing Chromosome 2 into Human Tumor Cells with Mutations in MSH2 and MSH6: 8.
- Valeri, N., P. Gasparini, C. Braconi, *et al.* (2010) MicroRNA-21 Induces Resistance to 5-Fluorouracil by down-Regulating Human DNA MutS Homolog 2 (HMSH2). *Proceedings of the National Academy of Sciences of the United States of America* 107(49): 21098–21103.

- Valeri, N., P. Gasparini, M. Fabbri, *et al.* (2010) Modulation of Mismatch Repair and Genomic Stability by MiR-155. *Proceedings of the National Academy of Sciences of the United States of America* 107(15): 6982–6987.
- Vasen, H. F. A., I. Blanco, K. Aktan-Collan, *et al.* (2013) Revised Guidelines for the Clinical Management of Lynch Syndrome (HNPCC): Recommendations by a Group of European Experts. *Gut* 62(6): 812–823.
- Vasen, H. F. A., Z. Ghorbanoghli, F. Bourdeaut, *et al.* (2014) Guidelines for Surveillance of Individuals with Constitutional Mismatch Repair-Deficiency Proposed by the European Consortium “Care for CMMR-D” (C4CMMR-D). *Journal of Medical Genetics* 51(5): 283–293.
- Vasen, H. F., P. Watson, J. P. Mecklin, and H. T. Lynch (1999) New Clinical Criteria for Hereditary Nonpolyposis Colorectal Cancer (HNPCC, Lynch Syndrome) Proposed by the International Collaborative Group on HNPCC. *Gastroenterology* 116(6): 1453–1456.
- Vembar, S. S., and J. L. Brodsky (2008) One Step at a Time: Endoplasmic Reticulum-Associated Degradation. *Nature Reviews. Molecular Cell Biology* 9(12): 944–957.
- Venderbosch, S., I. D. Nagtegaal, T. S. Maughan, *et al.* (2014) Mismatch Repair Status and BRAF Mutation Status in Metastatic Colorectal Cancer Patients: A Pooled Analysis of the CAIRO, CAIRO2, COIN, and FOCUS Studies. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 20(20): 5322–5330.
- Verma, L., M. F. Kane, C. Brassett, *et al.* (1999) Mononucleotide Microsatellite Instability and Germline MSH6 Mutation Analysis in Early Onset Colorectal Cancer. *Journal of Medical Genetics* 36(9): 678–682.
- Vogelsang, M., A. Comino, N. Zupanec, P. Hudler, and R. Komel (2009) Assessing Pathogenicity of MLH1 Variants by Co-Expression of Human MLH1 and PMS2 Genes in Yeast. *BMC Cancer* 9: 382.
- Wanat, J. J., H. Singh, and E. Alani (2007) The Effect of Genetic Background on the Function of *Saccharomyces Cerevisiae* Mlh1 Alleles That Correspond to HNPCC Missense Mutations. *Human Molecular Genetics* 16(4): 445–452.
- Wang, Q., C. Lasset, F. Desseigne, *et al.* (1999) Neurofibromatosis and Early Onset of Cancers in HMLH1-Deficient Children. *Cancer Research* 59(2): 294–297.
- Wang, Q., G. Montmain, E. Ruano, *et al.* (2003) Neurofibromatosis Type 1 Gene as a Mutational Target in a Mismatch Repair-Deficient Cell Type. *Human Genetics* 112(2): 117–123.
- Wang, Y., and J. Qin (2003) MSH2 and ATR Form a Signaling Module and Regulate Two Branches of the Damage Response to DNA Methylation. *Proceedings of the National Academy of Sciences* 100(26): 15387–15392.
- Warthin, A. S. (1985) *Classics in Oncology. Heredity with Reference to Carcinoma as Shown by the Study of the Cases Examined in the Pathological Laboratory of the University of Michigan, 1895-1913.* By Aldred Scott Warthin. 1913. *CA: A Cancer Journal for Clinicians* 35(6): 348–359.
- Wielders, E. A. L., R. J. Dekker, I. Holt, G. E. Morris, and H. te Riele (2011) Characterization of MSH2 Variants by Endogenous Gene Modification in Mouse Embryonic Stem Cells. *Human Mutation* 32(4): 389–396.
- Wielders, E. A. L., J. Hettinger, R. Dekker, *et al.* (2014) Functional Analysis of MSH2 Unclassified Variants Found in Suspected Lynch Syndrome Patients Reveals Pathogenicity Due to Attenuated Mismatch Repair. *Journal of Medical Genetics* 51(4): 245–253.
- Wielders, E. A. L., H. Houlleberghs, G. Isik, and H. te Riele (2013) Functional Analysis in Mouse Embryonic Stem Cells Reveals Wild-Type Activity for Three MSH6 Variants Found in Suspected Lynch Syndrome Patients. *PloS One* 8(9): e74766.

- Willenbacher, R. F., D. E. Aust, C. G. Chang, *et al.* (1999) Genomic Instability Is an Early Event during the Progression Pathway of Ulcerative-Colitis-Related Neoplasia. *The American Journal of Pathology* 154(6): 1825–1830.
- Wimmer, K., and J. Etzler (2008) Constitutional Mismatch Repair-Deficiency Syndrome: Have We so Far Seen Only the Tip of an Iceberg? *Human Genetics* 124(2): 105–122.
- Wimmer, K., C. P. Kratz, H. F. A. Vasen, *et al.* (2014) Diagnostic Criteria for Constitutional Mismatch Repair Deficiency Syndrome: Suggestions of the European Consortium “care for CMMRD” (C4CMMRD). *Journal of Medical Genetics* 51(6): 355–365.
- Win, A. K., M. A. Jenkins, J. G. Dowty, *et al.* (2017) Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology* 26(3): 404–412.
- de Wind, N., M. Dekker, N. Claij, *et al.* (1999) HNPCC-like Cancer Predisposition in Mice through Simultaneous Loss of Msh3 and Msh6 Mismatch-Repair Protein Functions. *Nature Genetics* 23(3): 359–362.
- Worrillow, L. J., L. B. Travis, A. G. Smith, *et al.* (2003) An Intron Splice Acceptor Polymorphism in HSMH2 and Risk of Leukemia after Treatment with Chemotherapeutic Alkylating Agents. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 9(8): 3012–3020.
- Wu, Z., C.-L. Chan, A. Eastman, and E. Bresnick (1991) Expression of Human O6-Methylguanine-DNA Methyltransferase in Chinese Hamster Ovary Cells and Restoration of Cellular Resistance to Certain N-Nitroso Compounds. *Molecular Carcinogenesis* 4(6): 482–488.
- Yamamoto, H., and K. Imai (2015) Microsatellite Instability: An Update. *Archives of Toxicology* 89(6): 899–921.
- Yang, Guohze, Stefan J. Scherer, Scarlet S. Shell, *et al.* (2004) Dominant Effects of an Msh6 Missense Mutation on DNA Repair and Cancer Susceptibility. *Cancer Cell* 6(2): 139–150.
- Yoon, K., S. Lee, T.-S. Han, *et al.* (2013) Comprehensive Genome- and Transcriptome-Wide Analyses of Mutations Associated with Microsatellite Instability in Korean Gastric Cancers. *Genome Research* 23(7): 1109–1117.
- Yoshioka, K.-I., Y. Yoshioka, and P. Hsieh (2006) ATR Kinase Activation Mediated by MutS $\alpha$  and MutL $\alpha$  in Response to Cytotoxic O6-Methylguanine Adducts. *Molecular Cell* 22(4): 501–510.
- Yurgelun, M. B., B. Allen, R. R. Kaldate, *et al.* (2015) Identification of a Variety of Mutations in Cancer Predisposition Genes in Patients With Suspected Lynch Syndrome. *Gastroenterology* 149(3): 604-613.e20.
- Zaidi, N. H., T. P. Pretlow, M. A. O’Riordan, *et al.* (1995) Transgenic Expression of Human MGMT Protects against Azoxymethane-Induced Aberrant Crypt Foci and G to A Mutations in the K-Ras Oncogene of Mouse Colon. *Carcinogenesis* 16(3): 451–456.
- Zerdoumi, Yasmine, Juliette Aury-Landas, Catherine Bonaiti-Pellié, *et al.* (2013) Drastic Effect of Germline TP53 Missense Mutations in Li-Fraumeni Patients. *Human Mutation* 34(3): 453–461.
- Zhang, L. (2008) Immunohistochemistry versus Microsatellite Instability Testing for Screening Colorectal Cancer Patients at Risk for Hereditary Nonpolyposis Colorectal Cancer Syndrome. Part II. The Utility of Microsatellite Instability Testing. *The Journal of Molecular Diagnostics: JMD* 10(4): 301–307.
- Zhang, X., S. Chen, J. Yu, *et al.* (2018) Analysis of Human MutS Homolog 2 Missense Mutations in Patients with Colorectal Cancer. *Oncology Letters* 15(5): 6275–6282.

Zhou, Z.-Q., D. Manguino, K. Kewitt, *et al.* (2001) Spontaneous Hepatocellular Carcinoma Is Reduced in Transgenic Mice Overexpressing Human O6- Methylguanine-DNA Methyltransferase. *Proceedings of the National Academy of Sciences* 98(22): 12566–12571.