

UNIVERSITE DE LA MEDITERRANEE
FACULTE DE MÉDECINE DE MARSEILLE
ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

T H È S E

Présentée et publiquement soutenue devant

LA FACULTÉ DE MÉDECINE DE MARSEILLE

Le 11 Avril 2014

Par Mr Philippe GRANDVAL

Né le 17 décembre 1969 à Marseille

**CARACTERISATION DES VARIATIONS GENETIQUES
CONSTITUTIONNELLES DE SIGNIFICATION INCONNUE DANS
LE SYNDROME DE LYNCH**

Pour obtenir le grade de DOCTEUR DE L'UNIVERSITE DE LA MEDITERRANEE

SPÉCIALITÉ : Génétique Médicale et Génomique Fonctionnelle

Membres du Jury de la Thèse :

M. le Professeur René LAUGIER	Président
M. Le Professeur Pierre LAURENT-PUIG	Rapporteur
M. le Professeur Jean Christophe SAURIN	Rapporteur
M. le Professeur Alex KARTHEUSER	Examineur
Mme le Docteur Sylviane OLSCHWANG	Directeur

REMERCIEMENTS

A MON DIRECTEUR DE THESE :

Madame le Docteur Sylviane OLCHWANG.

Ma chère Sylviane,

Quatre ans de collaboration déjà, des projets qui se concrétisent les uns après les autres, reflétant exactement ta façon d'être et de travailler, et une certitude de mener à bien nos objectifs scientifiques. Je te suis infiniment reconnaissant pour ton engagement et ton soutien sans faille, malgré les obstacles, et pour la qualité de ton enseignement. Je suis conscient de ma chance de pouvoir être ton élève et t'assure de mon dévouement le plus sincère.

A MON JURY DE THESE:

A Monsieur le Professeur René LAUGIER.

Cher Maitre, j'ai l'honneur d'être votre élève depuis de nombreuses années et suis touché de l'amitié que vous me témoignez en présidant ce jury de thèse. Soyez remercié pour votre soutien et assuré de mon profond respect.

A Monsieur le Professeur Jean-Christophe SAURIN.

Cher Jean Christophe, tu m'as initié à l'oncogénétique digestive et témoigné un soutien des plus touchants pour mener à bien mes projets médicaux et scientifiques. Ton parcours et ton engagement pour notre spécialité sont un exemple qui forcent le respect et sont un modèle pour moi et beaucoup d'autres de mes collègues. Soit assuré de ma gratitude.

A Monsieur le Professeur Pierre LAURENT PUIG.

Cher Maitre, vous me faites l'honneur de rapporter mon travail et c'est avec beaucoup d'humilité que je vous le sou mets, au regard de ce que vous représentez dans le domaine de l'oncologie et l'oncogénétique digestive. Soyez également assuré de ma gratitude.

A Monsieur le Professeur Alex KARTHEUSER.

Cher Maitre, c'est un honneur immense que vous me faites de siéger à mon jury de thèse. Le jugement que vous porterez sur ce travail, au travers de votre expérience dans le domaine des prédispositions héréditaires aux cancers digestifs est de la plus grande importance. Je vous suis très reconnaissant pour votre considération et vous assure de mon plus grand respect.

A MES MAITRES ET AMIS

Monsieur le Professeur Jean François SEITZ,

Cher Maître, j'ai la chance de profiter au quotidien de vos connaissances inégalables en oncologie digestive, de votre confiance, et de votre soutien. Aussi je me permets de vous adresser le témoignage de mon profond respect.

Monsieur le Professeur Marc BARTHET

Cher Marc, je me rappelle d'un soir de 2009 ou tu m'as conseillé de m'intéresser à l'oncogénétique.... Je suis fier aujourd'hui de pouvoir te présenter ce travail et compte sur ton soutien pour développer à Marseille cette belle discipline et la faire mieux connaître à nos collègues endoscopistes.

Monsieur Le Professeur Christophe BEROUD

Cher Christophe, je te suis très reconnaissant de m'avoir accueilli dans ton laboratoire et suis fier de faire partie des débuts de votre histoire marseillaise. J'espère y apporter en retour mon concours pour de fructueuses collaborations.

Madame le Docteur Hélène ZATTARA

Ma chère Hélène, tu m'as ouvert les portes de la génétique, et je te dois ma rencontre avec Sylviane OLSCHWANG. Je te suis infiniment reconnaissant pour ton soutien et ta confiance et t'assure de mon amitié et mon dévouement sans faille.

Je dédie cette thèse :

A mes trois enfants, Mathilde, Louise et Adrien. Vous avez décuplé mes forces et mon enthousiasme...

A ma mère et ma sœur. Qui êtes là. Tout le temps. Que ce travail vous récompense de votre Amour et votre soutien.

A mon père et mon frère, qui êtes indissociables de mes pensées, mes choix et mon envie de vivre une vie intense.

A mon Grand Père, modèle jamais atteignable, mais dont je suis l'enseignement intellectuel et spirituel depuis mon enfance.

TABLE DES MATIÈRES

RESUME.....	9
ABREVIATIONS.....	11
INTRODUCTION.....	12
Historique et définitions	12
Spectre « large » du syndrome de lynch : critères d'Amsterdam.....	14
Les critères de la conférence de Bethesda.....	17
Limites des critères d'Amsterdam et de Bethesda.....	18
Génétique du syndrome de Lynch.....	19
Instabilité des microsatellites	29
Le système MMR.....	30
Corrélations Génotype-Phénotype.....	30
Fréquence du syndrome de Lynch.....	31
Phénotype du syndrome de Lynch.....	32
Le cancer colo rectal.....	32
Les adénomes.....	33
Cancer de l'endomètre	33
Autres cancers associés au syndrome de Lynch.....	34
Dépistage et diagnostic du syndrome de Lynch.....	36
La consultation d'oncogénétique et les critères d'indication.....	36
Contenu et déroulement.....	39
Le suivi.....	41
Diagnostic biologique du syndrome de Lynch	42
Analyse immuno histochimique.....	42
Génotypage tumoral et statut MSI.....	45
Analyse moléculaire constitutionnelle	48
Caractérisation des variants.....	50
Bases de données :	52
Historique.....	52
Bases de données et syndrome de Lynch.....	54
BUTS DU TRAVAIL	56

MATERIELS ET METHODES..... 57

Méthodologie de classification des variants	57
Méthodes et outils de caractérisation des variants	58
Histoire familiale	58
Co-segrégation.....	58
Analyses in silico.....	59
Tests fonctionnels	62
Analyses d'épissage	65
Fréquence allélique	71
Co occurrence.....	71
Critères de qualité des bases LSDB et rôle des curateurs	71
Universal Mutation Database (UMD)	73

RESULTATS 74

Publication n°1	75
Publication n°2	78
Publication n°3	82
Publication n°4	84

CONCLUSIONS ET PERSPECTIVES 86

REFERENCES BIBLIOGRAPHIQUES 89

A.....	89
B.....	89
C.....	91
D	92
E.....	93
F.....	93
G.....	94
H	94
J.....	95
K.....	96
L.....	97
M.....	98

N	99
O	99
P	99
Q	101
R	101
S	102
T	104
U	104
V	104
W	105
Y	106
<u>ANNEXES.....</u>	107

RESUME

Le syndrome de Lynch, également appelé HNPCC (hereditary non polyposis colo rectal cancer) est une affection héréditaire autosomique dominante due à des mutations constitutionnelles des gènes du système de réparation de l'ADN (*MLH1*, *MSH2* et *MSH6*). Depuis 20 ans le réseau français des laboratoires impliqués dans le syndrome de Lynch a identifié un total de 6687 variants. Sept cent sept d'entre eux, essentiellement des variants faux sens, restent encore des variants de signification inconnue (VSI), sans utilité pour le conseil génétique. Le but de notre étude était de développer un algorithme permettant de classer les variants de signification inconnue. Les critères utilisés étaient les données des analyses in silico, phénotypiques (ségrégation, critères d'Amsterdam), l'état de la fonction MMR (DNA MisMatch Repair) dans les cellules tumorales, les tests fonctionnels et d'épissage, ainsi que les données publiées. Cet algorithme a été appliqué à l'ensemble des VSI de la base de données française et nous a permis de caractériser 370 variants en neutres (55), probablement neutres (54), probablement pathogènes (80) et pathogènes (181). Les données ont été intégrées dans la base de données française UMD des gènes MMR afin d'être disponibles pour la communauté scientifique. Grace aux données collectées par le réseau, nous avons également pu caractériser le phénotype du syndrome de Lynch. Nous avons ainsi confirmé que le cancer du sein ne fait pas partie du spectre du syndrome de Lynch et que les formes de ce syndrome associées à une mutation du gène *EPCAM* n'entraînent qu'un risque très faible de cancers de l'endomètre, permettant ainsi d'adapter les recommandations de suivi dans cette situation.

Mots clés : syndrome de Lynch, cancer colo rectal, système MMR, variants de signification inconnue, analyse fonctionnelle

ABSTRACT:

Lynch syndrome, also named hereditary non polyposis colo rectal cancer (HNPCC), is a frequent cancer predisposition with an autosomal dominant mode of inheritance and caused by heterozygous germ line mutations in one of the major DNA mismatch repair (MMR) genes (*MLH1*, *MSH2* and *MSH6*). For 20 years, the French laboratories network involved in Lynch syndrome identified a total of 6687 variations. Among them, 707, mainly missense variations, remained variants of uncertain significance (VUS), thus could not be used for reliable genetic counseling. The aim of our study was to develop an algorithm able to classify VUS, according to the international consensus (IARC). This algorithm was constructed based on criteria usually required for genetic characterization such as in silico analysis, phenotypical data (segregation, Amsterdam criteria's), MMR status in tumor cells, functional assays, splicing analyses and published data. Data were registered in the French database (<http://www.umd.be/MLH1/>, <http://www.umd.be/MSH2/> and <http://www.umd.be/MSH6/>). As a result of this work, we were able to classify 370 variants of the 707 (52,3%) as neutral (55), likely neutral (54), likely causal (80) and causal (181).

As part of this work, we also analyzed phenotypical data of patients with Lynch syndrome and showed that breast cancer can definitively be excluded from the spectrum of Lynch-related cancers, and that EPCAM mutations, which may lead to Lynch syndrome, are associated with a very low incidence of endometrial cancer and have probably to be considered as an allelic disease with specific clinical recommendations.

Key words: Lynch syndrome, colo rectal cancer, MMR system, unclassified variants, functional analysis

ABREVIATIONS

HNPCC: Hereditary Non Polyposis Colorectal Cancer

LOH: loss of heterozygosity

MMR: Mismatch Repair System

MSI: Micro Satellite Instability

SNP: single nucleotide polymorphism

VSI: variant de signification inconnue

VUS: variant of unknown significance

INTRODUCTION

Le syndrome de Lynch ou syndrome « HNPCC » (Hereditary Non-Polyposis Colorectal Cancer) est une affection héréditaire autosomique dominante à l'origine de 2 à 8% des cancers colo rectaux (CCR).

HISTORIQUE ET DEFINITIONS

La première description de ce syndrome date de 1913 (Warthin, 1913, Warthin, 1925). Aldred Warthin était le premier à décrire une famille, dite « famille G » dans laquelle le nombre anormal de cancers du côlon, de l'endomètre et de l'estomac dans les générations successives, laissait penser qu'une prédisposition héréditaire existait. Il était également le premier à suggérer le phénomène d'« anticipation », dans lequel l'âge de survenue des cancers diminue au fil des générations. Il fallut attendre 1960 et les travaux d'Henri Lynch pour évoquer le mode de transmission autosomique dominant de cette prédisposition (Lynch et al., 1966). En 1973, Boland et al montraient que deux types de phénotypes existaient, l'un comportant des cancers colo rectaux, et l'autre des cancers extra coliques (Boland et al., 1978 Boland et al., 1984). Les termes de syndrome de Lynch 1 et 2 ont alors été utilisés, puis remplacés par celui de « hereditary non polyposis colorectal cancer », proposé par Lynch lui-même. (Lynch et al., 1985).

En 1991, Vasen et al (Vasen et al., 1991) énonçaient des critères, dits d'Amsterdam I, définissant ce syndrome sur le plan clinique :

- trois apparentés atteints de cancer colorectal (CCR) histologiquement prouvés, un des apparentés devant être lié au premier degré avec les deux autres,
- au moins deux générations successives atteintes,
- un des cancers diagnostiqué avant l'âge de 50 ans,
- exclusion de la polypose rectocolique familiale.

Parallèlement, se créait un consortium intitulé “International Collaborative Group on Hereditary NonPolyposis Colorectal Cancer », plus tard reconnu sous le nom de “International Society for Gastrointestinal Hereditary Tumors”, ou InSiGHT, qui continue de se réunir tous les 2 ans (www.insight-group.org). Cette société œuvre pour la promotion de la recherche translationnelle dans le domaine des prédispositions héréditaires aux tumeurs digestives et constitue actuellement la plus importante base de données internationale des variants de syndrome de Lynch.

La compréhension des mécanismes de carcinogenèse a progressé avec la constatation en 1990 de pertes de segments d’ADN dans les tumeurs se développant dans le cadre d’une prédisposition de type HNPCC, ou « pertes d’hétérozygoties » (« LOH »). C’est M. Perucho qui le premier, en 1992, constata la présence de séquences anormalement répétées, appelées microsatellites, pouvant correspondre un mécanisme différent de carcinogénèse (Peinado et al., 1992) dans lequel des cancers colo rectaux étaient associés à une instabilité des microsatellites (« MSI » microsatellite instability).

Le rapprochement s’est fait rapidement avec les précédents travaux réalisés chez la bactérie et la levure, où un système de réparation de l’ADN appelé MMR (« Mismatch Repair System ») avait été découvert et dont le dysfonctionnement était associé à une instabilité des microsatellites.

Grâce à l’étude de marqueurs microsatellites, des analyses de liaison ont pu être réalisées et les résultats de ces travaux ont suggérés la présence d’un gène suppresseur de tumeur dans le chromosome 2p (Peltomaki et al., 1993).

Très rapidement, plusieurs équipes ont cloné et publié les gènes *MSH2* (Leach et al., 1993), *MLH1* (Bronner et al., 1994), *PMS2* et *PMS1* (Nicolaidis et al., 1994), et finalement *MSH6* (Miyaki et al., 1997).

Ainsi, les deux notions fondamentales et caractéristiques du syndrome de Lynch étaient définies : le lien entre une prédisposition héréditaire aux cancers colo rectaux et un système de réparation des mésappariements de l’ADN (MMR) déficient, et sa conséquence sur le génome, identifiable grâce à la présence d’une instabilité des microsatellites.

Les patients atteints de syndrome de Lynch ont un risque également élevé de développer d'autres cancers, en particulier de l'endomètre, mais également des adénocarcinomes d'autres organes (ovaire, estomac, intestin grêle, épithélium biliaire, urinaire).

Dans sa description initiale, Lynch avait proposé l'existence de deux syndromes génétiques. Le premier (appelé par la suite syndrome de Lynch I) ne prédisposerait qu'au cancer colorectal, alors que le deuxième (appelé syndrome de Lynch II) prédisposerait également à d'autres types de tumeurs, en particulier au cancer de l'endomètre (Lynch et al., 1985). L'étude de Mecklin et al. (Mecklin et al., 1991) a permis d'affirmer qu'il n'existait qu'un seul syndrome de Lynch et non deux. Les auteurs ont comparé la fréquence des différents sites des tumeurs observées entre deux types de familles, celles où l'on trouve des cancers de l'endomètre et celles où l'on n'en trouve pas. Ils n'ont trouvé aucune différence pour le panel (large) des tumeurs envisagées. Dans la mesure où il n'y a pas eu de comparaison à la population générale, cette étude ne peut pas nous servir à définir le spectre. Le même défaut rend inutilisable l'étude de Vasen et al. (Vasen et al., 1996). L'étude de Watson et Lynch (Watson et Lynch. 1993) est la première qui a tenté réellement de définir un spectre avec une méthodologie appropriée. Les familles ont été recensées avec des critères proches d'Amsterdam (au moins 3 cas de cancer du côlon ou de l'endomètre dont au moins 2 cancers du côlon avant 50 ans). Les nombres attendus de cancers ont été calculés pour 1317 sujets dits "à haut risque" (sujets atteints ou parents du 1^{er} degré d'un sujet atteint de cancer du côlon ou de l'endomètre) à partir des fréquences dans la population générale (obtenues à partir du registre du Connecticut). Ils ont alors comparé les nombres observés aux nombres attendus pour différentes localisations de cancers. Les résultats sont indiqués dans le tableau 1, ainsi que ceux de Vasen et al. et de Aarnio et al. (Aarnio et al., 1999) qui procèdent selon la même méthodologie.

Tableau 1 : Risques relatifs (et degré de signification¹) de tumeurs associées au syndrome HNPCC dans 3 études (*seules sont indiquées les tumeurs pour lesquelles le risque relatif a été trouvé significatif dans au moins une étude*)

Site de la tumeur	Watson et Lynch, 1993		Vasen et al., 1996 ²		Aarnio et al., 1999	
	RR	Signification	RR	Signification	RR	Signification
Estomac	4,1	p< 0,001	10,5	p< 0,001	6,9	p< 0,001
Intestin grêle	25,0	p< 0,001	166,7	p< 0,001	-	-
Système hépato-biliaire	4,9	p< 0,05	-	-	9,1	p< 0,05
Rein	3,2	p< 0,01	-	-	4,7	p< 0,05
Vessie	1,1	NS	-	-	7,6	p < 0,01
Uretère	22,0	p< 0,001	53,2	p< 10-8		
Ovaire	3,5	p< 0,001	4,1	NS	13	p< 0,001
Poumon	12,0	p< 0,01	-	-	1,0	NS
Cerveau	1,6	NS	-	-	4,5	p< 0,05

¹ Il s'agit de la valeur nominale du risque de 1^{ère} espèce, donc non corrigé pour les comparaisons multiples

² Les risques relatifs (RR) et degré de signification ont été calculés à partir des valeurs données dans l'article

Ce sont les risques relatifs qui doivent permettre de définir le spectre des tumeurs associées au syndrome HNPCC car ils ont la meilleure valeur prédictive de l'existence d'une mutation d'un des gènes MMR. Si l'on admet les valeurs suivantes du risque relatif pour classer la valeur prédictive positive attendue :

- $RR < 5$: valeur prédictive positive faible
- $5 \leq RR < 8$: valeur prédictive positive moyenne
- $RR \geq 8$: valeur prédictive forte,

et que l'on examine les valeurs données dans le tableau 1, il est logique "d'admettre" dans le spectre étroit le cancer de l'endomètre, de l'intestin grêle et de l'urothélium (bassin et uretère), dans le spectre large l'ovaire, l'estomac et le système hépato-biliaire, et d'exclure toutes les autres localisations.

En 1999, les critères d'Amsterdam I ont été élargis pour intégrer ces cancers extra-coliques à la définition clinique du syndrome de Lynch, et sont devenus les critères d'Amsterdam II (Vasen et al., 1999) :

Critères d'Amsterdam	Critères d'Amsterdam modifiés
Trois membres de la famille atteints de cancer du côlon dont un au premier degré par rapport aux autres	Trois membres de la famille atteints de cancer colorectal héréditaire sans polypose dont un au premier degré par rapport aux autres
Deux générations atteintes	Deux générations atteintes
Un des cancers diagnostiqué avant l'âge de 50 ans.	Un des cancers <u>du spectre HNPCC</u> avant l'âge de 50 ans.
Absence de polypose colique dans la famille	Absence de polypose colique dans la famille

Des tumeurs cutanées et cérébrales ont été également rapportées dans des familles atteintes de syndrome HNPCC, qui portent alors les noms de syndromes de Muir-Torre et Turcot (Hamilton et al., 1995), sans qu'il s'agisse de prédispositions génétiques distinctes.

En 2011, le spectre des cancers a été revu par Bonadona et al (Bonadona v. 2011) dans le cadre de l'étude ERISCAM (*Estimation des Risques de Cancer chez les porteurs de mutations des gènes MMR*). Dans ce travail, où l'ensemble des consultations d'oncogénétique du réseau français a participé, 537 familles informatives de syndrome de Lynch (10283 membres) ont été incluses. Un total de 303 mutations distinctes a été identifié (127 pour *MLH1*, 151 pour *MSH2* et 25 pour *MSH6*). Les risques cumulés de cancer étaient les suivants :

	50 ans	70 ans
Hommes	18% (95%IC 12-27)	45% (95%IC 32-59)
Femmes	19% (95%IC 14-28)	54% (95%IC 41-70)

Les risques cumulés de cancers étaient nuls en dessous de 30 ans, quel que soit l'organe. Par organe, en fonction du gène muté, le colon, l'endomètre et l'ovaire (et à moindre degré l'estomac) étaient les organes les plus à risques, contrairement à, l'urothélium, l'intestin grêle et les voies biliaires, où le risque était inférieur à 1% à 70 ans.

	Risque cumulé, % (IC 95%)											
Age	CCR				Cancer de l'endomètre				Cancer de l'ovaire			
	Total	MLH1	MSH2	MSH6	Total	MLH1	MSH2	MSH6	Total	MLH1	MSH2	MSH6
30	2(1-3)	1(0-3)	2(1-5)	0(0-1)	0(0-1)	0(0-1)	0(0-1)	0	0	0	0(0-1)	0
50	13(9-19)	14(8-27)	20(13-30)	3(2-6)	8(4-15)	9(3-19)	8(3-21)	3(1-8)	3(1-5)	4(0-11)	4(1-9)	0(0-1)
70	35(25-49)	41(25-70)	48(30-77)	12(8-22)	34(16-58)	54(20-80)	21(8-77)	16(8-32)	8(2-37)	20(1-65)	24(3-52)	1(0-3)

LES CRITERES DE LA CONFERENCE DE BETHESDA

En 1996, une conférence sur le syndrome de Lynch, soutenue par le National Cancer Institute (NCI) s'est tenue à Bethesda. L'objectif était de définir une nouvelle stratégie de détection des individus atteints, en intégrant aux données des critères d'Amsterdam, d'autres moins restrictives, telles que la présence d'adénomes ou l'histologie (Rodriguez-Bigas et al., 1997). Les auteurs ont ainsi défini **dans quelles situations cliniques, un statut MSI devait être recherché** :

- la présence de cancers primitifs multiples chez un individu,
- la présence d'adénome colorectal isolé avant 40 ans (ou de cancer avant 45 ans),
- certaines caractéristiques histologiques (faible différenciation, cellules en bagues à chaton).

Ces critères ont fait l'objet d'une révision en 2004 (Umar A. 2004), précisant le panel des marqueurs microsatellites utilisés et le seuil défini pour parler d'instabilité, les situations cliniques conduisant à une recherche d'instabilité, et finalement, des recommandations de prise en charge des patients, de la clinique jusqu'à l'analyse constitutionnelle, en passant par le conseil génétique. (cf. tableaux 2 et 2bis et annexe 1)

Situations devant faire l'objet d'une évaluation du statut MMR
Cancer colon avant 50 ans
Cancer synchrone ou métachrone colo rectal ou du spectre HNPCC quelque soit l'âge
CCR avec instabilité des microsatellites (MSI-H) avant 60 ans
CCR avec ATCD au 1 ^{er} degré de cancer du spectre HNPCC dont un avant 50 ans
CCR avec 2 ou plus cancers du spectre HNPCC au 1 ^{er} ou 2 ^{ème} degré quelque soit l'âge

Tableau 2 : critères révisés de Bethesda (Umar A et al., 2004)

Stratégie d'analyse moléculaire du syndrome de Lynch
Une recherche d'instabilité des microsatellites doit précéder l'analyse constitutionnelle des gènes <i>MSH2/MLH1</i> en cas de perte d'expression d'une de ces deux protéines
Après qu'une mutation a été identifiée, les apparentés doivent être référés à une consultation d'oncogénétique
En cas d'impossibilité des réaliser les analyses sur la tumeur, l'analyse constitutionnelle est une alternative
En l'absence de mutation retrouvée chez un cas index présentant un statut MSI-H avec histoire familiale, un conseil génétique doit être proposé ainsi que des recommandations de suivi identiques à celles du syndrome de Lynch

Tableau 2bis : critères révisés de Bethesda (Umar A et al., 2004)

LIMITES DES CRITERES D'AMSTERDAM ET DE BETHESDA

Les critères d'Amsterdam modifiés et ceux de Bethesda ont fait l'objet d'une évaluation de leurs performances car il apparaissait qu'un certain nombre de familles atteintes et porteuses d'une mutation délétère ne remplissait pas ces critères (Syngal et al., 2000). Une méta analyse a analysé l'ensemble des travaux (15 pour Amsterdam et 4 pour Bethesda). Les auteurs ont démontré une sensibilité globale de 78% pour les critères d'Amsterdam, signifiant que **22% des patients porteurs d'une mutation constitutionnelle n'étaient pas repérés**. La sensibilité des critères de Bethesda était de 89%. Les spécificités étaient de 46-68% et 53% respectivement. Les probabilités « post test » d'un test positif (probabilité pour qu'une personne remplissant les critères

soit porteuse d'une mutation délétère) étaient insuffisantes (0.46 et 0.41), incitant les auteurs à penser que ces critères étaient insuffisants pour prédire une mutation constitutionnelle de syndrome de Lynch (Kievit W et al., 2004)

GENETIQUE DU SYNDROME DE LYNCH

Les gènes « principaux »

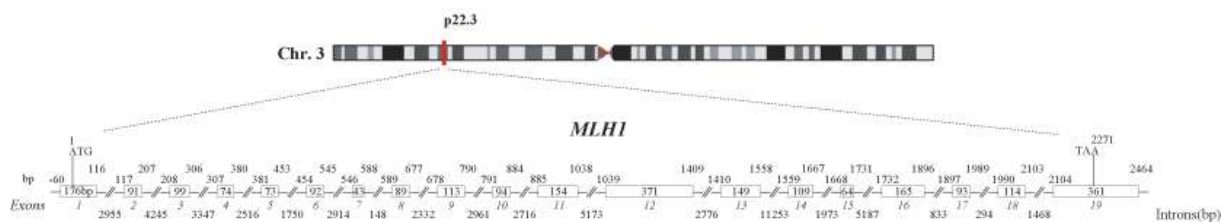
Les études de liaison génétique ont d'emblée révélé l'hétérogénéité génétique de ce syndrome en identifiant 2 locus morbides, 2p15-16 (plus tard redéfini en 2p22-21) et 3p21 (Peltomaki et al., 1993 ; Lindblom et al., 1993). Les gènes correspondant à ces locus morbides, appelés respectivement *MSH2* (MutS homolog 2, 16 exons) et *MLH1* (MutL homolog 1, 19 exons) ont pu être rapidement identifiés grâce à leur homologie avec les gènes bactériens MutS et MutL, impliqués dans la réparation des mésappariements de l'ADN (Leach et al., 1993 ; Fishel et al., 1993 ; Bronner et al., 1994 ; Papadopoulos et al., 1994). Dans des familles atteintes de syndrome de Lynch, des mutations délétères constitutionnelles ont été également identifiées dans le gène *MSH6* (2p16, 10 exons), autre homologue de MutS (Miyaki et al., 1997 ; Akiyama et al., 1997). Les études de liaison puis les études des mutations délétères ont montré que plus de la moitié des cas de syndrome de Lynch était due à des mutations constitutionnelles de *MSH2* et *MLH1* (Peltomaki et al., 1997 ; Wijnen et al., 1998).

En France, il n'a pas été observé d'effet fondateur majeur actuellement (Parc et al., 2003). Malgré l'existence de quelques mutations récurrentes, la très grande diversité des mutations limite la mise en place d'une approche orientée. Il n'existe pas actuellement de technique d'analyse simple et exhaustive, mais des améliorations sont attendues dans ce domaine, en termes d'automatisation tant expérimentale qu'analytique. Environ 20% des familles, ayant un syndrome de Lynch typique sans mutation délétère détectable de *MSH2* et *MLH1*, ont un remaniement génomique du gène *MSH2* impliquant un ou plusieurs exons et le plus souvent la région promotrice (Charbonnier et al., 2002). Le gène *MSH6* serait impliqué dans environ 4-5% des familles, de même que les remaniements génomiques du gène *MLH1*.

1. *MLH1* (Mut L Homolog 1)

Gène :

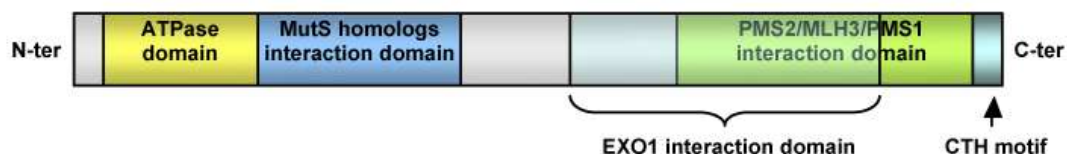
MLH1 (NM_000249) est un gène comportant 19 exons codant pour une protéine de 756 acides aminés. Ce gène est l'homologue humain de MutL qui est un gène de réparation de l'ADN du système MMR de *Escherichia Coli*. Il est localisé en 3p21.3.



Protéine:

La protéine *MLH1* fait partie du système MMR et forme un hétéro dimère avec la protéine du gène *PMS2* (*PMS1* protein homolog 2) afin de coordonner l'accès des autres protéines impliquées dans la réparation de l'ADN, telles que les hélicases, la protéine codée par *EXO 1*, *PCNA* (proliferating cell nuclear antigen), *RPA* (single-stranded-DNA binding-protein), et les DNA polymérases [Peltomaki et al., 2003]. Dans le syndrome de Lynch, les deux allèles de *MLH1* sont inactivés au sein des cellules.

MLH1 protein (756 aa ; 84.6 kDa)



From « Functional significance and clinical phenotype of nontruncating mismatch repair variants of *MLH1* » Raevaara T.E. et al., *Gastroenterology* 2005; 129:537-549
and « Functional Domains of the *Saccharomyces cerevisiae* Mlh1p and Pms1p DNA Mismatch Repair Proteins and Their Relevance to Human Hereditary Nonpolyposis Colorectal Cancer-Associated Mutations » Pang Q. et al., *Molecular and Cellular Biology* 1997; 17:4465-4473

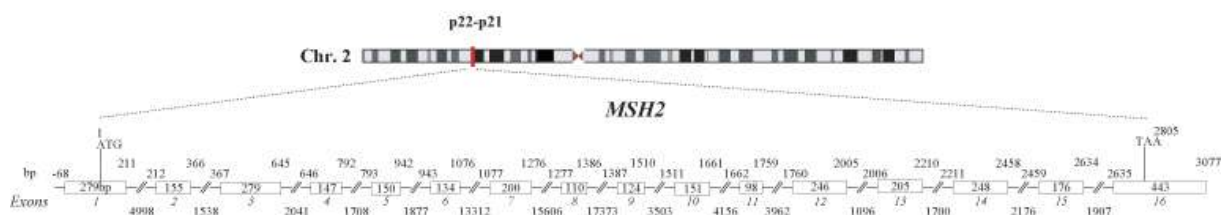
Mutations de *MLH1*:

Plus de 200 différentes mutations ont été décrites, avec 5-10% de délétions [Peltomaki et al., 2003, Peltomaki & Vasen, 2004] et 6211 variants sont actuellement décrits dans la base de données INSIGHT, dont 78.8% de substitutions et 13.9% de courtes délétions ou insertions nucléotidiques.

2. *MSH2* (Mut S Homolog 2)

Gène:

Le gène *MSH2* (NM_000251) comporte 16 exons codant pour une protéine de 934 acides aminés. Il est localisé en 2p21.



Protéine :

La protéine *MSH2* forme un hétéro dimère avec *MSH6*, autre protéine du système MMR, ou *MSH3*, dont la fonction est d'identifier les mésappariements.

MSH2 protein (934 aa ; 104.7 kDa)



From « Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA » Obmolova G. et al., Nature vol.407, 12 octobre 2000, 703-710

« The crystal structure of DNA mismatch repair protein MutS binding to a G×T mismatch » Lamers M.H. et al, Nature vol.407, 12 octobre 2000, 711-717

« Interactions of Human hMSH2 with hMSH3 and hMSH2 with hMSH6 : Examination of Mutations Found in Hereditary Nonpolyposis Colorectal Cancer » Guerrette S. et al., Mol. Cell. Biol. 1998, 18(11):6616

and « A conserved MutS homolog connector domain interface interacts with MutL homologs » Marc L. Mendilloa, Victoria V. Hargreavesa, Jonathan W. Jamisonb, Ashley O. Mob, Sheng Lib, Christopher D. Putnama,b, Virgil L. Woods, Jr.b,1, and Richard D. Kolodnera,b,c,d,1

Mutations de *MSH2*:

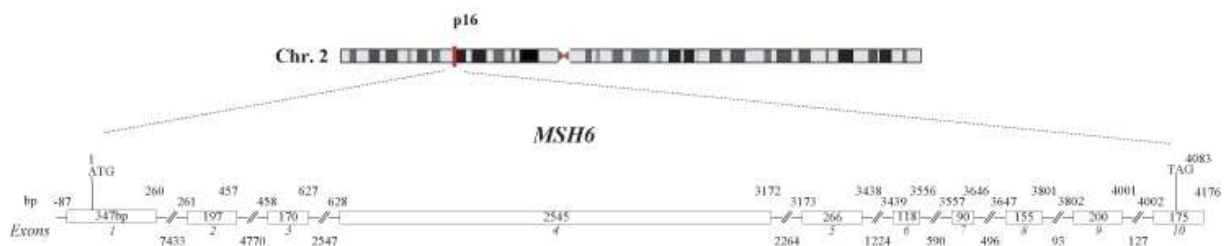
Plus de 170 mutations ont été identifiées, dont au moins 20% sont de grandes délétions [Peltomaki et al., 2003, Peltomaki & Vasen, 2004].

4561 variants sont rapportés dans la base de données INSIGHT (fev 2014), dont 70.5% de substitutions et 23.5% de délétions ou insertions.

3. *MSH6* (Mut S Homolog 6)

Gène:

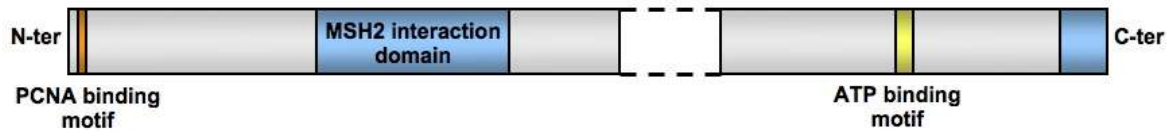
Le gène de *MSH6* (NM_000179) comporte 10 exons codant pour une protéine de 1360 acides aminés.



Protéine :

La protéine code par *MSH6* forme un hétéro dimère avec la protéine *MSH2* dans le système MMR dont les fonctions sont la reconnaissance des mésappariements selon le modèle de « pince coulissante » (« sliding clamp model ») [Fishel et al., 1993, Gruber & Kohlmann, 2003].

MSH6 protein (1360 aa ; 153 kDa)



From « Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome » Kariola R. et al., Human Molecular Genetics 2002; 11:1303-1310

Mutations de *MSH6*:

Plus de 30 mutations ont été identifiées [Peltomaki & Vasen, 2004], principalement de larges délétions.

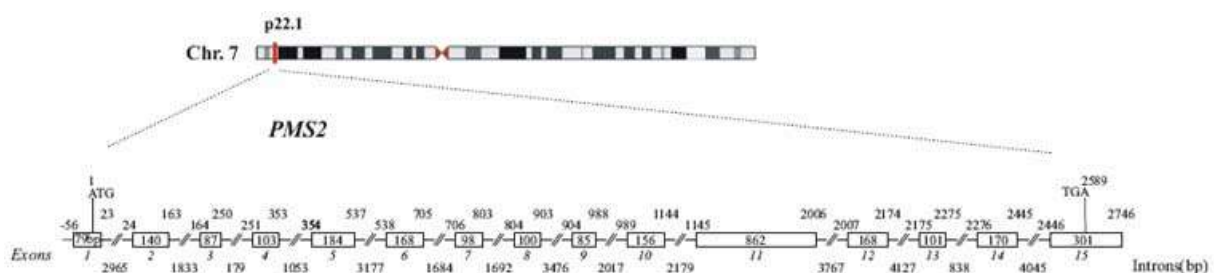
1489 variants sont rapportés dans la base de données INSIGHT (fév 2014), dont 69.9% de substitutions et 15.9% de délétions ou insertions.

4. PMS2 (Post Meiotic Segregation increased 2)

Gène :

Le gène PMS2 comporte 15 exons et code pour une protéine de 862 acides aminés. Il est localisé en 7p22.2.

Plusieurs pseudo gènes ont été identifiés en 7p22, 7p12-13, 7q11, and 7q22 [Nicolaidis et al., 1995].



Protéine :

Homologue humain du Pms1 de la levure, PMS2 forme un hétérodimère avec *MLH1*.

Mutations de PMS2 :

Les mutations germinales sont rares [Hendriks et al., 2006]. Des mutations ponctuelles et grands réarrangements ont été décrits jusqu'à 20%. Cependant ces délétions sont difficiles à rechercher en raison de la présence de pseudo gènes.

La première mutation a été rapportée par Nicolaidis (1994) sur une série de 22 familles porteuses d'un syndrome HNPCC étudiées, en utilisant la méthode de pré-criblage PTT, pour Protein Truncation Test. Il s'agissait d'une délétion en phase qui supprimait les codons 301-381. Dans la tumeur correspondante, une perte d'hétérozygotie avec un phénotype MSI a été observée, suggérant ainsi le rôle causal de cette mutation. Parmi les 50 familles ne présentant pas de mutations pour les gènes *MLH1* et *MSH2*, 5 variants caractérisés par substitutions d'une base ont été identifiés qui correspondaient a priori à des variants sans signification fonctionnelle puisque tous étaient retrouvés dans des témoins avec un taux de 0,5 à 4% (Wang et al., 1999). Une diminution, mais non une abolition, de la capacité d'interaction protéine-protéine avec *MLH1* pour 3 de ces variants (P511K, T597S, M622I) a été documentée (Yuan et al., 2002). Cependant, en absence d'études plus approfondies, le rôle de ces variants dans le développement du syndrome HNPCC reste encore discutable. D'après deux autres études effectuées sur des séries de 22 familles atteintes d'un syndrome HNPCC typique et 84 familles atteintes d'un syndrome HNPCC probable, aucune mutation délétère n'a pu être mise en évidence (Viel et al., 1998; Liu et al., 2001). Les mutations ponctuelles de PMS2 semblent plutôt rares dans des familles atteintes de HNPCC, cependant elles pourraient être associées au syndrome de Turcot. Dans une série de 14 familles atteintes de syndrome de Turcot, Hamilton (1995) a pu identifier 1 mutation non-sens dans PMS2, à côté de 10 mutations dans le gène APC et 1 mutation dans *MLH1*. A l'aide d'un test *in vitro* (Nicolaidis et al., 1998), il a été montré que cette mutation provoquait un effet " dominant négatif " induisant un phénotype MSI dans des cellules normales. Une mutation de type faux-sens a été ensuite décrite chez un autre patient (Miyaki et al., 1997) présentant un niveau de MSI élevé dans ses tissus normal et tumoral. Le caractère délétère de cette mutation reste néanmoins incertain du fait d'une transmission de la mutation par la mère indemne. Enfin, une femme porteuse d'une double mutation tronquante de PMS2 a été identifiée par De Rosa (De Rosa et al., 2000) dans une famille atteinte d'un syndrome de

Turcot. Chacune des mutations provenait cependant d'un parent non atteint, ce qui soulève l'hypothèse d'un effet additif des mutations PMS2 qui, isolément, semblent présenter un effet mineur.

En résumé, seulement 3 patients porteurs de mutations délétères de PMS2 ont été identifiés jusqu'à présent sur plus de 200 cas étudiés, dont 2 atteints du syndrome de Turcot. Une étude plus approfondie pourrait mettre en évidence une corrélation génotype/phénotype.

5. EPCAM (Epithelial Cell Adhesion Molecule)

Le gène d'EPCAM comporte 9 exons codant pour une protéine de 314 acides aminés.

Des délétions du signal de transcription terminal d'EPCAM sont impliquées dans 1 à 2.8% des familles atteintes de syndrome de Lynch [Sivagnanam et al., 2010] et surviennent par recombinaison médiées par des séquences répétées courtes de type SINE ou ALU [Kuiper et al 2011]. Ligtenberg et al. ont montré qu'une mutation de l'extrémité 3' du gène EPCAM (exons 8 et 9) pouvait être responsable d'un syndrome de Lynch en entraînant une extinction constitutionnelle de l'expression de *MSH2* par méthylation de son promoteur, situé en amont du gène (Cf. Figure 1)

Par ailleurs, l'expression d'EPCAM varie selon les tissus. De hauts niveaux sont retrouvés dans les cellules souches colo rectales alors qu'ils sont faibles dans les leucocytes (Ligtenberg et al., 2009).



Figure 1 : co localisation des gènes EPCAM et *MSH2*

Les gènes « secondaires »

Une mutation délétère constitutionnelle d'un des gènes MMR principaux est à l'origine d'environ 70% des cas de syndrome de Lynch. La cause des autres cas est par contre peu ou pas connue. De nombreuses études ont été réalisées afin d'identifier de nouveaux gènes, parmi deux groupes de gènes candidats : des partenaires des gènes *MLH1* ou *MSH2* participant au processus de la réparation des mésappariements de l'ADN (*MLH3*, *MSH3*, *PMS1*, *EXO1*, *MBD4/MED1*), et des gènes qui jouent un rôle important dans la tumorigenèse des cancers colorectaux et pour lesquels des mutations somatiques avaient été décrites (*TGFβRII*).

1. Gène *MLH3*

Homologue humain de *Mlh3* chez la levure, *MLH3* est également un partenaire de *MLH1*. Une mutation tronquante et 12 variants (substituant un nucléotide) ont été décrits (Wu et al., 2001). Aucun de ces variants n'a été retrouvé parmi 188 témoins, ce qui a conduit à écarter l'hypothèse de variants non délétères. Cependant une autre étude, conduite par Lipkin (Lipkin et al., 2001) n'a identifié aucune mutation constitutionnelle de *MLH3* parmi 60 familles. Afin d'évaluer le rôle de *MLH3*, 112 patients, porteurs ou non d'anomalies des gènes *MLH1* et *MSH2*, ont été analysés pour *MLH3* et 8 variants ont été identifiés, dont 2 chez des patients par ailleurs porteurs d'une mutation délétère de *MLH1* ou de *MSH2* (Wang, données non publiées). L'indisponibilité de tests fonctionnels et l'impossibilité d'étudier la coségrégation de ces variants ont rendu l'interprétation équivoque. Plus récemment, 1 mutation tronquante et 11 variants ont été trouvés dans une série de 70 familles, parmi lesquelles 3 sont présumées délétères du fait de la coségrégation du variant avec la pathologie et son absence dans la population générale (Liu et al., 2003).

Il est donc clair que l'altération du gène *MLH3* ne joue pas un rôle prépondérant dans le syndrome HNPCC. Néanmoins, la pénétrance faible de certains variants étant suggérée par différentes études (Wang, données non publiées; Liu et al., 2003), leurs conséquences fonctionnelles méritent d'être évaluées, en adaptant la méthode décrite par Trojan et al. (2002).

2. Gène MSH3

Partenaire du gène *MSH2*. Aucune mutation constitutionnelle n'a été jusqu'à présent mise en évidence dans le cadre du syndrome de Lynch.

3. Gène PMS1

Partenaire du gène *MLH1*, PMS1 ne semble pas jouer de rôle important dans la réparation post-réplivative de l'ADN chez les mammifères. Jusqu'à présent, une seule mutation constitutionnelle a été décrite par Nicolaides (Nicolaides et al., 1994) caractérisée par une délétion en phase de la séquence codante. Cependant une délétion génomique du gène *MSH2* a été identifiée par la suite chez le même patient, remettant en cause le caractère délétère de la mutation de PMS1 (Liu et al., 2001). Aucune mutation délétère n'a été mise en évidence parmi 50 patients sans mutation des gènes principaux ; 5 variants, présents avec une fréquence de 0.5% dans la population générale, ont été trouvés (Wang, données non publiées).

4. Gène EXO1 (5'-3' exonucléase)

EXO1 est l'homologue humain de l'exonucléase 1 de la levure. Chez la levure, l'exonucléase 1 interagit avec *MLH1*, *MSH2* et Msh3 et participe à la réparation de l'ADN en intervenant dans la catalyse de l'excision d'ADN de 5' en 3'. Une fonction similaire chez l'homme a été décrite dans un système *in vitro* (Genschel et al., 2002). Une mutation sur un site d'épissage et 8 variants ont été décrits chez 13 patients, pour la plupart porteurs d'une forme atypique de syndrome de Lynch (Wu et al., 2001). L'étude des tumeurs correspondantes a montré un phénotype MSI dans 6 cas. De manière surprenante, une perte de l'allèle mutant et une rétention de l'allèle normal ont été constatées dans 12 tumeurs. Les tests fonctionnels *in vitro* effectués par Sun (Sun et al., 2002) semblent montrer une perte partielle de la fonction pour 5 variants. Afin d'évaluer le rôle de l'EXO1, Jagmohan-Changur (Jagmohan-Changur., 2003) a publié une étude multi-populations européennes regroupant un total de 970 cas et 1007 témoins. Il

a montré que parmi ces 9 variants, 5 sont vraisemblablement non délétères, puisque présents chez les témoins avec la même fréquence ; 3 variants supplémentaires ont été identifiés, avec absence de perte allélique dans les tumeurs correspondantes.

L'implication d'EXO1 dans le syndrome HNPCC semble donc secondaire. Cependant l'hypothèse de mutations de "pénétrance faible" reste à évaluer.

5. MBD4/MED1 (methyl-CpG binding endonuclease 1)

Ce gène récemment cloné code pour un médiateur de *MLH1* qui participe à la réparation de l'ADN. Des mutations somatiques ont été détectées dans des tumeurs colorectales présentant un phénotype d'instabilité microsatellitaire (Ricchio et al., 1999). Aucune mutation constitutionnelle n'a été mise en évidence dans une série de 120 familles répondant aux critères stricts ou proches définissant le syndrome HNPCC (Wang, données non publiées).

6. Gène *TGFβRII* (récepteur de type II au TGF beta)

Le *TGFβRII* joue un rôle important dans le contrôle de la prolifération cellulaire. Des mutations somatiques de *TGFβRII*, caractérisées par une délétion ou une insertion d'une nucléotide dans une séquence codante et mononuléotidique (A)₁₀, sont des événements très fréquents et représentent ainsi une "signature" d'une tumeur MSI. Ceci suggère que *TGFβRII* est une cible préférentielle de l'instabilité du génome et que les mutations retrouvées sont associées à un avantage de croissance des cellules. Contrairement aux mutations somatiques, les mutations constitutionnelles de *TGFβRII* semblent être des événements rares. Jusqu'à présent, une seule mutation constitutionnelle a été décrite (Lu et al., 1998). Il s'agissait d'une mutation de type faux-sens, chez une famille atypique de HNPCC, avec un phénotype MSS dans la tumeur. Il a été néanmoins montré, en utilisant un système *in vitro*, que cette mutation affectait la fonction de la protéine. Une autre étude a été ensuite réalisée par Shin (Shin et al., 2000) sur 35 patients porteurs d'un syndrome HNPCC, 44 cas suspectés plus 45 cas sporadiques jeunes. Aucune mutation délétère n'a pu être mise en évidence.

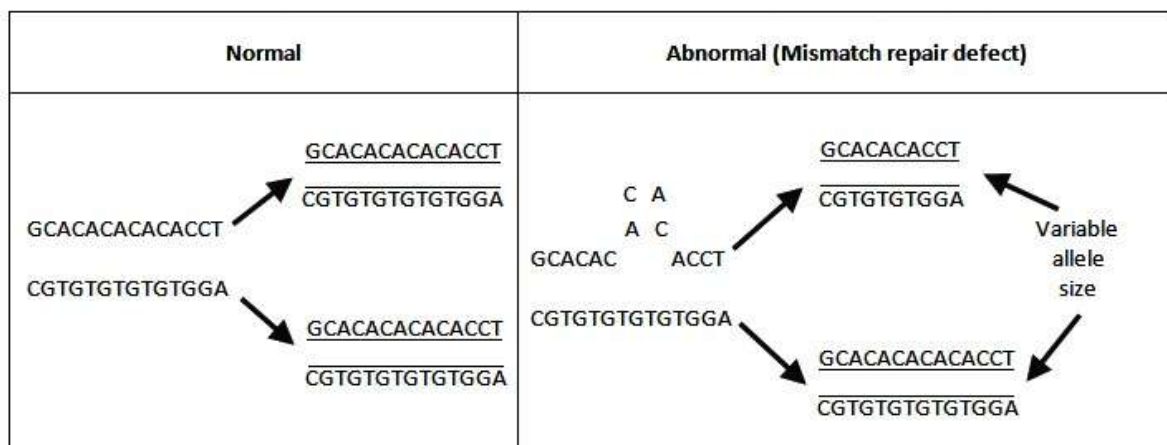
Système MMR

INSTABILITE DES MICROSATELLITES

La grande majorité des cellules déficientes dans leur système MMR développent un phénotype hyper muté (« mutator phenotype »), caractérisé par une augmentation de l'ordre de 10^2 - 10^3 du taux de mutations spontanées (Loeb et al., 1991). Ces mutations affectent l'ensemble du génome, mais en particulier les séquences microsatellites (MS). L'instabilité des MS de certains gènes qui les contiennent entraîne des anomalies des voies de signalisation de la transduction, la régulation de la transcription et de l'apoptose. Ainsi une instabilité des MS intra génique inactive le gène du TGF β -RII dans 80% des tumeurs dont le système MMR est déficient (Markowitz et al., 1995). La présence d'un haut niveau d'instabilité (MSI-H) est associée à une mutation des gènes *hMLH1* et *hMSH2* (Frayling et al., 1999). Un niveau bas d'instabilité (MSI-L) est plus volontiers associé à une mutation du gène *hMSH6* (10%) et *hPMS2* (5%) (Dietmaier et al., 1997).

La grande majorité des tumeurs du syndrome de Lynch présente une instabilité élevée (95%) contrairement aux cancers colo rectaux sporadiques qui ont un tel phénotype dans 10-15% des cas (Boland et al., 1998). Le diagnostic d'une instabilité des microsatellites est ainsi devenu l'indicateur d'un système MMR déficient.

Microsatellite Replication



Adapted from Gruber SB, Kohlmann W. The genetics of hereditary nonpolyposis colorectal cancer. *J Natl Cancer Net.* 2003;1:137-44.

Figure 1 : instabilité des microsatellites (Wendy Kohlmann. GeneReviews™)

LE SYSTEME MMR

Ce système permet une réparation post répllicative de l'ADN en corrigeant les erreurs commises par les DNA polymérases (Pol) epsilon ou delta au cours de la réplication du génome. Il s'agit d'erreurs d'appariement base-base causées par des incorporations accidentelles, des insertion/délétions, survenant notamment dans les séquences répétées des microsatellites. Le système MMR protège ainsi les cellules de mutations spontanées. Le processus comporte quatre étapes (annexes 2 et 3):

1. Le mésappariement est reconnu par l'hétéro dimère MutS alpha, formé par les protéines *MSH6* et *MSH2*, ou par l'hétérodimère MutS bêta, formé par *MSH2* et *MSH3*.
2. Cette reconnaissance entraîne le recrutement d'un autre hétéro dimère MutL alpha (*MLH1*-*PMS2*).
3. L'activité endonucléasique de *PMS2* génère la création d'une incision dans le brin complémentaire, va permettre à la 5'exonucléase (*EXO 1*) d'exciser le brin contenant le mésappariement.
4. Pour finir, une re-synthèse de la portion excisée est opérée grâce à la ADN polymérase.

CORRELATIONS GENOTYPE-PHENOTYPE

Environ 90% des cancers survenant dans le cadre d'un syndrome de Lynch sont dus à une mutation de *MLH1* ou *MSH2*. Seuls 10% sont associés à une mutation de *MSH6* ou *PMS2* (Lynch H et al., 2003)

Le risque de cancer varie avec les gènes concernés. Ainsi, une mutation de *MSH2* est associée à un risque élevé de cancers extra coliques, alors que ceux-ci (à l'exception des cancers de l'endomètre), sont rarement retrouvés avec les mutations de *MSH6*, *PMS2* ou *EPCAM*.

Les mutations de *MSH2* ont été plus fréquemment décrites que les 3 autres gènes chez les patients atteints du syndrome de Muir-Torre (South et al 2008).

Les mutations de *MSH6* sont plus fréquemment associées à un statut MSI-L. Les cancers surviennent plus tard et sont classiquement plus distaux que dans les syndromes de Lynch résultants des mutations des autres gènes MMR (Wu et al., 1999, Berends et al., 2002). Le cancer de l'endomètre est assez communément retrouvé, même si des risques moindres de cancers colo rectal et de l'endomètre ont été rapportés [Berends et al., 2002, Baglietto et al., 2010].

Les mutations du gène *PMS2* sont associées à un risqué plus faible de cancers du spectre du syndrome de Lynch (25%-32%) (Senter et al., 2008).

Pénétrance

La pénétrance des mutations du syndrome de Lynch du CCR dans le syndrome de Lynch est difficile à préciser et est probablement inférieure à 100%. Une étude de Lin et al. rapportait une pénétrance de 94-96% pour les hommes (*MLH1/MSH2*) et de 39-63% chez les femmes (Lin et al., 1998). Ainsi, certains individus atteints d'une mutation délétère ne développeront jamais de cancers.

Anticipation

Il n'existe pas d'effet d'anticipation démontré dans le syndrome de Lynch (c'est à dire l'apparition à un âge plus précoce dans les générations suivantes). Une étude suggérant ce phénomène n'a pas été confirmée

FREQUENCE DU SYNDROME DE LYNCH

La prévalence du syndrome de lynch est estimée à 1:440 (Chen et al., 2006). La proportion d'individus porteurs de CCR qui répondent aux critères d'Amsterdam varie de 0.3% à 4.4% selon les études (Riegler et al., 1999 ; Peel et al., 2000 ; Percepese et al., 2001 ; Raedle et al., 2001 ; Hemminki et Li, 2001 ; Katballe et al., 2002 ; Andrieu et al.,

2003), avec une moyenne de l'ordre de 1%. La seule estimation de la proportion de patients porteurs d'une mutation d'un gène MMR, réalisée à partir d'une recherche systématique de mutation chez des patients atteints de CCR présentant une MSI, est de 2,7% (Salovaara et al., 2000). Seulement 17% des patients porteurs de mutation répondaient aux critères d'Amsterdam stricts, confirmant ainsi le manque de sensibilité de ces critères. Cette étude a cependant été réalisée en Finlande où il existe deux mutations fondatrices et dont les chiffres ne sont pas nécessairement extrapolables à la population française.

La proportion de syndrome de Lynch chez les patientes traitées pour cancer de l'endomètre est de 6.6% dans une série hospitalière (Parc et al., 2000).

PHENOTYPE DU SYNDROME DE LYNCH

LE CANCER COLO RECTAL

Les cancers colorectaux qui se développent dans un contexte de syndrome de Lynch présentent quelques éléments non spécifiques mais trouvés plus fréquemment que dans les CCR sporadiques :

- précocité au diagnostic (médiane : 44 ans) ;
- localisation droite (2/3 des cas) ;
- association à un CCR synchrone ;
- caractéristiques histologiques particulières.

De nombreux paramètres ont été analysés, dans les buts de distinguer les tumeurs MSI des tumeurs MSS, mais leur valeur prédictive positive ne dépasse généralement pas 40% (Jass et al., 2002 ; Chapusot et al., 2003). Deux caractéristiques tumorales ont une spécificité supérieure à 0.90 :

- L'infiltration lymphocytaire péri tumorale est plus fréquente dans les tumeurs en relation avec une mutation délétère constitutionnelle d'un gène MMR, que dans

celles présentant 2 mutations délétères somatiques : 19/51(37%) versus 2/18(11%).

- Le type médullaire (OR 37.8) (Alexander et al., 2001)

Cependant, l'identification de ces caractéristiques tumorales nécessite une expertise certaine, et, en raison de leur valeur prédictive médiocre, leur recherche n'apparaît pas adaptée à la pratique médicale courante.

LES ADENOMES

Comme dans le cancer colo rectal sporadique, les adénomes représentent la lésion pré cancéreuse du syndrome de lynch. Les adénomes présentent des caractéristiques légèrement différentes des adénomes sporadiques, à savoir, une réaction inflammatoire lymphoïde « Crohn like », une composante mucineuse, et une faible différenciation. Ils sont, comme le cancer, plus fréquemment localisés au niveau du colon proximal (50 vs 26%)(Rijcken et al., 2002) et de plus petite taille. La fréquence de la dysplasie de haut grade est également significativement plus élevée que pour les adénomes sporadiques (36 vs 13% $p<0.02$. Rijcken et al., 2002), ce qui laisse penser que la séquence adénome-cancer est plus rapide dans le syndrome de Lynch. La prévalence des adénomes est directement liée à l'âge. Ainsi, dans une étude récente de Lillegren et al (Lillegren A et al., 2008) portant sur 695 sujets issus de l'étude CAPP2, la prévalence des adénomes augmentait avec la tranche d'âge (16-34 ans : 5%, 35-44 ans : 12.8%, 45-56 ans : 18.9% $P=0.0001$). On notait également dans cette étude, à moindre degré, un nombre constant de polypes hyperplasiques, lesquels se retrouvaient plutôt dans le colon distal.

CANCER DE L'ENDOMETRE

Comme les CCR, les cancers de l'endomètre qui se développent dans un contexte de syndrome de Lynch présentent quelques éléments non spécifiques mais plus fréquents que les cancers sporadiques :

- âge au diagnostic plus précoce de 15 ans en moyenne (Watson et Lynch, 1993);
- présence d'une hyperplasie endométriale sans atypie (Ichikawa et al., 2002 ; Berends et al., 2001 ; de Leeuw et al., 2000) ;
- présence d'un phénotype MSI (Berends et al., 2001 ; de Leeuw et al., 2000 ; Parc et al., 2000).

Le pronostic des deux formes de cancer de l'endomètre est comparable (Boks et al., 2002).

AUTRES CANCERS ASSOCIES AU SYNDROME DE LYNCH

Intestin grêle

Lorsque l'intestin grêle est le siège d'adénocarcinomes (1-4% des risques de cancers dans le Lynch), le duodénum est atteint dans 36% des cas (Rodriguez-Bigas et al., 1998). Par analogie, 6% des carcinomes de la papille sont associés à un génotype MMR (Agaram et al., 2010). Récemment la réalisation de vidéo capsule du grêle a permis de détecter 8.6% d'adénocarcinomes et d'adénomes (Saurin et al., 2010).

A ce jour, la prévalence des adénomes spécifiquement duodénaux reste rare et non publiée. Dans le syndrome de Lynch, les cancers du grêle sont de diagnostic plus précoce (âge moyen au diagnostic : 49 ans), majoritairement des adénocarcinomes (plus de 95%), et leur pronostic serait meilleur (Rodriguez-Bigas et al., 1998).

Voies urinaires

Les cancers des voies urinaires hautes à cellules transitionnelles sont plus fréquents que dans la population générale, de survenue plus précoce et le ratio homme/femme est plus équilibré (Sijmons et al., 1998).

Ovaire

Le cancer de l'ovaire est de survenue plus précoce (45 en moyenne). La forme épithéliale domine (plus de 90%) dont environ 5% de tumeurs "borderline". Dans les formes invasives, les tumeurs sont de stade plus précoce au diagnostic et assez ou bien différenciées. La présence d'un cancer de l'endomètre synchrone du cancer de l'ovaire est plus fréquente (Watson et al., 2001).

Estomac

La fréquence de cette tumeur est plus élevée dans les générations les plus anciennes et tend à diminuer dans les générations les plus jeunes dans la même proportion qu'en population générale (Lynch et Smyrk, 1996).

Voies biliaires

Les cancers du tractus biliaire regroupent les cancers de l'ampoule de Vater, les cancers des voies biliaires extra-hépatiques et de la vésicule biliaire. Ils sont le plus souvent étudiés ensemble. Dans les syndromes HNPCC, c'est le cholangiocarcinome qui est admis comme faisant partie du spectre tumoral (Mecklin et al., 1992).

LA CONSULTATION D'ONCOGENETIQUE ET LES CRITERES D'INDICATION

Indications

Il n'existe aucun critère pathognomonique permettant, à partir de la seule présentation clinique d'un cancer, d'identifier une prédisposition héréditaire et le gène responsable. Cependant, il existe des arguments d'orientation de trois ordres :

- individuels : diagnostic à un âge inhabituel, existence de cancers primitifs multiples. Ces informations sont généralement fournies par le compte-rendu d'hospitalisation ou de consultation spécialisée ;
- familiaux : nombre de cancers, localisation, distribution dans l'arbre familial. A cette étape, la vérification des données anamnestiques est recommandée ;
- tumoraux : caractéristiques histologiques, résultats des tests immunohistochimiques ou biologiques des cellules tumorales, d'où la nécessité de disposer systématiquement des comptes-rendus d'anatomie pathologique.

La logique médicale retenue pour définir des indications de consultations d'oncogénétique est la probabilité de l'existence d'une mutation délétère. Le choix initial des critères d'Amsterdam, comme définition clinique du syndrome de Lynch, avait pour principal objectif de sélectionner des familles très évocatrices de la ségrégation d'un gène majeur, pré-requis indispensable à l'identification des gènes impliqués dans une maladie génétique complexe. En revanche ces critères sont trop restrictifs et trop peu sensibles en pratique médicale courante. En particulier, ces critères ignorent les cas de mutation délétère constitutionnelle d'un gène MMR à expression tardive, ou ségrégeant dans des familles de petite taille, ou encore les mutations de novo, qui peuvent se manifester par des cancers primitifs multiples chez un même patient.

Progressivement, l'établissement du diagnostic de syndrome de Lynch, pour lequel la prise en charge est codifiée a intégré la prise en compte de localisations tumorales autres que le côlon et le rectum, et la reconnaissance du phénotype MSI comme caractéristique obligatoire des cancers se développant dans un contexte de mutation délétère constitutionnelle d'un des gènes MMR principaux. Il existe par contre des agrégations familiales de CCR répondant aux critères d'Amsterdam I avec tumeurs MSS, dont les gènes responsables sont encore inconnus (Rovella et al., 2001, Jenkins et al., 2002). Du fait de l'absence de connaissance des risques de cancer dans ces familles (localisations en particulier), seuls les conseils de prise en charge de la localisation colique apparaissent justifiables.

L'indication d'une consultation d'oncogénétique sera retenue d'emblée pour tous les patients répondant aux critères d'Amsterdam II élargis (2 apparentés au premier degré au minimum, et non 3). Les unités de consultation multidisciplinaire (UCP) de cancérologie pourront dans les autres configurations adopter le schéma suivant :

- adresser d'emblée tous les patients dont la probabilité de cancer de type MSI est supérieure à 50% (à savoir : CCR avant 40 ans, antécédent personnel de cancer colorectal ou de l'endomètre) ; en l'absence d'information concernant les valeurs prédictives des antécédents de cancer du spectre large du syndrome de Lynch, les experts proposent de les prendre en compte de la même manière.
- parmi les autres patients, sélectionner sur le caractère MSI, ceux dont l'âge est compatible avec une absence de méthylation de l'ADN liée à la sénescence (avant 60 ans) (Nakagawa et al., 2001), ou ayant un antécédent au premier degré de cancer du spectre large.

Ne seront pas retenus comme ayant un risque fort de CCR, génétiquement déterminé, les cas de CCR MSS après 40 ans, ni les cas de CCR MSI isolé après 60 ans, en raison du phénomène de méthylation lié à la sénescence. Dans l'impossibilité d'obtenir une détermination du phénotype tumoral, les CCR diagnostiqués entre 40 et 60 ans seront discutés par l'UCP de cancérologie.

Le thesaurus de cancérologie digestive de la Société Nationale Française de Gastro Entérologie (SNFGE) a intégré ces recommandations de la façon suivante (<http://www.snfge.org/data/ModuleDocument/publication/5/pdf/TNCD-chapitre-3.pdf>):

Dépistage des formes familiales :

Interrogatoire :

- *Recherche de cancers épidémiologiquement liés : endomètre, ovaire, grêle, uretère ou cavités excrétrices rénales,*
- *Recherche d'antécédents familiaux de cancers.*

Syndrome de Lynch (HNPCC) : présence de tous les critères d'Amsterdam II

- *Au moins 3 sujets atteints de cancers du spectre HNPCC (côlon-rectum, endomètre, ovaire, grêle, uretère ou cavités excrétrices rénales) dont 1 uni aux 2 autres au premier degré,*
- *Au moins 2 générations successives concernées,*
- *Au moins 1 cancer diagnostiqué avant l'âge de 50 ans ; tumeurs vérifiées par examen ana-path.*

Les formes incomplètes sont fréquentes et certains critères doivent faire entreprendre une enquête à la recherche d'un syndrome de Lynch. En particulier la conférence d'expertise collective française a retenu des critères cliniques élargis nécessitant la prescription d'une consultation d'oncogénétique d'emblée.

Indication d'une consultation d'oncogénétique (accord d'experts) :

- *Personnes ayant deux parents atteints par un cancer du spectre dont un avant l'âge de 50 ans*
- *Malades ayant un antécédent personnel de cancer du spectre HNPCC*
- *Malades de moins de 40 ans*
- *Présence d'une instabilité microsatellitaire chez un patient de moins de 60 ans ou quelque soit l'âge en cas d'antécédent au 1er degré d'un cancer du spectre HNPCC.*

La recherche du phénotype d'instabilité microsatellitaire (MSI) présent dans les cellules tumorales à l'aide des techniques de biologie moléculaire permet une reconnaissance de

formes potentiellement prédisposées. Au moins 3 des 5 marqueurs microsatellites testés devront être positifs pour retenir le phénotype d'instabilité microsatellitaire. En complément, l'immunohistochimie à la recherche d'une extinction de l'une ou l'autre des protéines de la réparation (MMR) MSH2, MLH1, MSH6 permettra alors d'orienter la recherche de l'altération génétique causale vers l'un ou l'autre gène. En cas d'extinction de MLH1, la recherche de la mutation BRAF V600E est recommandée. Si celle-ci est présente, il n'est pas nécessaire de réaliser un séquençage des gènes MMR. En l'absence de la mutation BRAF V600E ou d'emblée en cas d'extinction de MSH2 ou MSH6 un séquençage des gènes MMR peut être proposé. Cette altération sera recherchée par une prise de sang.

**Indications de recherche d'un phénotype d'instabilité microsatellitaire (MSI)
(accord d'experts) :**

- *Patient de moins de 60 ans atteint par un cancer du spectre HNPCC*
- *Patient quel que soit son âge, ayant un antécédent familial au premier degré de cancer du spectre HNPCC.*

CONTENU ET DEROULEMENT

Les consultations d'oncogénétique ont certaines particularités :

Le recueil initial de l'information passe par la reconstitution de l'arbre généalogique (histoire familiale). Cette étape est à la fois délicate car elle fait intervenir des tiers à la consultation, et douloureuse pour les consultants car pratiquement tous, même lorsque le risque génétique ne sera pas retenu, ont une histoire familiale où coexistent plusieurs cas de cancers ;

- La prise en charge des personnes à risque élevé va s'effectuer sur une longue période (des dizaines d'années) ;
- Parmi les possibilités d'interventions, certaines comme la chirurgie prophylactique sont irréversibles ;
- La consultation a un retentissement psychologique au niveau personnel, familial et social ;

- Il y a un décalage important entre les avancées scientifiques, leurs applications thérapeutiques et les attentes du public ;
- Enfin, il existe, pour les personnes ayant une mutation délétère, un risque élevé de survenue de pathologies graves, redoutées car encore mutilantes et fréquemment mortelles.

Pour ces raisons, il est très important de rappeler l'importance du respect de certaines règles relatives aux interventions médicales :

- Il faut que l'action envisagée pour la personne lui procure un bénéfice.
- Pour respecter cet objectif majeur, les médecins doivent affronter une double difficulté, celle de la nature du bénéfice attendu (espérance de vie, qualité de vie, dignité, conséquences psychologiques et sociales), et celle de leurs mesures. Les interventions proposées (simples consultations, surveillance par imagerie, coloscopies, endoscopies, chirurgie prophylactique ...) présentent toutes des inconvénients potentiels. Le bénéfice attendu est donc un compromis entre les risques et les avantages. La possibilité de sauver des vies ne peut cependant servir seule à justifier tous les types d'interventions, ni être mise en balance avec les échecs, les insuffisances et leur acceptabilité.
- Il convient d'examiner la réalité de la demande des personnes consultant et leur laisser le libre choix des décisions (principe désigné par la suite sous le terme d'autonomie).
- Il convient d'éviter toutes pressions à finalités médicales ou scientifiques et de chercher à limiter les pressions familiales afin que les personnes accèdent librement et volontairement aux consultations.
- Grâce aux informations et aux conseils qui leur sont délivrés, les personnes doivent avoir la maîtrise des choix et des décisions, y compris celle de refuser de réaliser l'analyse proposée ou d'en découvrir le résultat (droit de ne pas savoir).

Concernant les aspects pratiques de cette consultation, en particulier ceux portant sur la confidentialité, il est recommandé de ne pas diffuser les arbres représentant la répartition des pathologies familiales. Une anonymisation est possible. En revanche il

peut être utile de laisser un document écrit à l'issue de cette consultation (document-type et/ou annoté).

Il est interdit de contacter et, à plus forte raison, de convoquer une personne, qui n'aurait pas demandé explicitement à rencontrer un médecin généticien ou à réaliser une analyse génétique. Le médecin, dans le respect des règles relatives au secret médical, ne peut écrire à une personne s'il ne l'a pas déjà rencontrée, puisque le contenu du courrier inclurait nécessairement une information relative à la pathologie, dont est atteint le consultant.

LE SUIVI

Le suivi des personnes à haut risque (établi sur des arguments génétiques, familiaux ou moléculaires) se fera en collaboration avec l'équipe pluridisciplinaire d'oncogénétique dans le cadre d'un réseau compétent. La prise en charge médicale de ces personnes doit bénéficier :

- De l'assurance d'un soutien psychologique pour les personnes particulièrement anxieuses, du fait du risque accru, de l'impossibilité de se voir fournir une réponse de type binaire (analyse moléculaire infructueuse), même après un test négatif. Cela suppose une grande disponibilité de tous les intervenants.
- D'une évaluation, qui nécessite une communication des informations entre les différents acteurs de la prise en charge avec l'accord des personnes.
- D'un contrôle de qualité des procédures.

Ces objectifs seront d'autant mieux atteints que seront mis en œuvre des protocoles évaluatifs.

ANALYSE IMMUNO HISTOCHIMIQUE

L'immunohistochimie permet d'étudier sur une coupe histologique l'expression tissulaire des protéines de réparation des mésappariements de l'ADN. A l'état normal, ces protéines sont exprimées dans le noyau de nombreuses cellules de l'organisme, en particulier dans l'intestin par les cellules du tiers inférieur des cryptes de la muqueuse, par les lymphocytes du centre germinatif des follicules lymphoïdes et par les lymphocytes et les cellules endothéliales du stroma de la tumeur, qui servent ainsi de témoins internes positifs à la technique.

Le principe du test consiste à rechercher une perte de l'expression normale d'une ou plusieurs de ces protéines dans les cellules tumorales par rapport aux tissus normaux présents sur la coupe histologique. En effet, les individus porteurs d'une mutation délétère d'un gène MMR héritent un allèle muté d'un des gènes de réparation de l'ADN du parent atteint et un allèle du parent non atteint. L'activité de réparation de l'ADN est maintenue dans les tissus grâce à l'expression de la protéine produite par l'allèle normal. Pendant la croissance tumorale, l'allèle normal est inactivé par mutation, délétion ou méthylation, ce qui aboutit à une perte de l'expression de la protéine dans les cellules tumorales.

Trois protéines peuvent être actuellement étudiées à l'aide d'anticorps monoclonaux disponibles dans le commerce : *MLH1* (clone G-168-15, PharMingen, Le Pont de Claix, France), *MSH2* (clone FE11, Oncogene Research, Meudon, France), et *MSH6* (clone 44, Becton Dickinson, Le Pont de Claix, France). Des anticorps similaires commencent à être commercialisés par d'autres firmes, mais la quasi-totalité des études publiées ont utilisé les anticorps précités. Enfin, quelques travaux ont suggéré l'intérêt d'étudier l'expression de la protéine PMS2 (clone A16-4, Pharmingen) qui permettrait de détecter certaines mutations de *MLH1* qui entraînent une perte d'expression de PMS2 sans anomalie d'expression de *MLH1* (Young et al., 2001 ; de Jong et al., 2004). L'immunohistochimie a l'avantage de pouvoir être pratiquée sur des prélèvements fixés et inclus en paraffine, quel que soit le fixateur utilisé. Mais il est en pratique difficile

d'obtenir des résultats reproductibles dans toutes les conditions de fixation, les meilleurs résultats étant obtenus avec le formol à 10%.

En utilisant deux anticorps dirigés contre les protéines *MLH1* et *MSH2*, dont les gènes correspondants sont les plus fréquemment atteints dans le syndrome de Lynch, on peut mettre en évidence une perte d'expression d'une de ces deux protéines dans les noyaux des cellules tumorales dans les cancers colorectaux. La perte d'expression est exclusive, ce qui signifie qu'elle ne concerne qu'une seule des deux protéines *MLH1* ou *MSH2*, et est somatique, ce qui signifie qu'elle ne s'observe que dans les cellules tumorales. Pour pouvoir affirmer qu'il existe une perte d'expression d'une des deux protéines *MLH1* ou *MSH2*, on doit observer une absence de marquage homogène de la totalité des cellules tumorales présentes sur la coupe histologique, en présence d'un témoin interne positif c'est-à-dire d'un marquage nucléaire des cellules épithéliales de la muqueuse normale adjacente, des lymphocytes du centre germinatif des follicules lymphoïdes lorsqu'ils sont présents, ou encore des lymphocytes ou des cellules endothéliales du stroma de la tumeur situés entre les éléments tumoraux. Ceci implique de disposer pour l'immunohistochimie d'un prélèvement fait à cheval sur la tumeur et la muqueuse saine. Deux études ont montré que l'immunohistochimie pouvait être réalisée sur des prélèvements regroupés grâce à la technique de « tissue microarray », ce qui permet d'effectuer rapidement l'analyse de grandes séries de patients (Hendriks et al., 2003 ; Jourdan et al., 2004).

L'immunohistochimie permet d'identifier la protéine anormale (Cf. annexe 4), ce qui oriente la recherche directe de mutation délétère vers un des gènes de réparation des mésappariements de l'ADN. L'immunohistochimie est aussi une technique complémentaire de la recherche d'instabilité de microsatellites par génotypage. Les premières études ont montré qu'il existe une excellente corrélation entre la perte d'expression d'une protéine observée en immunohistochimie et l'instabilité de microsatellites mise en évidence par génotypage. En effet, les cancers colorectaux ayant une perte d'expression d'une protéine de réparation des mésappariements de l'ADN ont presque toujours un phénotype MSI, et les cancers colorectaux sans perte d'expression d'une protéine de réparation des mésappariements de l'ADN sont le plus souvent dépourvus d'instabilité de microsatellites. La conférence de Bethesda a confirmé que l'immunohistochimie était une méthode appropriée d'étude des tumeurs survenant chez

des malades à risque de syndrome de Lynch, au même titre que le génotypage (Umar et al., 2004).

En comparant les données issues de l'étude de Lindor et al. (2002), on obtient pour la recherche d'instabilité de microsatellites par rapport à l'immunohistochimie une sensibilité de 1, une spécificité de 0.97, une valeur prédictive positive de 0.92 et une valeur prédictive négative de 1; pour l'immunohistochimie, par rapport à la recherche d'instabilité de microsatellites, on obtient une sensibilité de 0.92, une spécificité de 1, une valeur prédictive positive de 1 et une valeur prédictive négative de 0.96. En résumé, si l'on considère le classement des tumeurs en MSI ou MSS par génotypage comme le « gold standard », l'absence d'expression d'une protéine MMR signe le phénotype MSI, alors que les 2 protéines *MSH2* et *MLH1* sont présentes dans 3% des tumeurs MSI.

Lorsque l'on ajoute l'étude par immunohistochimie de la protéine *MSH6* à celle des protéines *MLH1* et *MSH2*, la sensibilité de l'immunohistochimie par rapport au génotypage augmentait de 93 à 100% et la valeur prédictive négative de 99 à 100% dans une étude portant sur 204 cancers colorectaux (Rigau et al., 2003). Une autre étude (Chapusot et al., 2004) a comparé la caractérisation de la fonction MMR par immunohistochimie et génotypage sur une série de 462 adénocarcinomes colorectaux. Les prélèvements étaient fixés au formol et inclus en paraffine. Pour 87 cas, il a été nécessaire de refaire l'analyse immunohistochimique en modifiant le protocole. En dépit de ces deux essais, qui ont augmenté la durée de l'analyse et son coût, 8 cas sont restés ininterprétables. Au total, une discordance entre les 2 analyses a été observée dans 28 cas, dont 10 étaient probablement le fait d'une hétérogénéité intratumorale (marquage focal). Les raisons de ces discordances pourraient être le délai de prise en charge et la durée de fixation des pièces de résection par le pathologiste. L'analyse immunohistochimique des protéines MMR nécessite donc une standardisation du traitement des pièces opératoires, une interprétation par des pathologistes entraînés.

Les tumeurs se développant dans le cadre d'un syndrome de Lynch ont un phénotype instable qui se caractérise par l'apparition de nouveaux allèles lorsque l'on génotype des marqueurs microsatellites à partir de l'ADN extrait de cellules tumorales. Ce phénotype instable, à l'exception d'un cas insuffisamment documenté (Berends et al., 2002), est constant dans les cancers colorectaux associés à l'inactivation d'un des gènes MMR principaux.

La caractérisation de ce phénotype par les techniques de génétique moléculaire est actuellement standardisée par une conférence internationale organisée par le NIH et publiée dans Cancer Research en 1998 (Boland et al., 1998). Elle recommande le génotypage de 5 marqueurs D2S123, D5S346, D17S250; BAT-25, BAT-26. Les trois premiers sont des microsatellites dinucléotidiques, et les deux derniers des microsatellites mononucléotidiques quasi monomorphes. La réalisation de ce test, qui compare le génome constitutionnel et tumoral d'un patient, suppose un contrôle histologique de la qualité du prélèvement tumoral. Ceci implique donc que les laboratoires de génétique moléculaire réalisant cet examen soient en étroite relation avec les services d'anatomie pathologique. **Le phénotype instable appelé MSI-H (MicroSatellite Instability High) est déterminé par la présence d'au moins deux microsatellites instables sur les 5 testés.** La présence d'un seul microsatellite instable devant conduire à génotyper une nouvelle série de microsatellites afin de définir d'une manière non ambiguë le phénotype tumoral. Un taux d'instabilité non nul et inférieur à 30% fait porter le diagnostic de phénotype MSI-L (pour MicroSatellite Instability Low). L'ensemble des données de la littérature ne montre pas de différence clinico-biologique entre les tumeurs ayant un phénotype MSS (MicroSatellite Stable) et MSI-L (Laiho et al., 2002). Ce phénotype MSI-L n'est d'ailleurs pas observé dans les tumeurs des patients porteurs d'une mutation délétère constitutionnelle d'un gène MMR.

En conclusion, les tumeurs sont définies par le degré de stabilité de leur ADN microsatellite, les tumeurs MSI-H, où au moins 2 locus sur 5 sont instables, et les tumeurs non instables, où 1 locus est instable (MSI-L) ou aucun (MSS). Dans la suite, nous parlerons donc de tumeurs MSI ou MSS.

D'autres propositions de marqueurs ont été faites en particulier à l'aide de microsatellites mononucléotidiques afin de pallier l'absence de disponibilité d'ADN provenant de cellules normales. L'équipe de Richard Hamelin propose de génotyper 5 marqueurs quasi monomorphes BAT-26, BAT-25, NR-21, NR-22, NR-24. Le statut MSI est alors affirmé par la présence d'une instabilité de 3 des 5 marqueurs (Suraweera et al., 2002). Cette technique permet d'identifier les tumeurs MSI en l'absence d'ADN provenant de cellules normales, à condition que le contingent cellulaire tumoral du prélèvement soit d'au moins 10%. Les résultats de la conférence internationale retiennent ce nouveau panel comme une alternative au précédent (Umar et al., 2004).

Le but de ce test biologique sur cellules tumorales est de permettre un criblage des patients atteints de cancers colorectaux afin d'isoler le sous-groupe de patients pouvant bénéficier d'une recherche de mutation délétère constitutionnelle des gènes MMR. La sensibilité et la spécificité ainsi que la valeur prédictive positive et négative du test MSI, pour la recherche d'une mutation délétère des gènes *MSH2* et *MLH1*, doivent donc être appréciées ; ceci est possible à partir de quelques articles publiés (Aaltonen et al., 1998 ; Salovaara et al., 2000; Dunlop et al., 1997; Katballe et al., 2002).

La sensibilité est probablement très proche de 1. A notre connaissance, la présence d'un faux négatif est rapportée dans un cas de mutation du gène *MLH1* (substitution au niveau de l'exon 16), dont le caractère délétère ne peut être cependant formellement affirmé, conduisant à une sensibilité minimale de 0.97 dans cette série (Liu et al., 2000).

Ainsi il est possible de réserver la réalisation d'une analyse génétique constitutionnelle des gènes MMR aux personnes porteuses d'un cancer MSI, et par conséquent, de s'abstenir d'une telle analyse chez les personnes porteuses de cancers MSS.

Des études ont cherché à identifier la valeur prédictive positive (VPP) du phénotype MSI pour l'identification d'une mutation délétère des gènes MMR dans différentes populations, en particulier chez des patients répondant aux critères d'Amsterdam ou de Bethesda (Christensen et al., 2002 ; Terdiman et al., 2002). Cette VPP peut être estimée à 0.55 IC_{95%} [0.43-0.66] ; la valeur de 0.66 est atteinte pour les familles répondant aux critères d'Amsterdam II, pour lesquelles seules des mutations ponctuelles des gènes

MSH2 et *MLH1* ont été cherchées ; il est vraisemblable qu'une analyse du gène *MSH6* et une recherche d'anomalies génomiques complexe de ces gènes augmenterait cette VPP au-delà de 0.90. Cette valeur est assez comparable à celles trouvées en population (Aaltonen et al., 1998 ; Salovaara et al., 2000) pour les sous-groupes de patients atteints de :

- un cancer du côlon avant 50 ans (0.45-0.78) ;
- un cancer du côlon gauche (0.21-0.52) ;
- un antécédent familial de cancer du côlon ou du corps de l'endomètre (0.56-0.625) ;
- un antécédent personnel pour une de ces deux localisations tumorales (0.36-0.78).

Par contre cette valeur prédictive positive est beaucoup plus faible pour les patients atteints du cancer du côlon droit (0.14-0.16). Ceci est dû à la fréquence des cancers du côlon droit sporadiques ayant un phénotype MSI lié à une méthylation acquise du promoteur du gène *MLH1*.

La valeur prédictive négative du test MSI (i.e. probabilité d'être non porteur d'une mutation délétère MMR si la tumeur est MSS) été évaluée dans 3 analyses (Katballe et al., 2002; Christensen et al., 2002; Wahlberg et al., 2002). Elle est de 0.71, IC_{95%} [0.59-0.81]. L'analyse des gènes MMR avait porté sur la recherche de mutations ponctuelles sur les 2 gènes *MSH2* et *MLH1*. On sait maintenant qu'il convient de considérer en outre les anomalies génomiques complexes et les mutations ponctuelles du gène *MSH6* (respectivement 20% et 10% de la totalité des anomalies génétiques des 3 gènes). La VPN est donc très peu différente de 1.

La stratégie de détection des porteurs de mutations délétères des gènes *MSH2* et *MLH1* parmi les sujets atteints par un cancer colorectal repose à la fois sur des critères cliniques et éventuellement sur une approche de pré-criblage par la détermination du statut MSI des cancers. Une étude coût/efficacité montre que l'attitude qui consiste à rechercher de manière systématique une mutation délétère des gènes *MLH1* et *MSH2* chez les patients répondant aux critères d'Amsterdam II, et après un criblage du

phénotype MSI chez ceux ayant des critères incomplets, est celle qui a le meilleur rapport.

ANALYSE MOLECULAIRE CONSTITUTIONNELLE

Différentes méthodes ont permis l'identification de mutations délétères dans des gènes prédisposant au syndrome HNPCC, basées sur la génétique inverse (analyses de liaison génétique suivie d'un clonage positionnel) et sur la connaissance de gènes candidats.

Une fois un gène cloné et sa séquence codante établie (cartographie des régions génomiques correspondant aux exons, aux introns et établissement de la séquence transcrite du gène), son analyse moléculaire à des fins médicales est alors techniquement possible. La séquence d'un gène peut être modifiée de différentes manières, avec des conséquences variables. Ces modifications peuvent impliquer la séquence codante (mutations) ou non (phénomènes épigénétiques tels que la méthylation d'un promoteur). Les mutations peuvent être délétères, lorsque une des fonctions de la protéine est abolie, neutres (ce sont les variants sans conséquence fonctionnelle), ou sans conséquence connue (il s'agit des variants de signification inconnue). Les mutations délétères des gènes MMR sont des mutations ponctuelles (délétion, insertion, duplication, substitution nucléotidique) ou complexes (délétion, duplication d'un ou plusieurs exons). Les mutations ponctuelles introduisant un codon stop prématuré sur la protéine et les mutations complexes sont considérées comme délétères. Les substitutions nucléotidiques de type faux-sens sont d'interprétation plus équivoque, et beaucoup sont actuellement classées dans les variants de signification inconnue.

Il existe plusieurs stratégies d'analyse pour mettre en évidence les mutations. Nous ne citerons que les plus couramment employées en médecine. Elles ont toutes en commun qu'elles se basent sur une technique de type PCR, appliquée séquentiellement à chaque exon, ou fragment d'exon. Ces produits de PCR, appelés amplicons sont ensuite analysés avec ou sans méthode de pré-criblage. Le séquençage direct apparaît comme étant le « gold standard » dans le sens où l'on reconnaît à cette technique la sensibilité la plus élevée. Elle peut être entièrement automatisée, et, malgré la diminution des volumes

réactionnels que cette automatisation a permis, son coût reste encore élevé. Des progrès techniques sont attendus à court terme avec les nouvelles générations d'appareils (rapidité d'analyse, facilité d'interprétation des résultats).

Les autres méthodes, dites de pré-criblage, permettent de repérer les régions d'un gène où se trouvent des variations de séquence, qui seront ensuite caractérisées par séquençage direct. Les temps d'analyse sont réduits. Plusieurs méthodes existent :

- le PTT (protein truncation test), consiste en une transcription et une traduction in vitro, qui met en évidence des protéines tronquées par électrophorèse ;
- le SSCP (single strand conformation polymorphism) et le DGGE (denaturing gradient gel electrophoresis). Ces deux techniques basées sur la migration de fragments d'ADN permettent de visualiser des variations électrophorétiques résultant de l'existence d'espèces moléculaires différentes bien que de même taille ;
- la DHPLC (denaturing high pressure liquid chromatography) est un dérivé automatique de la méthode précédente, diminuant les variations inter-manipulateurs et donc augmentant sa sensibilité.

Une étude coût-efficacité récente réalisée dans le contexte du cancer du sein (BRCA1), dont les résultats sont transposables aux gènes MMR, a montré que le séquençage direct des gènes de grande taille était la technique la plus coûteuse, et que d'autres stratégies se révélaient plus coût-efficaces (Sevilla et al., 2003). Cette étude permet actuellement de faire un choix en fonction des contraintes économiques et des volumes à traiter ainsi que de la sensibilité considérée comme acceptable. Il s'agit donc de trouver un équilibre entre le nombre de tests à réaliser, le coût et la sensibilité (nombre de mutations manquées) de la technique adoptée, qui nécessitera une spécialisation du personnel technique très variable.

La plus part des variations génétiques sont représentées par des polymorphismes (single nucleotide polymorphism-SNP), dont au moins 10 millions ayant une fréquence supérieure à 1% ont été identifiées dans le génome humain [Kruglyak and Nickerson, 2001]. Une majorité de ces variations sont situées dans des régions inter géniques et n'ont pas de traduction phénotypique. Cependant, certaines sont localisées dans des régions codantes, ou vont perturber l'épissage de gènes, entraînant ainsi une conséquence phénotypique. Ces mutations représentent 55% des gènes répertoriés dans HGMD database (82176 sur 148413)(<http://www.hgmd.cf.ac.uk/ac/>).

Certaines de ces variations sont associées sans ambiguïté à un syndrome de Lynch car elles entraînent la synthèse d'une protéine tronquée, et de fait, un système MMR non fonctionnel.

Ces variations sont classiquement :

- des mutations faux sens (entraînent un codon stop) (« non sens »)
- des delins (délétions/insertions) avec décalage du cadre de lecture (« frameshift »)
- des mutations au niveau des jonctions intron/exons
- des mutations introniques dans les régions consensus (+/-1 ou 2)
- des mutations silencieuses et introniques situées dans le voisinage de site d'épissage entraînant la formation de transcrits anormaux majeurs
- des grandes délétions (exons) (« large genomic rearrangement »)
- des grandes duplications entraînant un décalage du cadre de lecture avant une jonction intron/exon

Ces variations pathogènes sont largement représentées dans les bases de données. Cependant, un grand nombre de variations coexistent, essentiellement dites faux sens (« missense »), ou un nucléotide est substitué, dupliqué ou délété. Les substitutions nucléotidiques entraînant un changement d'un acide aminé, ou les délétions de 3 nucléotides, ne changent pas la phase de lecture, et posent le problème de leur caractère

délétère, puisque la quasi-totalité de la protéine est intacte. Ce type de mutation, dont on ne connaît que partiellement les conséquences fonctionnelles, représente environ 30% des variations du gène *MLH1*, 15% du gène *MSH2* et jusqu'à 37% du gène *MSH6* identifiées chez des patients répondant aux critères d'analyse génétique constitutionnelle à la recherche d'une mutation délétère des gènes MMR (cf. tableau 4).

Gènes	Faux sens	silencieuse	In frame	Faux sens/ frameshift	Intronique	Grand réarrangement génomique	Sites d'épissage	Epissage altéré	Total (%)
MLH1	445	48	21	129	439	82	96	21	41.7
MSH2	272	42	22	84	431	88	64	6	32.8
MSH6	203	42	12	96	180	8	10	1	18
PMS2	122	10	2	23	47	20	5	1	7.5

Tableau 4 : données de la base InSiGT concernant la nature des variants MMR (2012).

Lorsque l'on peut démontrer grâce à des arguments de poids comme les tests fonctionnels, que l'un de ces variants perturbe la fonction du gène, celui-ci est catalogué délétère. Dans le cas contraire, si aucun effet n'est démontré, il est qualifié de neutre. Enfin, dans un nombre de cas non négligeable, il manque des éléments de caractérisation pour apprécier le lien entre le variant et une prédisposition héréditaire. Ces variants sont appelés « variants de signification inconnue » (VSI).

Plusieurs critères ont été proposés pour parvenir à caractériser la pathogénicité des variants:

- caractère de novo du variant
- ségrégation du variant parallèlement à la pathologie dans une famille
- absence démontrée dans une population contrôle

- changement de polarité ou de taille d'un acide aminé
- modification d'un acide aminé dans un domaine protéique conservé
- effet significatif du variant sur l'épissage ou la fonction, démontré par un test chez l'animal ou dans des essais in vitro
- perte de l'allèle muté dans la tumeur (LOH)
- perte d'expression de la protéine dans la tumeur.

Malheureusement, plusieurs de ces critères sont difficiles à évaluer. Par exemple, le nombre de sujets contrôle d'une population est imprécis. La présence d'un variant chez des sujets sains est probablement suffisante pour affirmer sa neutralité et son absence ne permet pas d'évoquer sa pathogénicité. Enfin, la perte d'expression d'une protéine en immunohistochimie ou une instabilité des microsatellites ne peuvent garantir qu'ils ne sont pas la conséquence d'un autre variant pathogène non identifié.

De façon globale, les critères requis sont les suivants :

- cliniques et généalogiques ;
- tumoraux (type MSI, extinction protéique sélective) ;
- moléculaires (caractère unique de la variation dans l'ADN analysé, co ségrégation intrafamiliale) ;
- phylogénétiques et biochimiques (différents degrés de conservation, prédiction des structures tertiaires) ;
- bibliographiques (descriptions antérieures, études en population normale) ;
- fonctionnels (tests in vitro, in vivo).

BASES DE DONNEES :

HISTORIQUE

Dans les années 80, les variations des gènes étaient reportées dans les bases de données existantes telles que DataBase (GDB) (Pearson et al., 1991), GenBank (Bilofsky, H. S 1986,), EMBL (Hamm G. H. & Cameron, 1986), et Swiss-Prot (Bairoch, A. & Boeckmann,

B 1991). Cependant, ces dernières étaient peu compatibles avec une extraction des informations concernant les mutations. La publication en 2003 du génome humain, a offert au monde médical et scientifique une profusion de données utilisables pour la compréhension du mécanisme des maladies, leur diagnostic et les perspectives thérapeutiques. Depuis, des milliers de variations sont identifiées chaque année dans les laboratoires.

Deux approches principales ont été développées à la fin des années 90, notamment sous l'impulsion de la Human Genome Variation Society (HGVS) (Horaitis et al., 2004.):

- « core » database (ou « central » database)
- « locus specific » database (LSDB)

La première récence l'ensemble des mutations de tous les gènes, l'archétype étant la Human Gene Mutation Database ((HGMD; www.hgmd.org), dans laquelle chaque mutation est entrée une fois, et où la description phénotypique est limitée, rendant difficiles les études de variabilité phénotypiques.

A l'inverse les LSDB collectent l'ensemble des mutations, publiées ou non, pour chaque gène avec une description moléculaire, clinique et biologique. Ces données sont validées par des experts appelés « curateurs », dont le rôle est de limiter au maximum les erreurs (classiquement 10% des informations publiées sont erronées) et de standardiser les descriptions cliniques et biologiques de chaque patient. Sur le même modèle, les bases de données se sont également développées pour les séquences non codantes (introniques, 5' et 3') ainsi que pour les polymorphismes (SNP).

Plusieurs bases de données MMR existent, à commencer par la base historique InSiGHT et MMR UV (www.MMRUV.info). Ces deux bases recensent exhaustivement les variants des gènes MMR, cependant sans décrire précisément les modalités de caractérisation et de validation de la pathogénicité des variants recensés.

D'autres bases publiées font également référence. Nous détaillerons les caractéristiques des deux plus représentatives.

Ou J et al (Ou J et al., 2008), ont extrait dans Pubmed l'ensemble des données concernant les variants des gènes MMR. Les critères recherchés étaient les suivants :

- phénotype, histoire familiale,
- statut MSI (génotype et IHC),
- fréquence des variants chez les contrôles, analyse de ségrégation
- conservation dans les espèces,
- type de tests fonctionnels et in silico.

Dans cette base, également en lien avec InSiGHT, il n'apparaît pas qu'il existe de panel d'expert. 513 variants (308 *MLH1*, 180 *MSH2*, 15 *MSH6*, 3 *PMS2* et 7 *MLH3*) sont décrits, dont 50% pathogènes.

The screenshot shows the website of the University Medical Center Groningen, Department of Genetics - Oncogenetics Research Group. The page displays details for a specific mutation: MSH2: A834T. The mutation is described as a missense mutation (c.2500G>A, p.Ala834Thr) located at position 15 of the MSH2 gene. The classification is 'Missense mutation' with a 'Further classification' of 'na' and a 'Priority change' of 'na'. The 'Domain' is 'MSH2_0062' and the 'LOVD' is 'MSH2_0062'. Below this, a table titled 'Functional and in silico analysis' provides details on the mutation's impact. The table has columns for 'Article', 'Type', 'Methods', 'Result', 'Outcome', and 'Comments'. Two articles are listed: 'Ohta H, Oosterbaard M, 2008, 1406-17' and 'Ohta H, Oosterbaard M, 2008, 1406-17'. The first article describes in relevant MMR deficient cell line + cell free in vitro MMR assays, showing repair efficiency comparable to the WT (48.8% versus 44.7%) and normal expression in vivo cells compared to WT. The second article describes in silico analysis using SIFT score, resulting in a score of 0.10. The outcome for both is 'nonpathogenic'. The comments for the second article state: 'Nonpathogenic: if SIFT score < 0.05 then the AA substitution is predicted to affect protein function'.

Mutation	DNA change	Protein change	Exon	Codon	Classification	Further classification	Priority change	Conservation	Domain	LOVD
A834T	c.2500G>A	p.Ala834Thr	15	834	Missense mutation	na	na	na	MSH2_0062	MSH2_0062

Article	Type	Methods	Result	Outcome	Comments
Ohta H, Oosterbaard M, 2008, 1406-17	In relevant MMR deficient cell line + cell free in vitro MMR assays		repair efficiency comparable to the WT (48.8% versus 44.7%) normal expression in vivo cells compared to WT	nonpathogenic	
Ohta H, Oosterbaard M, 2008, 1406-17	In silico analysis	SIFT score	0.10	Nonpathogenic	if SIFT score < 0.05 then the AA substitution is predicted to affect protein function

Barnetson et al (Barnetson et al., 2008) ont publié les données d'un registre écossais sur une période allant de 1999 à 2004, de 932 CCR de moins de 55 ans et l'on comparé à une population contrôle. Les critères étudiés étaient :

- Statut MSI (géotypage, IHC)
- Données cliniques, ségrégation
- Analyses in silico: Clustal W, Genscan, NNSplice
- Analyse d'épissage (2mutations *MSH2*).

En complément de la description des variants, les auteurs décrivaient un score de pathogénicité basé sur une pondération par critère:

- Variant présent chez les contrôles (-6), ségrégation / 2 membres de la famille (2) / 2-5 membres (4) / > 5 membres (6)
- MMS (0), MSI-L (2), MSI-H (4), IHC + (4)
- Conservation (1), SIFT non tolérée (2); PolyPhen 3 (2), 2 (1); In vitro MMR (1)

Cependant, la méthodologie finale du calcul du score n'était pas précisée, ni même le seuil permettant de définir la pathogénicité des variants. Au final, 110 mutations sur 932 CCR (12%) étaient caractérisées, dont :

- 33% de mutations tronquantes
- 67% de variants inconnus

TABLE 1. Mutations Identified in Colorectal Cancer Cases and Evidence for Pathogenicity*

Mutation identified	No. of cases (n = 932)	Cases frequency (95% CI)	No. of controls	Controls frequency (95% CI)	Effect size	95% CI	No. Affected cases	Tumor assessed	Tumor MSI status ^a				Conserved (ClustalM)	SIFT tolerated	PolyPhen annotated ^c	In vitro MMR (References) ^d	Category of pathogenicity	Score ^e	
									MLH1	MSH2	MSH6	MSH3							
MLH1																			
c.422C>G p.Q224	1	0.001 (0.0002-0.004)	2/302	0.002 (0.0006-0.007)	0.18	0.04-0.32 ^f	0	1	MSI	+	+	+	+	Yes	No	2		Design	0
c.2050G>A p.D678	1	0.001 (0.0002-0.004)	0/308	0 (0-0.002)	0.43	0.26-0.61 ^g	1	1	MSI-H	-	N/A	N/A	N/A	Yes	No	2		Pathogenic	19
c.277A>G p.R930	1	0.001 (0.0002-0.004)	0/187	0 (0-0.003)	9.73	0.18-46.12 ^h	0	1	MSI	+	+	+	+	No	Yes	2	+ (12.2)	Unclassified	2
c.293C>T p.P309S	1	0.001 (0.0002-0.004)	0/100	0 (0-0.004)	7.96	0.16-40.12 ^h	0	1	MSI	+	+	+	+	Yes	No	2		Unclassified	5
c.1217G>A p.S405Y	1	0.001 (0.0002-0.004)	1/322	0.003 (0.0002-0.009)	1.1	0.07-1.53	0	1	MSI	+	+	+	+	No	Yes	2	+ (3.4)	Design	0
c.1820T>A p.L597H	1	0.001 (0.0002-0.004)	0/303	0 (0-0.004)	7.97	0.16-40.20 ^h	0	1	MSI	+	+	+	+	No	Yes	1		Unclassified	0
c.1822AA>GC p.R576A	16	0.007 (0.002-0.028)	14/153	0.009 (0.005-0.016)	1.30	0.53-3.95	0	6	MSI	+	+	+	+	No	Yes	2	-/+ (2.35-7)	Design	0
c.2041G>A p.A681T	2	0.002 (0.0004-0.008)	0/303	0 (0-0.003)	12.6	0.24-23.22 ^h	2	2	MSI-H	-	+	+	+	Yes	No	2	-/+ (2-4.6/8/9)	Pathogenic	19
c.2366A>G p.Q688R	2	0.002 (0.0004-0.008)	5/106	0.005 (0.002-0.012)	0.46	0.09-2.56	0	2	MSI	+	+	+	+	No	Yes	1	+ (9)	Design	0
c.2366G>A p.Q718R	1	0.001 (0.0002-0.004)	6/309	0.008 (0.005-0.012)	0.16	0.02-1.48	0	1	MSI	+	+	+	+	No	Yes	1	-/+ (2.23)	Design	0
c.1232C>T p.H791R	1	0.001 (0.0002-0.004)	1/100	0.008 (0.0002-0.006)	1.07	0.07-12.18	0	0	N/A	N/A	N/A	N/A	N/A	Yes	No	3	-/+ (3.4)	Design	N/A
c.2252A>G p.K751R	2	0.002 (0.0004-0.008)	1/100	0.008 (0.0002-0.006)	2.15	0.19-23.7	0	2	MSI	+	+	+	+	Yes	Yes	1	-/+ (3.0)	Design	0
MSH2																			
c.1A>C	1	0.001 (0.0002-0.004)	0/322	0 (0-0.003)	12.58	0.23-68.7 ^h	0	1	MSI-L	+	+	+	+	No	No	1		Unclassified	4
c.139C>G p.H46Q	1	0.001 (0.0002-0.004)	0/106	0 (0-0.004)	6.2	0.16-40.42 ^h	0	1	MSI	+	+	+	+	No	No	3		Unclassified	4
c.217G>A p.R306K	1	0.001 (0.0002-0.004)	1/323	0.004 (0.0002-0.009)	1.09	0.07-12.62	0	1	MSI	+	+	+	+	No	Yes	1		Design	0
c.1965G>A p.G322D	17	0.008 (0.001-0.029)	15/496	0.03 (0.009-0.054)	0.56	0.27-1.17	1	7	MSI	+	+	+	+	Yes	Yes	2	-/+ (3.1)	Design	0
c.1281A>G p.N366S	1	0.001 (0.0002-0.004)	0/312	0 (0-0.004)	8.09	0.16-40.97 ^h	0	1	MSI-L	+	+	+	+	No	Yes	1		Unclassified	2
c.2096C>T p.P309L	1	0.001 (0.0002-0.004)	0/322	0 (0-0.003)	8.20	0.17-43.60 ^h	0	1	MSI	+	+	+	+	Yes	No	2		Unclassified	5
c.2337G>A p.M791I	1	0.001 (0.0002-0.004)	0/106	0 (0-0.003)	8.5	0.17-41.79 ^h	0	1	MSI	+	+	+	+	No	Yes	2		Unclassified	5
c.2486C>G p.T807R	1	0.001 (0.0002-0.004)	0/106	0 (0-0.003)	8.5	0.17-41.79 ^h	0	1	MSI-H	-	N/A	+	+	No	Yes	1		Unclassified	N/A
c.2503A>C p.N363H	1	0.001 (0.0002-0.004)	1/318	0.004 (0.0002-0.009)	1.08	0.07-12.93	0	1	MSI-L	+	+	+	+	No	Yes	1		Design	0
c.2634A>G>C	1	0.001 (0.0002-0.004)	0/312	0 (0-0.004)	0.09	0.00-409.5 ^h	1	1	MSI-H	+	+	+	+	N/A	N/A	N/A		Pathogenic ^g	N/A
c.2722T>G p.L941R	1	0.001 (0.0002-0.004)	0/118	0 (0-0.004)	8.1	0.16-40.97 ^h	0	0	N/A	N/A	N/A	N/A	N/A	No	No	2		Design	N/A

BUTS DU TRAVAIL

Depuis vingt ans, le réseau des laboratoires impliqués dans le diagnostic génétique du syndrome de Lynch en France a recensé 6700 variations ; 700 d'entre elles sont des variants de signification inconnues. L'objectif de notre travail était de :

- proposer un algorithme décisionnel permettant, à partir des données disponibles, de caractériser au mieux ces variants, selon une classification en 5 classes (causal, probablement causal, VSI, probablement neutre, neutre).
- Classer l'ensemble des variants soumis par le réseau afin de développer une base de données des gènes MMR de type LSDB avec l'outil UMD
- Analyser les données phénotypiques de formes rares de syndrome de Lynch grâce aux données de la base

MATERIELS ET METHODES

METHODOLOGIE DE CLASSIFICATION DES VARIANTS

En 2008, Plon et al avec le IARC (*International Agency for Research on Cancer*) ont proposé un système de classification des variants en 5 classes, adapté aux problématiques des mutations des gènes BRCA et MMR. (Plon et al., 2008). Dans ce système, **les variants des classes I (neutres) et V (pathogènes)** sont associés à une probabilité d'être pathogènes inférieure à 0.1% et supérieure à 99% respectivement (cf. tableau) et peuvent ainsi être utilisés en conseil génétique. Deux classes regroupent les variants dont la probabilité d'être pathogènes ou non (respectivement classes IV et II) est comprise entre 95-99% et 0.1-0.5%. Ces variants dits « probablement » pathogènes ou neutres, nécessitent d'autres arguments pour être caractérisés, mais en revanche peuvent permettre de proposer des indications de surveillance chez les sujets porteurs. Reste enfin la classe III, contenant les variants de signification inconnue.

Classe	Définition	Probabilité de pathogénicité	Surveillance
V	Définitivement pathogène	>0.99	Selon recommandations pour hauts risques
IV	Probablement pathogène	0.95-0.99	Selon recommandations pour hauts risques
III	Variant de Signification Inconnue (VSI)	0.05-0.949	Selon histoire familiale
II	Probablement neutre	0.001-0.049	Considéré comme non muté
I	Définitivement neutre	<0.001	Considéré comme non muté

Dans tous les cas, les auteurs recommandent que la pathogénicité de chaque variant soit validée par un panel d'experts, et que des règles formalisant l'approche basée sur l'expérience des généticiens soient élaborées, en particulier pour les gènes MMR ou l'absence de probabilité absolue est un frein à la classification.

HISTOIRE FAMILIALE

Par principe, si un variant est pathogène, il aura tendance à être associé à une histoire familiale forte, comparable à celle des variants pathogènes connus, et inversement. Cela sous-entend d'étudier à la consultation, le nombre d'individus atteints de cancers, ainsi que les âges de survenue de ces cancers.

Pour chaque variant, les données de la consultation d'oncogénétique concernant le nombre d'apparentés atteints, les âges de survenues des cancers, et les organes concernés, sont rapportés. Les critères d'Amsterdam II, lorsqu'ils sont présents, constituent le critère positif qui sera intégré dans l'algorithme décisionnel de caractérisation.

CO-SEGREGATION

L'étude de la co-ségrégation, décrite par Thompson et al (Thompson et al., 2003) et reprise entre autres pour l'étude des gènes MMR par Goldar et al (Goldar et al., 2008), a pour principes de comparer la probabilité pour un individu porteur de la mutation d'être atteint, à celle (hypothèse nulle) d'être indemne (cas où le variant est neutre). Il s'agit d'une approche similaire aux analyses de liaison, ciblée sur un variant et non sur plusieurs polymorphismes. La probabilité est dérivée du LOD score. Elle est fonction de la pénétrance et intègre le risque de développer la maladie pour chaque génotype, et la fréquence de l'allèle pathologique (en règle générale très rare). Ces paramètres de pénétrance sont classiquement issus de larges études familiales ou ségrègent des mutations pathogéniques à haut risque. Cette méthode dépend des possibilités de disposer des ADN de plusieurs individus dans une même famille. Son inconvénient est qu'elle dépend du nombre de méioses informatives. Ce paramètre possède donc un fort poids décisionnel quand il est disponible, mais est rarement suffisant à lui seul en raison des difficultés à recueillir les données cliniques et biologiques.

L'étude de l'effet de la substitution d'un acide aminé, dans une protéine est une des manières de caractériser la nature neutre ou délétère d'un variant. Plusieurs outils informatiques, appelés "in silico", sont à la disposition des biologistes afin de fournir les informations complémentaires aux données cliniques, in vitro et in vivo.

Conservation, propriétés physico-chimiques, prédictions sur la fonction protéique

L'étude de la conservation de l'acide aminé concerné par la mutation, peut être faite grâce à des logiciels comme ClustalW software (www.ebi.ac.uk/clustalw), qui analyse si celui-ci est conservé au cours de l'évolution dans plusieurs espèces (humaines, souris, drosophile, zébra Fish et levure).

SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>) (Henikoff S. 2003) est un algorithme capable de prédire l'effet de la substitution d'un acide aminé sur la fonction de la protéine et de suspecter le caractère délétère d'un variant. L'algorithme utilise la position de l'acide aminé et son type. Il choisit ensuite des séquences homologues afin de réaliser un alignement, puis calcule la probabilité qu'un autre acide aminé soit toléré à la place de l'acide normal. Un score est déterminé, et si celui-ci est inférieur à 0.05, le variant est prédit comme délétère (Cf. annexe 5).

Deux autres programmes, BLOSUM62 [Henikoff and Henikoff, 1992] et PolyPhen (<http://www.bork.embl-heidelberg.de/PolyPhen/>) [Ramensky et al., 2002], permettent d'évaluer les conséquences physico-chimiques de l'acide aminé substitué sur la structure et la fonction de la protéine.

La caractérisation d'un variant est également possible grâce à la prédiction des conséquences sur les propriétés physico-chimiques de l'acide aminé (AA) grâce à une matrice développée par Yu. Celle-ci analyse 48 propriétés physico-chimiques (groupes fonctionnels, propriétés optiques, hydrophobicité, charge, propriétés acidobasiques et taille).

(<http://cmgm.stanford.edu/biochem218/Projects%202001/Yu.pdf>).

Prédiction de l'effet d'un variant sur l'épissage:

Plusieurs outils permettent l'étude *in silico* des variants de signification inconnue. Deux algorithmes permettent d'identifier les zones supposées de séquences ESE (cf. chapitre épissage), et de prédire les conséquences d'un variant donné sur l'épissage :

- ESEFinder (Cartegni et al., 2003) fournit un score basé sur la fréquence des nucléotides à chaque position des motifs reconnus spécifiquement par 4 protéines SR (SF2/ASF, SC35, SRp46, SRp55).
- RESCUE-ESE (Fairbrother et al., 2002, et 2004) est un algorithme prédisant si un hexamère donné peut agir comme candidat ESE et évalue les conséquences telles que l'interruption ou l'altération de sites ESE présumés par une mutation au sein de l'hexamère.
- Human Splicing Finder (Desmet et al., 2009) est un outil informatique développé par l'équipe de Beroud et al., qui permet de rechercher les sites potentiels d'épissages (donneurs/ accepteurs, points de branchement, motifs auxiliaires), et calculer l'effet d'une mutation sur l'épissage.

Les performances de ces outils ont été évaluées en les comparant aux tests *in vivo* et *ex vivo* de référence. Ainsi, Auclair et al (Auclair et al., 2006) ont corrélé les prédictions de ces deux algorithmes avec l'analyse de l'ARN des patients porteurs de variants pathogènes ou neutres, et démontré une sensibilité de 80% et une spécificité de seulement 42%.

Lastella et al (Lastella et al., 2006) ont également analysé 99 variants de *MLH1* et *MSH2* avec 6 algorithmes :

- ESEfinder, RESCUE-ESE et PESX : identifient les sites présumés de motifs ESE et prédisent si ceux-ci seront affectés par les variants étudiés.
- NNSPLICE, SpliceSite finder et GENSCAN : calculent un score prédictif de l'effet du variant sur l'épissage.

En comparant les résultats de ces outils à ceux de tests *in vivo* (minigènes artificiels), les auteurs ont montrés que même si toutes les mutations tombaient sur des sites ESE

prédits par les algorithmes, la moitié seulement d'entre eux entraînait une anomalie de l'épissage. Inversement, lorsque les mutations étaient situées dans de petites régions exoniques, elles avaient un potentiel délétère sur l'épissage pas toujours corrélé aux données fournies par les outils in silico.

Prédictions de l'effet d'un variant sur la fonction MMR:

PON-MMR (<http://bioinf.uta.fi/PON-MMR>)

Cet algorithme développé par Ali H et al (Ali et al., 2012), utilise les résultats publiés pour les variants MMR pathogènes caractérisés principalement par des données des tests fonctionnels MMR (tests d'activité MMR in vitro, in vivo, et double-hybride en levure). Un pipeline d'algorithmes en ligne permettant le calcul d'un score de pathogénicité est proposé par les auteurs (Polyphen, SIFT, Mutation Taster, Mut Pred...) avec un seuil de 0.76 au-dessus duquel les variants sont pathogènes (neutre si <0.35 , VSI entre les deux). L'évaluation par les auteurs de leur outil fait état d'une précision de 0.87, sensibilité de 0.97, spécificité de 0.77 et VPN de 0.65.

Modèles prédictifs intégrant l'ensemble des caractéristiques

UMD Predictor®

Ce logiciel, développé par l'équipe de Beroud et al. (Frederic et al., 2009) offre une approche combinée des différents arguments pour une même variation. Il prend en compte

- l'acide aminé au niveau protéique, c'est-à-dire, en analysant son implication dans un domaine associé à une fonction biologique ou structurale. Ces domaines sont en général hautement conservés et peuvent être reconnus selon les critères suivants :
 - Cystéine impliquée dans un pont di sulfure (repliement protéique)
 - AA impliqué dans le transport calcique
 - Glycine impliquée dans un packing domaine-domaine
 - AA impliqué dans des modifications post traductionnelles telles que la glycosylation
 - Sites métallo protéases

- AA phosphorylés
- AA impliqués dans le transport de l'ATP
- Le degré de conservation (à partir des données de SIFT)
- Les différences physico-chimiques entre les AA sauvage et variant (BLOSUM62, Yu Matrix)
- Les conséquences sur l'épissage : UMD calcule une valeur consensus (CV) pour chaque site théorique donneur ou accepteur d'épissage, ainsi que les éléments régulateurs (ESE, ESS).

UMD Predictor® compile tous ces éléments et calcule un score permettant la caractérisation du variant (Cf. annexe 6) :

- <50 : non pathogène (polymorphisme)
- 50-64 : polymorphisme probable
- 65-74 : pathogène probable
- >74 : pathogène

Les performances de cet outil ont été évaluées par rapport aux outils disponibles. La sensibilité, spécificité, VPP et VPN étaient respectivement de 95.5%, 92.2%, 99.5% et 56%.

TESTS FONCTIONNELS

Les connaissances du fonctionnement du système de réparation des mésappariements de l'ADN chez l'homme viennent essentiellement des études réalisées dans des modèles bactériens et de levure. Les domaines fonctionnels des protéines *MLH1* et *MSH2* sont identifiés. La protéine *MLH1* comprend, en N-terminal, un domaine ATPase, en C-terminal un domaine d'interaction avec PMS1 et un motif CTH constitué des 13 derniers acides aminés, impliqué dans l'interaction avec PMS2. Quant à la protéine *MSH2*, elle possède un domaine ATPase en C-terminal et des domaines de reconnaissance et de liaison à l'ADN ainsi qu'un domaine d'homodimérisation en N-terminal. La cristallisation de la protéine MutL (partie N-terminale) et de la protéine MutS a confirmé et précisé l'existence de ces domaines.

Des souches bactériennes ou de levures portant une mutation de type faux-sens dans ces domaines fonctionnels montraient en outre un défaut de réparation de l'ADN (Pang et al., 1997, Studamire et al., 1999). Ces dernières observations ont été à la base du développement de plusieurs systèmes ayant pour objectif de tester, de façon spécifique, des variants identifiés dans des familles atteintes de syndrome HNPCC sur le plan fonctionnel. Le premier système a été conçu dans la levure par Shimodaira et al. (Shimodaira et al., 1998), basé sur la fonction appelée "effet de mutateur dominant" (*dominant mutator effect*). Ils ont pu ainsi tester 19 variants, et les classer. Sur le même principe, Quaresima et coll. (Quaresima et coll., 2003) a proposé un test similaire dans *E. Coli*. Une autre étude publiée par Ellison et coll. (2001) décrit la mise au point d'un système de mutagenèse dirigée dans la levure sur le gène *MLH1*, où 22 variants ont été analysés. En utilisant une stratégie similaire, Drotschmann et coll. (Drotschmann et al., 1999) ont testé une série de variants du gène *MSH2* et ont soulevé l'importance de la nature des tests de réparation et donc les limites de l'interprétation de ces tests.

Chez l'homme, deux approches ont été développées, tirant parti des interactions protéine-protéine, suivies d'essais de réparation *in vitro* (Guerrette et al., 1998; Nyström-Lahti et al., 2002), ou de transfections transitoires d'ADN variants dans des lignées cellulaires dont la fonction MMR est altérée, avec suivi de la fonction MMR et de fonctions dérivées (apoptose, prolifération...) (Trojan et al., 2002 ; Brieger et al., 2002).

Les objectifs d'un test fonctionnels peuvent être divers. Ils peuvent évaluer la capacité du système MMR à former des hétéro-dimères, permettre le cycle ADP-ATP, évaluer la localisation intracellulaire des protéines mutantes et enfin, évaluer la capacité d'une protéine mutée à restaurer la fonction MMR d'une lignée déficiente.

Trois tests permettent d'évaluer un système MMR fonctionnel par l'analyse de l'interaction de deux de ses protéines :

- « Yeast Two Hybrid Assay »

Le but de cet essai *in vivo* est de démontrer l'interaction des deux protéines partenaires, comme le sont les hétérodimères du système MMR.

Des souches hybrides de levures sont fabriquées, l'une associant la protéine d'intérêt (par exemple *MSH2*) au domaine de liaison de l'ADN du facteur transcriptionnel d'un

gène reporter situé en amont (exemple LacZ, His3). L'autre protéine (ex. *MSH6*) est fusionnée au domaine d'activation du même facteur transcriptionnel. Si les deux protéines sont capables d'interagir, les deux domaines du facteur transcriptionnel sont réunis, induisant l'expression du gène reporter, avec le phénotype correspondant (ex. coloration de la colonie de levure). Ce test a été utilisé pour caractériser des VSI de *MLH1* par Kondo et al. (Kondo et al., 2003)

- « GST pull down Assay »

Dans cet essai in vitro, une des deux protéines est couplée à un marqueur (Glutathionne S-transférase (GST) et fixée sur un support glutathion. L'ajout de la deuxième protéine dans le milieu permet de visualiser la protéine marquée si l'interaction a eu lieu. Guerrette et al. (Guerrette et al., 1999) ont analysé des VSI de *MLH1* et montrés une liaison diminuée avec PMS2.

- Expression transitoire de gènes MMR dans une lignée déficiente

Dans des lignées exprimant un mutant de *MSH2* ou *MLH1*, l'expression de *MSH6* ou PMS2 n'est possible qu'après transfection d'un variant non pathogène (fonctionnel) de *MSH2* ou *MLH1*. (Brieger et al., 2002)

Les tests suivants ont été utilisés pour caractériser des variants sur d'autres aspects de la fonction MMR :

- Tests basés sur le cycle ADP-ATP de MutS

La reconnaissance d'un mésappariement de l'ADN par le dimère *MSH2-MSH6* implique une réaction de liaison de l'ATP à l'ADN, suivie de son hydrolyse en ADP. Un variant d'une de ces deux protéines peut théoriquement affecter ce processus et empêcher le recrutement du dimère *MLH1-PMS2*. Des tests in vitro ont été développés dans ce sens (Cyr et al., 2012 ; Heinen et al., 2008)

- Localisation intracellulaire des protéines

Les protéines MMR sont transportées dans le noyau pour y assumer leur fonction, et ce grâce à la présence dans leur séquence protéique d'un signal de localisation nucléaire ou « nuclear localization signal » (NLS). Brieger et al ont identifiés ces séquences chez *MLH1* et PMS2, et démontré que des variants faux sens pouvaient affecter le transport

intra cellulaire des deux protéines en les retenant dans le cytoplasme. Knudsen et al ont démontré un mécanisme similaire pour des variants de EXO-1 (Knudsen et al., 2007). Cette étude de localisation, bien que techniquement difficile, peut être faite par immunohistochimie ou expression de protéines fluorescentes et permet d'apporter un élément de compréhension dans le mécanisme à l'origine d'un système MMR déficient.

ANALYSES D'ÉPISSAGE

L'épissage est un processus par lequel les ARN transcrits à partir de l'ADN génomique peuvent subir des étapes de coupure et ligature qui conduisent à l'élimination de certaines régions dans l'ARN final. Les segments conservés s'appellent des exons et ceux qui sont éliminés s'appellent des introns. La compréhension de la régulation de l'épissage est cruciale dans la mesure où des variations génomiques responsables d'épissages aberrants représentent plus de 50% de l'ensemble des mutations responsables de dysfonctions de gènes. Les mutations peuvent en effet non seulement altérer la séquence qui sera traduite en protéine, par exemple, une substitution de base modifiant un codon d'acide aminé pour un autre ou en un codon STOP, mais aussi en altérant l'épissage, générer une protéine tronquée ou absente.

Lors de la transcription de gènes codant des protéines, un ARN pré-messenger, ou transcrit primaire, est synthétisé puis épissé dans le noyau de la cellule pour donner lieu à l'ARN messenger dit mature. L'ARNm mature, constitué des seuls exons, est alors exporté vers le cytoplasme pour être traduit en protéine.

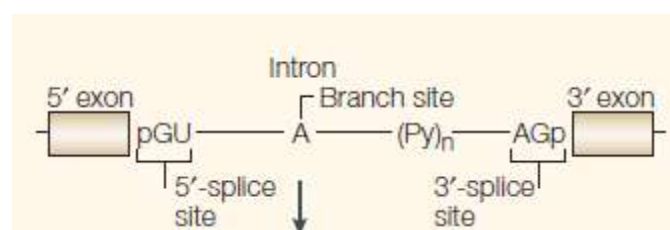
L'épissage est assuré par un ensemble de complexes ribonucléoprotéiques appelé collectivement « splicéosome ». Chaque complexe, fait de petites ribonucléoprotéines nucléaires, contient un ARN et plusieurs protéines.

Le mécanisme catalytique du splicéosome est encore inconnu, mais par analogie avec le fonctionnement du ribosome, on pense que c'est l'ARN qui est catalytique (et donc que le splicéosome est un ribozyme).

Les jonctions intron/exon contiennent des séquences nucléotidiques caractéristiques qui sont conservées. Ces séquences sont reconnues par le spliceosome. L'intron contient en plus une séquence interne, dite **boite de branchement**. Cette boite de branchement comporte une adénosine qui joue un rôle central dans le processus d'épissage.

L'épissage proprement dit s'effectue en deux temps. Premièrement, une attaque nucléophile du 2'-OH du ribose de l'adénosine de la boite de branchement sur le phosphate de la jonction exon-intron en 5'. Après cette coupure, le 3'-OH libéré au niveau de l'exon amont attaque le phosphate de la jonction intron-exon en aval. Les produits de cette réaction sont d'une part les deux exons ligaturés correctement et d'autre part, l'intron cyclisé au niveau de l'adénosine de la boite de branchement. En raison de sa forme particulière, cette forme de l'intron est appelée **lasso**. Le lasso est finalement ouvert par une enzyme de débranchement afin de pouvoir être recyclé.

Les variants affectant l'épissage vont être responsables de sauts d'exons (« exon skipping »), d'activation de sites cryptiques d'épissage, ou d'altération de la régulation de l'épissage alternatif. Ces mutations vont principalement affecter des régions clefs, hautement conservées, telles que les sites d'épissage en 5'(GU) et 3' (AG), les répétitions polypopyrimidine (Py_n), et les points de branchements, indispensables au processus d'excision des introns (figure 1 Pagani 2004)



En plus de ces éléments clefs, d'autres régions régulatrices de l'épissage sont désormais connues et susceptibles, en cas de mutation, d'affecter le processus. Il s'agit de signaux régulateurs capables de stimuler (« Exonic Splicing Enhancers »-ESE) ou d'inhiber (« Exonic Splicing Silencers »-ESS) l'épissage.

Les ESE sont des courtes séquences de 6-8 nucléotides réalisant un point de branchement de protéines dites « SR », riches en arginine et sérine et caractérisées par un domaine RRM (1-2 RNA Recognition Motif). Ces deux domaines permettent la fixation de la protéine sur les séquences ESE, puis le recrutement du spliceosome.

Les ESS sont de mécanismes moins connus, mais ces éléments semblent agir en interférant avec des éléments régulateurs négatifs tels que les hnRNP (« heterogeneous nuclear ribonucleoprotein »).

Une compétition entre ESE et ESS intervient souvent et l'épissage des exons dépend de l'affinité de chaque élément régulateur.

Les mutations affectant l'épissage ont été incriminées dans les pathologies telles que l'ataxie télangiectasie (ATM), où près de 50% des mutations du gène affectent l'épissage en créant de nouveaux sites cryptiques ou générant des sauts d'exons (Teraoka et al., 1999)

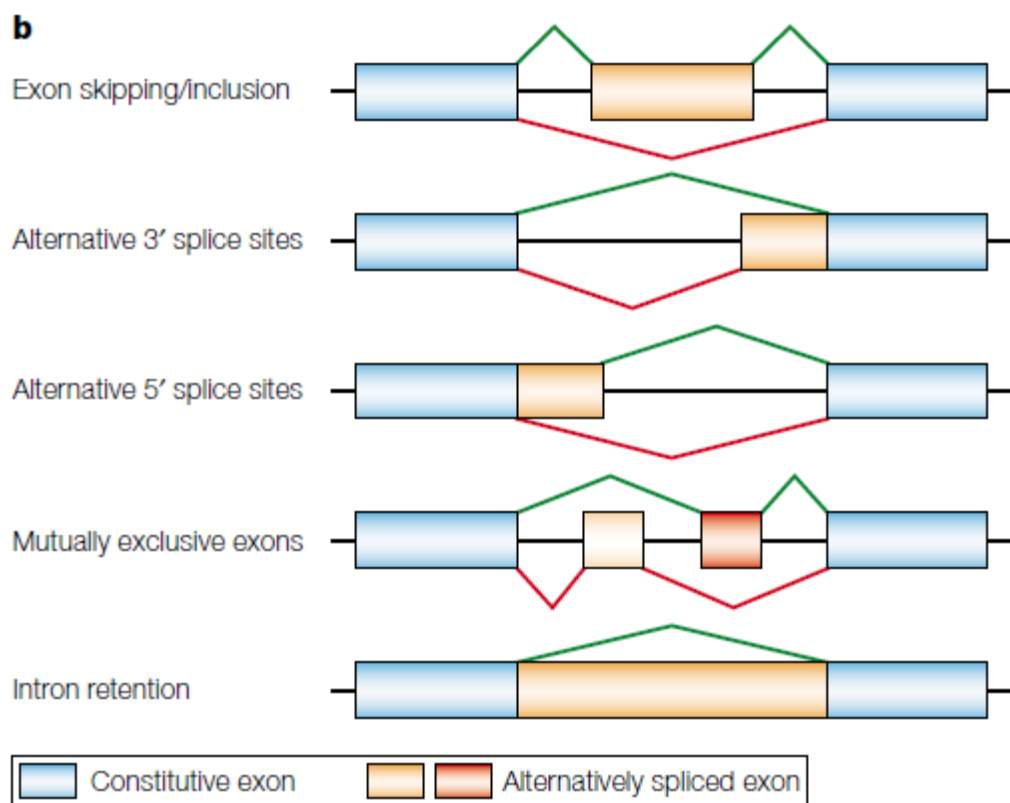


Figure 2 (d'après Cartegni et al . Nat Genet Rev 2002)

L'étude de l'épissage des gènes porteurs d'un variant de signification inconnue est complexe, mais son intérêt est fondamental, de par la fréquence des variants potentiellement concernés, et de la nécessité de caractériser les mutations identifiées chez les familles atteintes. Deux approches sont utilisées. La première consiste à étudier directement l'expression des ARN et identifier des transcrits anormaux. La deuxième utilise la technique de minigène dans lequel a été inséré une séquence comprenant ou non la mutation. Cette séquence idéalement est un ARN extrait à partir de lymphocytes (test in vivo), ou à partir d'ADN génomique dont on a amplifié les séquences (test in vitro).

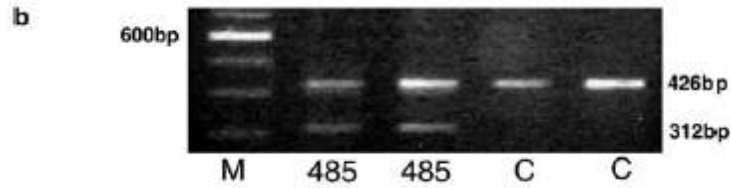
La première technique est décrite par Pagentecher et al (Pagentecher et al., 2006). Les étapes sont les suivantes :

- Extraction des ARN (PAXgene blood RNA kit)
- Synthèse d'ADNc (RT-PCR)
- Séparation par électrophorèse des fragments
- Séquençage des fragments

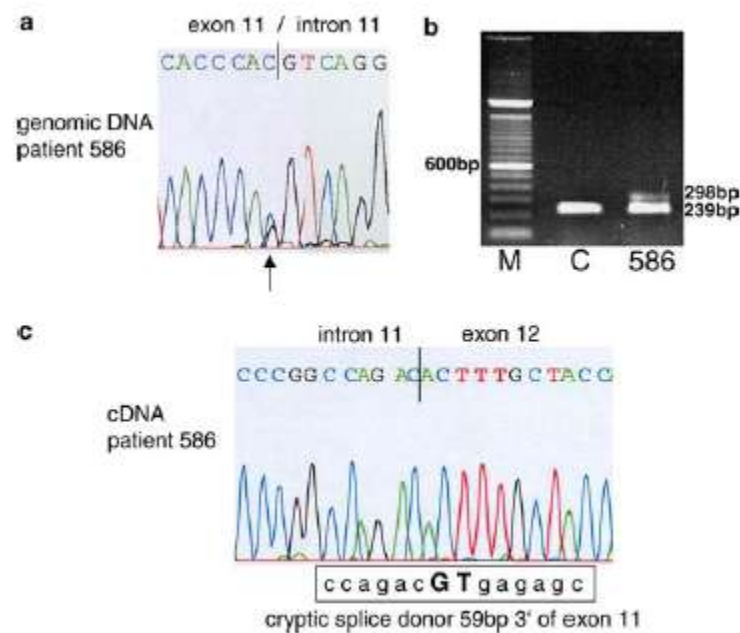
Dans cette technique, une mutation est considérée comme pathogène lorsqu'aucun transcrit de l'allèle mutant n'est observé, neutre quand les transcrits sont identiques à ceux de l'ADN génomique. Un effet partiel sur l'épissage est décrit lorsque des transcrits différents ne sont pas retrouvés chez les contrôles.

Pagentecher et al ont analysé 19 substitutions de séquences codantes ou non de *MSH2* et *MLH1* et démontré que 10 d'entre eux affectaient l'épissage.

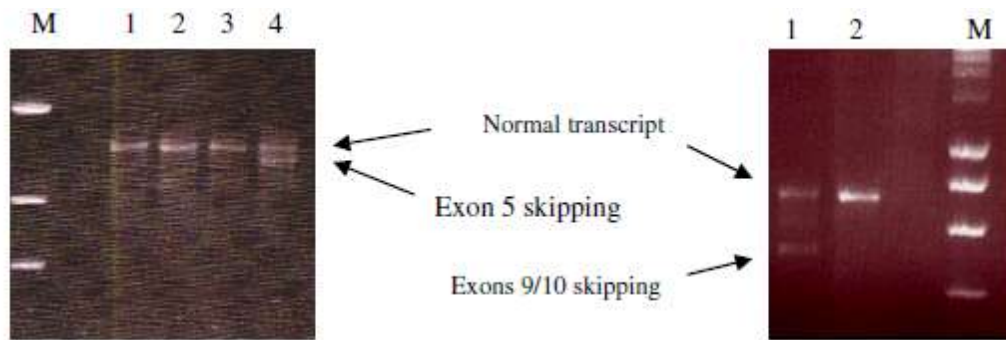
Par exemple, la mutation de *MLH1* c.2013G>C de l'exon 18 (p.Gln701His) a été étudiée en amplifiant par RT-PCR les exons 16 à 19. Un fragment de 312pb en plus d'un fragment de 426 pb attendu a été observé. Après séquençage, ce fragment correspond à un saut de l'exon 18.



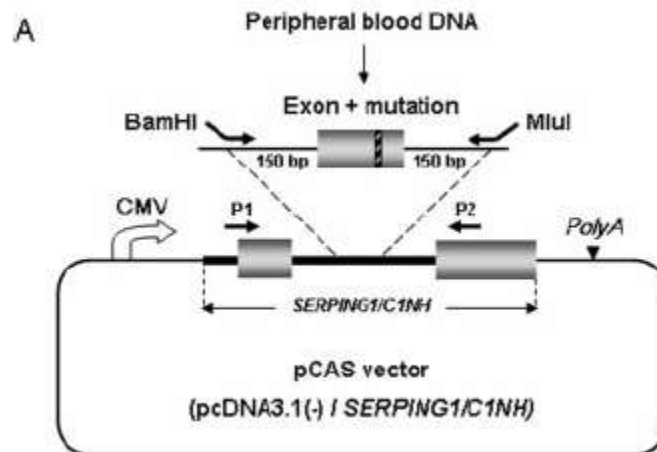
De même, la substitution c.1038G>C de l'exon 11 de *MLH1* (p.Gln346His) entraîne l'apparition d'un transcrit légèrement plus grand que celui du contrôle. Après séquençage, il s'agit d'une séquence additionnelle de 59pb secondaire à l'apparition d'un site cryptique d'épissage dans l'intron 11, affectant l'épissage des exons 10 et 11.



La deuxième technique décrite par Auclair et al utilise la technique de minigène à partir d'ARN extraits de lignées cellulaires lymphoblastoïdes issues de patients porteur d'un variant MMR. Les ARN sont transformés en ADNc par RT-PCR puis intégrés à un vecteur, lequel est transfecté dans des cellules COS-7. Les ARN extraits des lignées sont ensuite séparés et séquencés. Dans ce travail, 60 variants étaient étudiés (45 exoniques, 15 introniques). Quinze d'entre eux (5 et 10) étaient associés à un épissage anormal avec principalement un saut d'exon.



Pour finir, Tournier et al (Tournier et al., 2008), soulevant le fait que l'ARN est rarement disponible, ont développé une technique permettant d'analyser les conséquences sur l'épissage de variants à partir d'ADN génomique (test ex vivo). Dans cette technique, les exons comportant la substitution, flanqués d'extrémités 5' et 3' de 150 pb (incluant un site de restriction) sont amplifiés puis insérés dans un plasmide. Ce dernier est transfecté dans des cellules Hela, puis les ARN sont extraits, séparés et séquencés.



Gorlov et al (Gorlov et al., 2003) ont démontré que les variants pathogènes étaient significativement associés aux motifs ESE (57 vs 24%. $P=0.02$ pour *MHS2*, 58 vs 38%. $p=0.3$ pour *MLH1*). Par ailleurs, la distribution des variants dans le gène était

également très significativement corrélée à celle des motifs ESE, puisque plus de 80% d'entre eux se trouvaient dans les 20 derniers nucléotides de l'exon.

FREQUENCE ALLELIQUE

Ce paramètre est parfois difficile à obtenir car par définition, les variants étudiés sont rares et ont le plus souvent des spécificités géographiques ou ethniques. En pratique, si un variant est retrouvé chez une centaine d'individus, ou à une fréquence supérieure à 1% dans la population générale, il est peu probable qu'il soit un variant à haut risque.

CO OCCURRENCE

Le principe général repose sur le fait que si un variant pathogène est associé à l'agrégation familiale d'une pathologie, il est peu probable qu'il y ait un deuxième variant associé. De fait si c'est le cas, ce dernier aura peu de chance d'être un variant pathogène. Par ailleurs, le phénotype associé aux variants permet de raisonner sur la causalité d'un variant. Il est ainsi démontré par exemple, qu'un individu porteur de deux allèles pathogènes du gène BRCA1 n'est pas viable (Liu et al., 1996). De même, la présence en *trans*, d'un deuxième variant du gène BRCA2 entraîne l'apparition d'un phénotype de type maladie de Fanconi. Dans ce cas la pathogénicité du variant est démontrée par la constatation de ce phénotype (Howlett et al., 2002).

En ce qui concerne les gènes MMR, la co-occurrence de variants pathogènes est associée à un phénotype très agressif avec des âges très précoces d'apparition de cancers (Agostini et al., 2005, Auclair et al., 2006, Rahman et al., 2007), ou, pour le gène *MLH1*, à un phénotype de neurofibromatose (Wang et al., 1999)

CRITERES DE QUALITE DES BASES LSDB ET ROLE DES CURATEURS

La gestion des bases de données implique une grande rigueur, tant les conséquences des erreurs sont importantes, notamment en termes de conseils génétique (Soussi T et

Bérout C. 2006). Une confusion peut exister, sachant par exemple qu'un variant de *MLH1* (R217C) est défini comme pathogène (Fan et al., 2007), neutre (Takahashi et al., 2007, Trojan et al., 2007) ou VSI (Ellison et al., 2001).

Soussi et al, on ainsi proposé un référentiel à l'intention des gestionnaires de bases de données afin de garantir la qualité des données, et notamment lors de la publication dans les revues spécialisées de ces bases. En effet, la présence d'erreurs peut avoir d'importantes conséquences surtout lors que celles-ci sont consultées par des non-spécialistes, mais également peut masquer des informations pertinentes si celles-ci ne sont pas reprises dans la base. Ainsi, il est recommandé :

- que toutes les mutations soient reportées, y compris les mutations « neutres » qui ne changent pas l'acide aminé, dans la mesure où elles sont susceptibles d'affecter l'épissage ou la stabilité de l'ARN
- De réaliser une deuxième analyse sur une deuxième extraction d'ADN, ou un contrôle négatif sur un prélèvement inclus en paraffine
- D'être prudent sur l'interprétation de variants fréquemment retrouvés, ayant une activité normale, ou rarement reportés dans la littérature
- De réaliser des vérifications systématiques afin de dépister des erreurs typographiques (10-20% selon les bases)
- D'utiliser la nomenclature officielle pour la description des variants.

Deux particularités importantes conditionnent également la qualité de ces bases (Claustres et al., 2002):

- leur mise à jour régulière par les curateurs est indispensable
- chaque base présente un contenu et une structure dépendants du gène et des caractéristiques de la maladie. Il est pourtant essentiel que des éléments centraux soient utilisés, tels que :
 - identifiant unique pour chaque allèle
 - source de la donnée (publication, auteur)
 - contexte de l'allèle : espèce, gène, séquence de référence
 - description de l'allèle : nom, type, changements nucléotidiques...

Claustres et al ont réalisé une revue de 94 bases de données indépendantes en 2002, montrant que 75% des données étaient issues de la littérature. Les mutations étaient soumises aux curateurs selon un questionnaire type (HUGO-MIDI) dans 68% des cas afin de permettre la validation par les experts. La moitié seulement des bases fournissait une information sur la maladie et le gène, l'autre donnant simplement un lien vers OMIN ou d'autres sites cliniques. La bonne nomenclature des mutations n'était utilisée que dans 42% des et 36% des bases ne mentionnaient pas de date de mise à jour. Les auteurs de ce travail ont élaboré un type « idéal » de LSDB (cf. annexe 7)

UNIVERSAL MUTATION DATABASE (UMD)

Au début des années 90, Bérout et al. décidaient de développer une BDD dynamique, incluant des outils permettant de caractériser les variants inclus dans leur base. Ils développèrent ainsi un logiciel appelé « UMD » (Bérout et al., 2000), actuellement reconnu par la Human Genome Organization (HUGO) et la Human Genome Variation Society (HGVS) comme un outil de référence pour élaborer une BDD de type LSDB. Ils ont ainsi utilisé ce logiciel pour développer la première base de variants du gène p53 puis APC (cancer du colon) (Bérout, C. & Soussi, T. 1997), BRCA1 and BRCA2 (cancer du sein), MEN1 (multiple endocrine neoplasia type 1) (Wautot et al., 2002), sulfonyleurea receptor, SUR1 (hyperinsulinism21), RB1 (retinoblastoma), VHL (von Hippel-Lindau Syndrome) (Bérout et al., 1998) et WT1 (Wilms tumour) (Jeanpierre et al., 1998). Cette base de donnée utilise le langage 4D (www.4D.com) et a fait l'objet d'une mise à jour en 2005 (Beroud et al., 2005). Dans cette base, l'ensemble des séquences codantes et non codantes est disponible, ainsi que les positions des SNP. Les données cliniques et biologiques de chaque variant sont collectées et consultables (Cf. annexe 8). L'outil dispose d'une fonction de vérification de la séquence des variants soumis afin d'éviter les erreurs typographiques ou de nombre. Par ailleurs, le logiciel évalue les sites potentiels donneur/accepteur d'épissage en calculant un score (« consensus value ») dérivé de l'algorithme de Cartegni (Cartegni et al., 2002). Plus ce score est proche de 100, plus la séquence est susceptible d'être un site d'épissage (Cf. UMD Predictor. Annexe 6). En outre, le logiciel dispose de données de corrélation génotype/phénotype et montre la distribution géographique des variants sur le gène.

RESULTATS

Grandval P, Fabre AJ, Olschwang S. Design of a core classification process for DNA mismatch repair variations of a priori unknown functional significance. Hum Mutat. 2013 Jun;34(6):920-2. doi: 10.1002/humu.22312.

L'analyse de la littérature concernant les différentes méthodologies utilisées pour classer les VSI montre une grande hétérogénéité, traduisant la difficulté d'intégrer l'ensemble des critères et outils vus précédemment. Sur la base de leur expertise en termes de tests fonctionnels, l'équipe de N. De Wind (Rasmussen et al., 2012) a proposé un algorithme de classement des VSI comportant trois étapes :

1. analyse de l'histoire familiale, statut MMR (immunohistochimie et analyse constitutionnelle)
2. analyses fonctionnelles du système MMR, in silico et expérimentale
3. si insuffisant :
 - a. analyse de l'épissage
 - b. caractérisation détaillée des variants (localisation des protéines, propriétés biochimiques et biophysiques des dimères Mut S et L).

Au début de notre travail, 707 variants des gènes MMR de la base de données française étaient des VSI. Il n'était pas possible en France d'effectuer des tests de la fonction MMR (cf. matériels et méthodes). En revanche, parmi les laboratoires français impliqués dans l'analyse et la caractérisation des gènes du syndrome de Lynch, certains ont une expertise reconnue dans les analyses de l'épissage (Lyon, Rouen, Lille). Par ailleurs, le fonctionnement en réseau des laboratoires, et l'organisation des consultations d'oncogénétique, permettent de disposer de données cliniques et biologiques dans un grand nombre de situations. Nous avons fait l'hypothèse qu'une approche systématisée,

tenant compte de critères cliniques, biologiques, informatiques, en complément des analyses d'épissage lorsqu'elles étaient pratiquées, nous permettrait de classer un certain nombre de VSI, sans recours systématique aux tests fonctionnels.

MATERIELS ET METHODES :

En complément de l'algorithme proposé par Rasmussen et al. que nous trouvions très restrictif, nous avons établi un algorithme permettant de classer les VSI selon la nouvelle classification internationale en 5 classes (I : neutre, II : probablement neutre, III : VSI, IV : probablement causal, V : causal). Trois types de données étaient analysés :

1. in silico
2. cliniques (critères d'Amsterdam-ségrégation)/génotypiques (statut MMR)/épissage
3. données publiées (tests fonctionnels et fréquence allélique)

Nous avons ainsi défini les critères nécessaires pour la caractérisation des VSI :

- données in silico: Conservation, SIFT, UMD Predictor, NNSplice
- données cliniques: critères d'Amsterdam, analyse de co-ségrégation (LOD score>1.5)
- statut des Microsatellites (génotypage, IHC)
- tests fonctionnels
- analyses d'épissage (splicing reporter Minigene / RT-PCR)
- co occurrence d'une mutation délétère (sur un autre gène, ou sur le même gène en trans, sans phénotype aggravé)
- fréquence allélique.

Pour chaque classe, nous avons défini les différentes situations possibles, et les critères minima requis pour classer correctement les variants. Nous avons considéré par exemple, qu'une anomalie majeure de l'épissage, une fonction MMR déficiente, une ségrégation significative (LOD score>1.5), suffisaient à eux seuls à classer un variant

comme causal. Dans les situations où ces tests étaient absents ou normaux, l'interprétation des données familiales, biologiques et in silico permettaient, selon qu'elles étaient concordantes ou non, de proposer une classification en neutre ou causal dans certains cas, ou probablement neutre ou causal la plus part du temps. Dans ces cas, une analyse complémentaire ou une demande de renseignements clinique était formulée par le réseau afin de préciser la pathogénicité de ces variants.

RESULTATS

Grace à cet algorithme, nous avons pu caractériser :

- 84 variants neutres (classe 1), présentant des discordances d'analyses somatiques et/ou une co occurrence de mutation délétère
- 6 variants, publiés comme délétères (classe 5), ne remplissant pas les critères et reclassés en classe 4 (probablement délétères)
- 13 variants délétères (classe 5) du fait d'une fonction MMR altérée (4), d'un épissage altéré (3), ou de données de ségrégation et clinico biologiques dans plusieurs familles (consensus du réseau)



Grandval P, Fabre AJ, Gaildrat P, Baert-Desurmont S, Buisine MP, Ferrari A, Wang Q, Bérout C, Olschwang S. UMD-*MLH1/MSH2/MSH6* databases: description and analysis of genetic variations in French Lynch syndrome families. Database (Oxford). 2013 May 31;2013:bat036. doi: 10.1093/database/bat036.

A l'origine de ce travail, l'Institut National du Cancer (INCA) a subventionné la création d'une base de données afin de recenser les variants des gènes du syndrome de Lynch au sein de la population française. Le réseau français des laboratoires impliqués dans le syndrome de Lynch (liste en annexe 8), sous la coordination du Dr S. OLSCHWANG, a recensé depuis 1995 les variants identifiés chez les individus suspects de ce syndrome. Le développement d'un algorithme de caractérisation des variants (Cf. publication n°1), et du logiciel UMD par l'équipe du Pr. C. BEROUD, déjà opérationnel pour d'autres gènes nous a incité à soumettre à la communauté scientifique l'ensemble des données françaises au travers d'un outils performant, et ayant la possibilité d'intégrer un grand nombre de données cliniques, génétiques et biologiques.

MATERIELS ET METHODES

La collecte des données respectait les règles de bonnes pratiques de la consultation d'oncogénétique et des contrôles qualités des laboratoires. Chaque nouveau variants était soumis au curateur de la base (Dr S. OLSCHWANG).

La caractérisation de la pathogénicité était faite selon l'algorithme décrit précédemment, et tous les variants de classe 2 (probablement neutre) ou 4 (probablement causal)

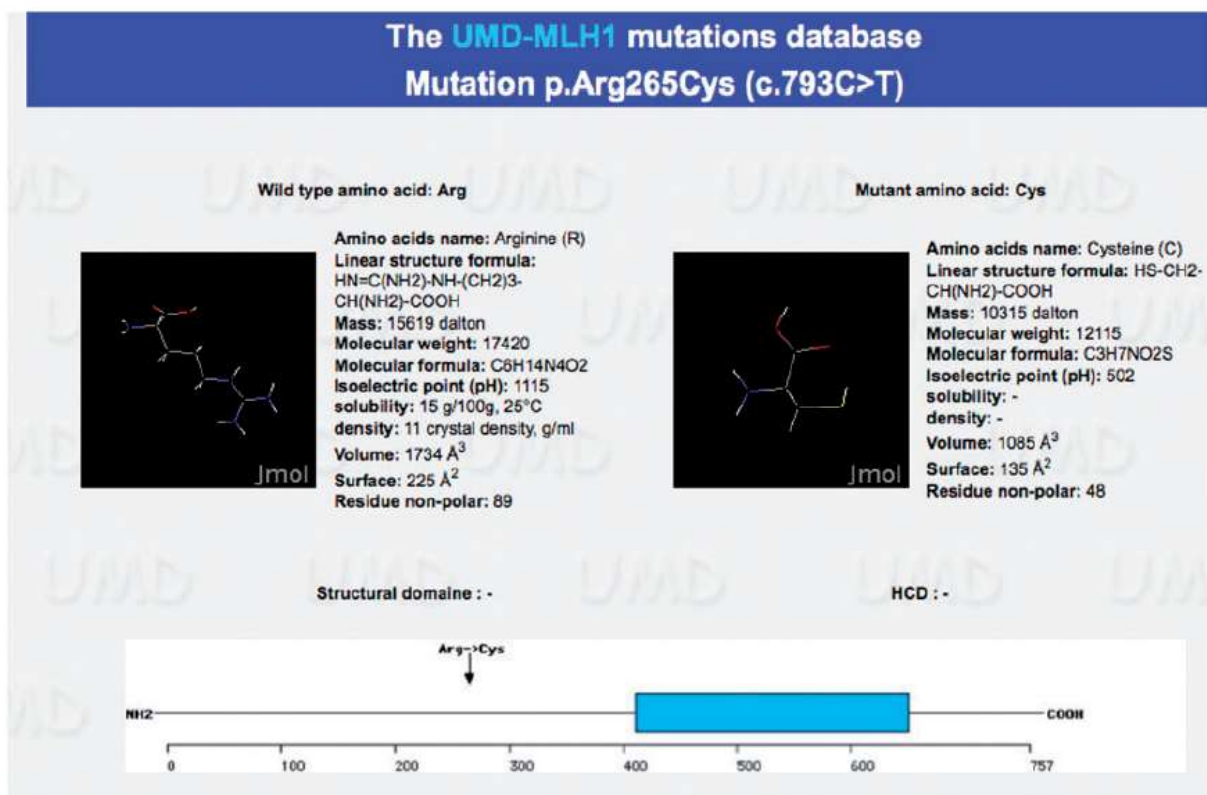
étaient régulièrement revus en fonction des éventuelles données cliniques ou biologiques.

Une base de donnée spécifique aux trois gènes MMR (*MLH1*, *MSH2* et *MSH6*) était élaborée dans l'outil UMD (Cf. chapitre matériel et méthodes).

RESULTATS

Une base pour chaque gène MMR a été développée (www.umd.be/MLH1/, www.umd.be/MSH2/ et www.umd.be/MSH6/). Six parties distinctes sont développées sur le site. Trois concernent le gène, la protéine et les éléments cliniques. Deux permettent de consulter l'ensemble des variants et avoir une vue globale des types de mutations (« statistiques »). Enfin, le dernier outil offre des fonctions analytiques permettant de rechercher un variant selon le type de mutation, sa localisation intronique ou exonique ou encore sa caractérisation biologique.

Pour chaque variant, les propriétés physico-chimiques de l'acide aminé normal et muté sont décrites ainsi que la position sur la séquence protéique.



Les données des tests in silico (SIFT et UMD –predictor) ainsi que l'ensemble des données cliniques (critères d'Amsterdam, ségrégation), biologiques (statut MMR), et fonctionnelles (fonction MMR et épissage) sont disponibles.

Conservation (0-1)	SIFT (0-1)	UMD-predictor (0-100)	mmruv database
1	0.05 (non pathogenous)	100 (Pathogenous)	p.Arg226Gln

Analysis	Result	Date	Origin	PMID/dbSNP
MMR function in tumor cells	<i>MLH1</i> -	17/10/08	31 (Montpellier)	---
In vitro analysis	RT-PCR: loss of exon 8	23/05/08	"Pagenstecher et al. Hum Genet 2006;119:9-22"	16341550
MMR function in tumor cells	<i>MLH1</i> - (patient B)	5/07/02	SO (Marseille)	---
MMR function in tumor cells	MSI	17/10/08	31 (Montpellier)	---
Clinical phenotype	Amsterdam I +	17/10/08	31 (Montpellier)	---
Segregation analysis	LOD score = 0.6	17/10/08	31 (Montpellier)	---
Ex vivo analysis	Splicing reporter minigene pCAS: loss of exon 8	26/05/08	19 (Rouen)	---
MMR function in tumor cells	<i>MLH1</i> - (patient A)	24/12/02	SO (Marseille)	---
MMR function in tumor cells	MSI (patient A)	24/12/02	SO (Marseille)	---

Finalement, la caractérisation du variant selon le modèle 5 classes (1 : neutre/ 2 : probablement neutre/ 3 : VSI/ 4 : probablement causal/ 5 : causal) est proposé après validation par le panel d'expert du réseau et le curateur.

Biological significance	Date	Comment
Causal	16/06/08	---

Les données publiées concernent les celles des 3 gènes principaux MMR telles que nous les avons analysées en juin 2012. Sur les 20000 individus adressés en consultation d'oncogénétique pour suspicion de syndrome de Lynch entre 1995 et 2012, 7047 variants ont été retrouvés. Après exclusion des variants associés à de grands réarrangements génomiques, une insertion ou délétion dans des sites d'épissage ou dans les extrémités 5' et 3' UTR, 1174 variants distincts ont été analysés. Parmi eux, 40%

étaient des mutations tronquantes (171-*MLH1*, 199-*MSH2*, 97 *MSH6*), et 707 restaient des VSI, dont 370 d'entre eux ont pu être caractérisés en classes 1/2/4/5.

A la suite de la publication, l'actualisation pour l'année 2013 de l'ensemble des variants de la base MMR a permis de classer de nouveaux variants.

	TOTAL	Classe 1 Neutre	Classe 2 Proba. Neutre	Classe 3 VSI	Classe 4 Proba. causal	Classe 5 Causal
<i>MLH1</i>	255	63	18	100	23	51
<i>MSH2</i>	280	64	33	136	24	23
<i>MSH6</i>	203	57	31	104	9	2
TOTAL	738	184	82	340	56	76

Ainsi, sur un total de 738 VSI, 260 variants ont été classés en causal ou neutre, dans le respect d'une méthodologie validée, et donc utilisables pour le conseil génétique. La répartition des caractéristiques principales des variants de classe 5 est précisée dans le graphe suivant.

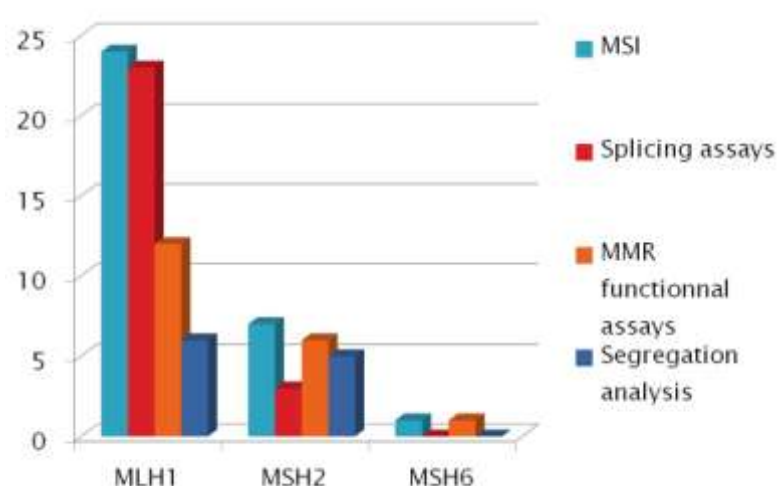


Figure : répartition des variants de classe 5 (causal), selon la catégorie de tests anormaux et données génétiques, et le gène

Grandval P, Baert-Desurmont S, Bonnet F, Bronner M, Buisine MP, Colas C, Noguchi T, North MO, Rey JM, Tinat J, Toulas C, Olschwang S. Colon-specific phenotype in Lynch syndrome associated with EPCAM deletion. Clin Genet. 2012 Jul;82(1):97-9.

En 2009, Ligtenberg et al (Ligtenberg et al., 2009) ont décrit dans des familles chinoises et des Pays Bas, des cancers colorectaux (CCR) avec absence d'expression de la protéine *MSH2*, associées à des délétions constitutionnelles des deux exons du gène *TACSTD1* codant pour la protéine EPCAM. Le promoteur de *MSH2*, situé en amont du gène *TACSTD1* et en *cis*, s'en trouve inactivé par un phénomène de méthylation directement lié à la mutation du gène EPCAM.

Cette extinction est donc dépendante de l'expression de la protéine EPCAM, et donc du niveau d'expression tissulaire de cette dernière. De forts taux ont été rapportés dans les cellules souches colo rectales expliquant les risques élevés de CCR chez les patients mutés. En revanche, les niveaux dans les tissus extra coliques sont très variables et les risques et l'évolution des cancers associés sont eux même potentiellement modifiés en fonction des conséquences des mutations du gène EPCAM sur la protéine, cette dernière étant une molécule d'adhésion et de prolifération.

Kempers et al (Kempers et al., 2011) ont ainsi étudiés 194 porteurs issus de 41 familles et comparé les risques de cancers avec une cohorte de patients mutés pour un des 3 gènes MMR, ou présentant une double mutation MMR et EPCAM. Sur les 194, 93 avaient développé un CCR. Le risque cumulé de CCR avant 70 ans était estimé à 75% (IC 95% 65-85), à un âge moyen de 43 ans, ce qui était similaire aux porteurs d'une double mutation ou d'une mutation de *MLH1* et *MSH2*. De façon surprenante, la prévalence du cancer de l'endomètre chez la femme était en revanche significativement plus faible que pour les mutations des gènes MMR (12%-IC 95% 0-27 vs 34-55% pour *MSH2* et *MSH6*. $P < 0.0001$ et 0.03).

Une deuxième publication de Lynch et al., parue en 2011, analysant deux grandes familles mutées pour EPCAM, confirmait également la faible incidence de cancers extra coliques, et en particulier, l'absence de cancers de l'endomètre (Lynch et al., 2011).

Le but de notre travail était d'évaluer l'adéquation des recommandations de dépistage au regard de la traduction phénotypique des délétions du gène EPCAM.

MATERIELS ET METHODES

720 cas index non liés génétiquement étaient recensés dans la base nationale Française des mutations des gènes MMR. L'étude portait sur les patients chez qui aucune altération du gène *MSH2* n'avait été identifiée bien que son inactivation ait été mise en évidence dans les tumeurs par immunohistochimie. Une étude de l'extrémité 3' du gène EPCAM était réalisée par MLPA ou QMPSF. Les données phénotypiques des familles avec délétion étaient collectées.

RESULTATS

Sept cas index présentant une délétion du gène EPCAM étaient identifiés. Quatre patients remplissaient les critères d'Amsterdam et 3 présentaient un cancer colorectal précoce sporadique, suggérant une mutation de novo. Les tests prédictifs réalisés dans les familles de ces patients permettaient la détection de 12 porteurs de la délétion et de 5 porteurs obligatoires (6 patients décédés n'ont pas été testés). Au total, 27 patients (17H/10F) présentaient un cancer digestif colorectal (à l'exception d'un adénocarcinome duodénal), pour lequel l'âge moyen de survenue était de 43,6 ans. Huit patients avaient développé un second cancer colorectal, 2 à 39 ans après le premier. Trois patients porteurs de la mutation, dont une femme de 83 ans, ne présentaient pas de cancer, suggérant une pénétrance incomplète. Aucun cas de cancer de l'endomètre n'était diagnostiqué chez les 10 femmes.

En conclusion, les mutations du gène EPCAM expliquent environ 1% des syndromes de Lynch et sont responsables d'une inactivation de *MSH2*. L'âge de survenue des cancers colorectaux y est similaire à celui observé pour les autres mutations des gènes MMR. Cette étude confirme les données récentes suggérant que l'atteinte de l'endomètre est beaucoup plus rare que dans le syndrome de Lynch classique, ce qui pourrait inciter à modifier les recommandations de prise en charge.

Grandval P, Barouk-Simonet E, Bronner M, Buisine MP, Moretta J, Tinat J, Olschwang S. Is the controversy on breast cancer as part of the Lynch-related tumor spectrum still open? Fam Cancer. 2012 Dec;11(4):681-3.

Dans sa description du spectre large du syndrome HNPCC, Lynch décrivait ces localisations à risque de cancers extra coliques (endomètre, ovaire, estomac, voies biliaires, urothélium, intestin grêle). Une controverse concernant d'autres localisations est apparue, renforcée par la publication par Win et al. (Win et al., 2012) de cancers pancréatiques et du sein. Dans ce travail, décrivant les risques de cancers dans les familles principalement américaines, les auteurs rapportaient des incidences ratio standardisés de 10.68 (95% CI, 2.68 à 47.70; P.001) et 3.95 (95% CI, 1.59 à 8.13; P .001) respectivement. Il est important de noter, que dans ce travail, aucune mention d'instabilité des microsatellites n'était faite. En revanche, la cohorte de patients français suivis dans le cadre de l'enquête ERISCAM (Bonadona et al., 2011), ne rapporte pas de sur risque significatif pour ces deux localisations.

Le but du travail était d'analyser les données de la base française des variants du syndrome de Lynch afin d'identifier les situations où un cancer du sein était identifié.

MATERIELS ET METHODES

Nous avons recensé l'ensemble des familles françaises incluses dans la base du réseau, pour lesquelles un des apparentés, porteur d'une mutation d'un gène MMR, avait présenté un cancer du pancréas ou du sein. Un génotypage de chaque tumeur était réalisé à la recherche d'une instabilité des microsatellites, complétée lorsque cela était possible d'une immuno histochimie.

RESULTATS

Quinze individus chez qui on retrouvait des critères d'Amsterdam, étaient porteur d'un variant causal du syndrome de Lynch et ont pu être analysés. Quatorze d'entre eux avaient eu un cancer du sein et un patient avait eu un cancer pancréatique. Seul deux

patients n'avaient présenté qu'un cancer du sein, sans autre cancer du spectre HNPCC (mutations BRCA 1 et 2, a priori exclues).

Aucune des 15 tumeurs testées ne présentait d'instabilité des microsatellites. Une seule était MSI-L, mais l'immunohistochimie ne révélait aucune perte d'expression des 4 protéines MMR. En revanche, les tumeurs du spectre de Lynch observées chez ces mêmes patients démontraient, lorsque la tumeur était disponible, une instabilité associée à une perte d'expression de la protéine correspondant au gène muté.

Ces résultats, bien que concernant un faible nombre de patients, laissent penser que le cancer du pancréas et le cancer du sein ne font pas partie du spectre du syndrome de Lynch. Dans tous les cas, une analyse tumorale à la recherche d'une instabilité des microsatellites doit être réalisée.

CONCLUSIONS ET PERSPECTIVES

Le but de ce travail était de soumettre une description exhaustive des variants du syndrome de Lynch recensés par le réseau français des laboratoires, à l'ensemble de la communauté scientifique et médicale impliquée dans la prise en charge des prédispositions héréditaires aux cancers digestifs. Cela a été possible grâce à l'expérience dans le domaine, de l'ensemble des généticiens et biologistes du réseau, et également grâce à l'expertise reconnue du laboratoire de génétique médicale et génomique fonctionnelle, plus spécifiquement du Pr BEROUD, responsable de l'équipe de bioinformatique. Cette dernière, en plus des compétences en génétique, a en effet accumulé une expérience importante dans l'élaboration des bases de données, et connaît les exigences requises pour la caractérisation des variants. Dans ce domaine, les modalités de validation de la pathogénicité d'un variant, et les critères de construction et de suivi d'une base de données sont principalement issues du collégial InSiGHT, dont les membres sont à l'origine d'une grande partie des publications dans le domaine du syndrome de Lynch. La multiplication, ces dernières années, des variants du syndrome de Lynch, a rendu la gestion des bases de mutations et l'accessibilité aux données d'autant plus complexes. Il apparaissait donc primordial que la base utilisée par le réseau français affiche une méthodologie de caractérisation adaptée et validée, afin de fournir aux cliniciens en charge des familles des résultats a priori certains, et/ou re examinés dans le temps au gré des nouvelles données et événements.

Parallèlement à notre réflexion, Rasmussen et al. (Rasmussen et al., 2011) ont proposé une méthodologie de classification essentiellement basée sur les tests fonctionnels, ces derniers étant, il est vrai, le « gold standard » pour évaluer les conséquences d'un variant sur la fonction MMR. Cependant, à l'exception de ce groupe et de quelques autres, cet algorithme n'est pas utilisable tel quel, car trop d'étapes dépendent de tests rarement disponibles en laboratoire (levures, notamment). Nous avons considéré que la caractérisation des variants ne pouvait s'exonérer des données cliniques et tumorales, et que l'organisation de la consultation génétique en France permettait d'obtenir dans la plupart des cas ces éléments qui, en revanche, sont rarement accessibles à travers les bases publiées. Nous avons donc défini les situations permettant de classer les variants, et cette méthodologie est actuellement utilisée à chaque réunion semestrielle du réseau

des laboratoires français. Ainsi, en présence d'un test fonctionnel et/ou d'une anomalie majeure de l'épissage, la caractérisation d'un variant en classe 5 (causale) est évidente. Mais en l'absence de tels éléments, faute de pouvoir réaliser les études, la présence d'une coségrégation avec un LOD score > 1.5 , et/ou une concordance entre le génotypage MMR et l'immunohistochimie, en association avec une présentation clinique typique (critères d'Amsterdam avec au moins deux tumeurs analysables dans la famille), nous avons considéré que le variant était causal. Rasmussen et al. dans la même revue, ont répondu à notre publication, reconnaissant que les tests fonctionnels restaient l'apanage de laboratoires spécialisés, mais argumentant également le fait que selon les pays, la réalisation des analyses constitutionnelles ciblées chez les apparentés n'était pas systématique, compliquant la possibilité de conduire une analyse de ségrégation. Par ailleurs, les auteurs soulevaient à juste titre que la pénétrance du syndrome de Lynch était incomplète, autre obstacle à la caractérisation clinique, et que les génotypages et immunohistochimies pouvaient comporter des faux positifs et négatifs. Rasmussen et al., concluaient de la façon suivante : "Until that time has come, the approach taken by Grandval et al. certainly is the best alternative" (Rasmussen et al., 2013).

Consciente de cette problématique, la même équipe a proposé un test *in vitro*, rapide et reproductible afin d'évaluer la fonction MMR des variants (Drost et al., 2012). Dans ce test, un plasmide contenant un mésappariement G-T est introduit dans un site de restriction, le rendant ainsi non clivable par l'enzyme, sauf à être réparé par un hétérodimère MutS (*MSH2-MSH6*) sauvage donc fonctionnel, produit après PCR par des cellules en culture (LoVo). Les auteurs ont évalué cette technique sur une série de 25 variants du gène *MSH2* et ont démontré une corrélation avec les tests *in silico* dans 83% des cas. Cependant la comparaison n'a pas été faite avec les tests fonctionnels de référence, recommandés d'ailleurs par la même équipe.

Comme il a été souligné précédemment, l'analyse de l'épissage est un élément crucial dans la caractérisation des variants, et dans ce domaine, plusieurs laboratoires français font référence et proposent, lorsque le réseau des laboratoires le demande, de réaliser un complément d'analyse pour des variants de classe 2, 3 ou 4. Dans notre travail, il apparaît effectivement que les variants pathogènes (classe 5) sont plus souvent caractérisés par une anomalie d'épissage que par un test fonctionnel anormal (14.1 vs 10.3% des variants classés), et que le gène *MLH1* est celui qui est associé le plus

fréquemment à ce type d'anomalie (36,5% pour *MLH1*, 4,7% pour *MSH2* et 0,0% pour *MSH6*). Dans le précédent travail de Drost et al., les auteurs rapportaient que les 20 variants pathogènes de *MHS6* testés étaient, de façon surprenante, capables de restaurer une fonction MMR déficiente contrairement à l'effet attendu. Les auteurs suggéraient la possibilité qu'une anomalie de l'épissage puisse expliquer le caractère pathogène des variants. Cependant, en analysant les variants testés dans ce travail, il apparaît que 8 sur les 20 sont déjà décrits dans notre base. Six variants sont classés neutres (c.73G>T, c.431G>T, c.1186C>G, c.1508C>G, c.2633T>C, c.3062C>A) et 2 sont des VSI (c.1474A>G, c.3260C>G). Quoiqu'il en soit, aucun de ces variants n'a été classé comme délétère (classe 5) selon notre algorithme, et en particulier aucun n'avait d'anomalie décrite de l'épissage. Ceci souligne l'extrême difficulté à caractériser correctement la pathogénicité des variants, alors même que les conséquences en termes de conseils génétique sont majeures. C'est également pour cette raison qu'il est excessivement difficile de proposer une évaluation des performances de notre algorithme et qu'un panel d'experts reste indispensable pour valider les résultats comme le suggérait Plon et al. (Plon et al., 2008).

En alternative aux tests *in vivo* et *in vitro*, les analyses *in silico* représentées par les outils de prédiction informatique, sont certainement une perspective concrète, rapide et peu coûteuse, d'amélioration de la caractérisation des variants. Ces outils offrent l'énorme intérêt de pouvoir répondre en temps réel à la prévisible masse de données produites par les techniques de séquençage haut débit dites NGS (Next generation sequencing), pour lesquelles il semble difficile d'imaginer la réalisation de plusieurs tests nécessitant par exemple des techniques de mutagénèse.

Pour conclure, ce travail collaboratif a permis de développer une méthode validée et standardisée de caractérisation des variants de signification inconnue dans le syndrome de Lynch. Une base de données spécifique à chacun des trois gènes majeurs de ce syndrome a été élaborée selon les recommandations internationales de classification et de recueil des données. Cet outil, évolutif dans le temps, fait l'objet de mises à jour périodiques et est accessible en ligne pour l'ensemble de la communauté scientifique et médicale. Son développement pour les gènes de la polypose familiale (*APC* et *MUTYH*) est en cours.

REFERENCES BIBLIOGRAPHIQUES

A

Aaltonen LA, Salovaara R, Kristo P, Canzian F, Hemminki A, Peltomaki P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med*. 1998;338:1481-1487.

Aarnio M, Sankila R, Pukkala E. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer* 1999;81:214-218.

Agaram NP, Shia J, Tang LH, Klimstra DS. DNA mismatch repair deficiency in ampullary carcinoma: a morphologic and immunohistochemical study of 54 cases. *Am J Clin Pathol*. 2010 May;133(5):772-80.

Agostini M, Tibiletti MG, Lucci-Cordisco E, Chiaravalli A, Morreau H, Furlan D, et al. Two PMS2 mutations in a Turcot syndrome family with small bowel cancers. *Am J Gastroenterol* 2005;100:1886–1891.

Alexander J, Watanabe T, Wu T, Rashid A, Li S, Hamilton S. Histopathological identification of colon cancer with microsatellite instability. *Am J Pathol* 2001;158:527-535.

Ali H, Olatubosun A, Vihinen M. Classification of mismatch repair gene missense variants with PON-MMR. *Hum Mutat*. 2012 Apr;33(4):642-50.

Andrieu N, Launoy G, Guillois R, Ory-Paoletti C, Gignoux M. Familial relative risk of colorectal cancer: a population-based study. *Eur J Cancer* 2003;39:1904-1911.

Auclair J, Busine MP, Navarro C, Ruano E, Montmain G, Desseigne F, et al. Systematic mRNA analysis for the effect of *MLH1* and *MSH2* missense and silent mutations on aberrant splicing. *Hum Mutat* 2006;27:145–154.

B

Baglietto L, Lindor NM, Dowty JG, White DM, Wagner A, Gomez Garcia EB, et al. Risks of Lynch syndrome cancers for *MSH6* mutation carriers. J Natl Cancer Inst. 2010 Feb 3;102(3):193-201.

Bairoch, A. & Boeckmann, B. The SWISS-PROT protein sequence data bank. Nucleic Acids Res. 1991;19, 2247–2249.

Barnetson RA, Cartwright N, van Vliet A, Haq N, Drew K, Farrington S, et al. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat. 2008 Mar;29(3):367-74.

Barraya R, Benhamiche AM, Rassiat E, Phelip JM, Jouve JL, Faivre J. Incidence et modalités de prise en charge des cancers de l'intestin grêle en Bourgogne. Gastroenterol Clin Biol 1999;23:215-220.

Berends MJ, Wu Y, Sijmons RH, Mensink RG, van der Sluis T, Hordijk-Hos JM, et al. Molecular and clinical characteristics of *MSH6* variants: an analysis of 25 index carriers of a germline variant. Am J Hum Genet. 2002 Jan;70(1):26-37.

Bérout C, Collod-Bérout G, Boileau C, Soussi T, Junien C. UMD (Universal mutation database): a generic software to build and analyze locus-specific databases. Hum Mutat. 2000;15(1):86-94.

Bérout C, Joly D, Gallou C, Staroz F, Orfanelli MT, Junien C. Software and database for the analysis of mutations in the VHL gene. Nucleic Acids Res. 1998 Jan 1;26(1):256-8.

Bérout C, Soussi T. p53 and APC gene mutations: software and databases. Nucleic Acids Res. 1997 Jan 1;25(1):138.

Bilofsky HS, Burks C, Fickett JW, Goad WB, Lewitter FI, Rindone WP, Swindell CD, Tung CS. The GenBank genetic sequence databank. Nucleic Acids Res. 1986 Jan 10;14(1):1-4.

Boks D, Trujillo A, Voogd A, Morreau H, Kenter G, Vasen H. Survival analysis of endometrial carcinoma associated with hereditary nonpolyposis colorectal cancer. Int J Cancer 2002;102:198-200.

Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, 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 Nov 15;58(22):5248-57.

Boland CR, Troncale FJ. Familial colonic cancer without antecedent polyposis. *Ann Intern Med* 1984;100(5):700–701.

Boland CR. Cancer family syndrome. A case report and literature review. *Am J Dig Dis.* 1978; 23(5):25s–27s .18.

Bonadona V, Bonaïti B, Olschwang S, Grandjouan S, Huiart L, Longy M, et al. Cancer risks associated with germline mutations in *MLH1*, *MSH2*, and *MSH6* genes in Lynch syndrome. *JAMA.* 2011 Jun 8;305(22):2304-10.

Brieger A, Plotz G, Raedle J, Weber N, Baum W, Caspary WF, et al. Characterization of the nuclear import of human MutL alpha. *Mol Carcinog.* 2005;43:51–58.

Brieger A, Trojan J, Raedle J, Plotz G, Zeuzem S. Transient mismatch repair gene transfection for functional analysis of genetic *hMLH1* and *hMSH2* variants. *Gut* 2002;51:677–684.

Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK. Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature* 1994;368:258-261.

C

Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet.* 2002 Apr;3(4):285-98.

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* 2003 Jul 1;31(13):3568- 1.

Chapusot C, Martin L, Mungra N, Rageot D, Bouvier AM, Bonithon Kopp C, et al. Sporadic colorectal cancers with defective mismatch repair display a number of specific morphological characteristics: relationship between the expression of *hMLH1* and *hMSH2* proteins and clinicopathological features of 273 adenocarcinomas. *Histopathology.* 2003;43:40-47.

Charbonnier F, Olschwang S, Wang Q, Boisson C, Martin C, Buisine MP, et al. *MSH2* in contrast to *MLH1* and *MSH6* is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer. *Cancer Res* 2002;62:848-853.

Chen S, Wang W, Lee S, Nafa K, Lee J, Romans K, et al. Prediction of germline mutations and cancer risk in the Lynch syndrome. *JAMA*. 2006;296:1479–87.

Christensen M, Katballe N, Wikman F, Primdahl H, Sorensen FB, Laurberg S, et al. Antibody-based screening for hereditary nonpolyposis colorectal carcinoma compared with microsatellite analysis and sequencing. *Cancer* 2002;95:2422-2430

Clarke AB, Cook ME, Tran HT, Gordenin DA, Resnick MA, Kunkel TA. Functional analysis of human MutS alpha and MutS beta complexes in yeast. *Nucleic Acids Res*. 1999;27:736–742.

Claustres M, Horaitis O, Vanevski M, Cotton RG. Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases. *Genome Res*. 2002 May;12(5):680-8.

Collod G, Bérout C, Soussi T, Junien C, Boileau C. Software and database for the analysis of mutations in the human FBN1 gene. *Nucleic Acids Res*. 1996 Jan 1;24(1):137-40.

Cyr JL, Brown GD, Stroop J, Heinen CD. The predicted truncation from a cancer-associated variant of the *MSH2* initiation codon alters activity of the *MSH2-MSH6* mismatch repair complex. *Mol Carcinog*. 2012 Aug;51(8):647-58.

D

de Jong AE, van Puijenbroek M, Hendriks Y, Tops C, Wijnen J, Ausems MGEM, et al. Microsatellite instability, immunohistochemistry, and additional PMS2 staining in suspected hereditary nonpolyposis colorectal cancer. *Clin Cancer Res* 2004;10:972-980.

de Leeuw WJ, Dierssen J, Vasen HF, Wijnen JT, Kenter GG, Meijers-Heijboer H, et al. Prediction of a mismatch repair gene defect by microsatellite instability and immunohistochemical analysis in endometrial tumours from HNPCC patients. *J Pathol* 2000;192: 328-335.

de Rosa M. Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the PMS2 gene. *Oncogene* 2000;19:1719-1723.

Desmet FO, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009 May;37(9):e67.

Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, Rüschoff J. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. *Cancer Res.* 1997 Nov 1;57(21):4749-56.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, de Wind N. A rapid and cell-free assay to test the activity of lynch syndrome-associated *MSH2* and *MSH6* missense variants. *Hum Mutat.* 2012 Mar;33(3):488-94.

Drotschmann K, Clark AB, Kunkel TA. Mutator phenotypes of common polymorphisms and missense mutations in *MSH2*. *Curr Biol* 1999; 9:907–910.

Drotschmann K, Clark AB, Tran HT, Resnick MA, Gordenin DA, Kunkel TA. Mutator phenotypes of yeast strains heterozygous for mutations in the *MSH2* gene. *Proc Natl Acad Sci USA* 1999; 96:2970–2975.

Dunlop MG, Farrington SM, Carothers AD, Wyllie AH, Sharp L, Burn J, Liu B, et al. Cancer risk associated with germline DNA mismatch repair gene mutations. *Hum Mol Genet* 1997;6:105-110.

E

Ellison AR, Lofing J, Bitter GA. Functional analysis of human *MLH1* and *MSH2* missense variants and hybrid human-yeast *MLH1* proteins in *Saccharomyces cerevisiae*. *Hum Mol Genet.* 2001 Sep 1;10(18):1889-900.

F

Fairbrother WG, Yeh RF, Sharp PA, Burge CB. Predictive identification of exonic splicing enhancers in human genes. *Science.* 2002 Aug 9;297(5583):1007-13.

Fairbrother WG, Yeo GW, Yeh R, Goldstein P, Mawson M, Sharp PA, Burge CB. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* 2004 Jul 1;32

Fan Y, Wang W, Zhu M, Zhou J, Peng J, Xu L, et al. Analysis of *hMLH1* missense mutations in East Asian patients with suspected hereditary nonpolyposis colorectal cancer. *Clin Cancer Res.* 2007 Dec 15;13(24):7515-21.

Frayling IM. Microsatellite instability. *Gut.* 1999 Jul;45(1):1-4.

Frédéric MY, Lalande M, Boileau C, Hamroun D, Claustres M, Bérout C, et al. UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity -- application to four genes: *FBN1*, *FBN2*, *TGFBR1*, and *TGFBR2*. *Hum Mutat.* 2009 Jun;30(6):952-9.

G

Genschel J. Human exonuclease I is required for 5' and 3' mismatch repair. *J Biol Chem* 2002;277:13302-13311.

Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS; IARC Unclassified Genetic Variants Working Group. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Hum Mutat.* 2008 Nov;29(11):1265-72.

Gorlov IP, Gorlova OY, Frazier ML, Amos CI. Missense mutations in *hMLH1* and *hMSH2* are associated with exonic splicing enhancers. *Am J Hum Genet.* 2003 Nov;73(5):1157-61.

Guerrette S, Wilson T, Gradia S, Fishel R. Interactions of human *hMSH2* with *hMSH3* and *hMSH2* with *hMSH6*: examination of mutations found in hereditary nonpolyposis colorectal cancer. *Mol Cell Biol* 1998;18:6616–6623.

H

Hamilton SR, Liu B, Parsons RE, Papadopoulos N, Jen J, Powell SM, et al. The molecular basis of Turcot's syndrome. *N Engl J Med*, 1995;332:839-847.

Hamm GH, Cameron GN. The EMBL data library. Nucleic Acids Res. 1986 Jan 10;14(1):5-9.

Heinen CD, Wilson T, Mazurek A, Berardini M, Butz C, Fishel R. HNPCC mutations in *hMSH2* result in reduced *hMSH2-hMSH6* molecular switch functions. Cancer Cell 2002;1:469-478.

Hemminki K, Li X. Familial colorectal adenocarcinoma and HNPCC: a nationwide epidemiological study from Sweden. Br J Cancer 2001;84:969-974.

Hendriks Y, Franken P, Dierssen JW, de Leeuw W, Wijnen J, Dreef E, et al. Conventional and tissue microarray immunohistochemical expression analysis of mismatch repair in hereditary colorectal tumors. Am J Pathol 2003;162:469-477.

Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 1992 Nov 15;89(22):10915-9.

Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812-4. 34.

Holmes J Jr, Clark S, Modrich P. Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. Proc Natl Acad Sci USA 1990;87:5837-5841.

Horaitis O, Cotton RG. The challenge of documenting mutation across the genome: the human genome variation society approach. Hum Mutat. 2004 May;23(5):447-52.

Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, et al. Biallelic inactivation of *BRCA2* in Fanconi anemia. Science 2002 ;297:606-609

J

Jagmohan-Changur S, Poikonen T, Vilkkilä S, Launonen V, Wikman F, Orntoft TF, et al. *EXO1* variants occur commonly in normal population: evidence against a role in hereditary nonpolyposis colorectal cancer. Cancer Res. 2003 Jan 1;63(1):154-8.

Jass J, Walsh M, Barker M, Simms L, Young J, Leggett B. Distinction between familial and sporadic forms of colorectal cancer showing DNA microsatellite instability. *Eur J Cancer* 2002;38:858-866.

Jeanpierre, C., Bérout, C., Niaudet, P. & Junien, C. Software and database for the analysis of mutations in the human WT1 gene. *Nucleic Acids Res* 1998; 26, 271–274.

Jenkins MA, Baglietto L, Dite GS, Jolley DJ, Southey MC, Whitty J, et al. After *MSH2* and *MLH1*, what next ? Analysis of 3-generational, population-based, early-onset colorectal cancer families. *Int J Cancer* 2002;102:166-171.

Jourdan F, Sebbagh N, Comperat E, Mourra N, Flahault A, Olschwang S, et al. Tissue microarray technology: validation in colorectal carcinoma and analysis of p53, h*MLH1*, and h*MSH2* immunohistochemical expression. *Virchows Arch.* 2003 ; 443:115-121.

K

Katballe N, Christensen M, Wilkman FP, Orntoft TF, Laurberg S. Frequency of HNPCC in Danish colorectal cancer patients. *Gut* 2002;50:43-51.

Kempers MJ, Kuiper RP, Ockeloen CW, et al. Risk of colorectal and endometrial cancers in EPCAM deletion-positive Lynch syndrome: a cohort study. *Lancet Oncol* 2011; 12: 49-55.

Kievit W, de Bruin JH, Adang EM, Ligtenberg MJ, Nagengast FM, van Krieken JH, Hoogerbrugge N. Current clinical selection strategies for identification of hereditary non-polyposis colorectal cancer families are inadequate: a meta-analysis. *Clin Genet.* 2004 Apr;65(4):308-16.

Knudsen NØ, Nielsen FC, Vinther L, Bertelsen R, Holten-Andersen S, Liberti SE, et al. 2007. Nuclear localization of human DNA mismatch repair protein exonuclease 1 (hEXO1). *Nucleic Acids Res* 35:2609–2619.

Kohlmann W, Gruber SB. Lynch Syndrome. 2004 Feb 05 [updated 2012 Sep 20]. In: Pagon RA, Adam MP, Bird TD, Dolan CR, Fong CT, Stephens K, editors. *GeneReviews™*

Kondo E, Suzuki H, Horii A, Fukushima S. A yeast two-hybrid assay provides a simple way to evaluate the vast majority of h*MLH1* germ-line mutations. *Cancer Res* 2003;63:3302–3308.

L

Laiho P, Launonen V, Lahermo P, Esteller M, Guo M, Herman JG, et al. Low-level microsatellite instability in most colorectal carcinomas. *Cancer Res* 2002;62:1166-70.

Lastella P, Surdo NC, Resta N, Guanti G, Stella A. In silico and in vivo splicing analysis of *MLH1* and *MSH2* missense mutations shows exon- and tissue-specific effects. *BMC Genomics*. 2006 Sep 22;7:243.

Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 1993;75:1215-1225.

Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, et al. Heritable somatic methylation and inactivation of *MSH2* in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat Genet*. 2009 Jan;41(1):112-7.

Liljegren A, Barker G, Elliott F, Bertario L, Bisgaard ML, Eccles D, et al. Prevalence of adenomas and hyperplastic polyps in mismatch repair mutation carriers among CAPP2 participants: report by the colorectal adenoma/carcinoma prevention programme 2. *J Clin Oncol*. 2008 Jul 10;26(20):3434-9

Lin KM, Shashidharan M, Thorson AG, Ternent CA, Blatchford GJ, Christensen MA, et al. Cumulative incidence of colorectal and extracolonic cancers in *MLH1* and *MSH2* mutation carriers of hereditary nonpolyposis colorectal cancer. *J Gastrointest Surg*. 1998 Jan-Feb;2(1):67-71.

Lindblom A, Tannergard P, Werelius B, Nordenskjold M. Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nat Genet* 1993;5:279-282.

Lipkin SM, Wang V, Stoler DL, Anderson GR, Kirsch I, Hadley D, Lynch HT, Collins FS. Germline and somatic mutation analyses in the DNA mismatch repair gene *MLH3*: Evidence for somatic mutation in colorectal cancers. *Hum Mutat*. 2001 May;17(5):389-96.

Liu CY, Flesken-Nikitin A, Li S, Zeng Y, Lee WH. Inactivation of the mouse Brca1 gene leads to failure in the morphogenesis of the egg cylinder in early postimplantation development. *Genes Dev* 1996;10:1835–1843

Liu HX. The role of hMLH3 in familial colorectal cancer. *Cancer Res* 2003;63:1894-1899.

Liu T. The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. *Cancer Res* 2001;61:7798-7802.

Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res*. 1991 Jun 15;51(12):3075-9.

Lu SL, Kawabata M, Imamura T, Akiyama Y, Nomizu T, Miyazono K, Yuasa Y. HNPCC associated with germline mutation in the TGF-beta type II receptor gene. *Nat Genet*. 1998 May;19(1):17-8.

Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med*. 2003;348(10):919-932.

Lynch HT, Kimberling W, Albano WA et al (1985) Hereditary nonpolyposis colorectal cancer (Lynch syndromes I and II). I. Clinical description of resource. *Cancer* 56(4):934–938. 21.

Lynch HT, Riegert-Johnson DL, Snyder C, Lynch JF, Hagenkord J, Boland CR, et al. Lynch syndrome-associated extracolonic tumors are rare in two extended families with the same EPCAM deletion. *Am J Gastroenterol*. 2011 Oct;106(10):1829-36.

Lynch HT, Schuelke GS, Kimberling WJ et al. Hereditary nonpolyposis colorectal cancer (Lynch syndromes I and II). II. Biomarker studies. *Cancer* 1985;56(4):939–951

M

Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, Lutterbaugh J, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science*. 1995 Jun 2;268(5215):1336-8.

Mecklin JP, Järvinen HJ. Tumor spectrum in cancer family syndrome (hereditary nonpolyposis colorectal cancer). *Cancer*. 1991 Sep 1;68(5):1109-12.

Mecklin JP, Jarvinen JHJ, Virolainen M. The association between cholangiocarcinoma and hereditary non-polyposis colorectal carcinoma. *Cancer*, 1992; 69:1112-1114.

Miyaki M, Konishi M, Tanaka K et al. Germline mutation of *MSH6* as the cause of hereditary nonpolyposis colorectal cancer. *Nat Genet* 1997;17(3):271–272.

N

Nagakawa H, Nuovo G, Zervos E, Martin E, Salovaara R, Aaltonen L et de la Chapelle A. Age-related hypermethylation of the 5' region of *MLH1* in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. *Cancer Res* 2001;61:6991-6995.

Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature*. 1994 Sep 1;371(6492):75-80.

Nystrom-Lahti M, Perrera C, Raschle M, Panyushkina-Seiler E, Marra G, Curci A, et al. Functional analysis of *MLH1* mutations linked to hereditary nonpolyposis colon cancer. *Genes Chromosomes Cancer* 2002;33:160–167.

O

Ou J, Niessen RC, Vonk J, Westers H, Hofstra RM, Sijmons RH. A database to support the interpretation of human mismatch repair gene variants. *Hum Mutat*. 2008 Nov;29(11):1337-41.

P

Pagenstecher C, Wehner M, Friedl W, Rahner N, Aretz S, Friedrichs N, et al. Aberrant splicing in *MLH1* and *MSH2* due to exonic and intronic variants. *Hum Genet*. 2006 Mar;119(1-2):9-22.

Pagon RA, Adam MP, Bird TD, Dolan CR, Fong CT, Stephens K, editors. *GeneReviews*™

Pang Q et coll. Functional domains of the *Saccharomyces cerevisiae* *MLH1p* and *Pms1p* DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations. *Mol Cell Biol* 1997;17:4465-4473.

Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, Rosen CA, et al. Mutation of a *mutL* homolog in hereditary colon cancer. *Science* 1994;263:1625-1629.

Parc Y, Boisson C, Thomas G, Olschwang S. Cancer risk in 348 French *MSH2* or *MLH1* gene carriers. *J Med Genet* 2003;40:208-213.

Parc YR, Halling KC, Burgart LJ, McDonnell SK, Schaid DJ, Thibodeau SN, et al. Microsatellite instability and *hMLH1/hMSH2* expression in young endometrial carcinoma patients: associations with family history and histopathology. *Int J Cancer* 2000;86:60-66.

Pearson, P. L. The genome data base (GDB) — human gene mapping repository. *Nucleic Acids Res* 1991; 19 (Suppl.), 2237–2239.

Peel DJ, Ziogas A, Fox EA, Gildea M, Laham B, Clements E, et al. Characterization of hereditary nonpolyposis colorectal cancer families from a population-based series of cases. *J Natl Cancer Inst* 2000;92:1517-1522.

Peinado MA, Malkhosyan S, Velazquez A, Perucho M. Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. *Proc Natl Acad Sci USA* 1992;89(21):10065–10069).

Peltomäki P, Aaltonen LA, Sistonen P et al. Genetic mapping of a locus predisposing to human colorectal cancer. *Science* 1993;260(5109):810–812)

Peltomäki, P, and Vasen, H, The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer: Mutations predisposing to hereditary nonpolyposis colorectal cancer : database and results of a collaborative study. *Gastroenterology*, 1997;113:1146-1158.

Percesepe A, Benatti P, Roncucci L, Sassatelli R, Fante R, Ganazzi D, et al. Survival analysis in families affected by hereditary non-polyposis colorectal cancer. *Int J Cancer* 1997;71:373-376.

Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat.* 2008 Nov;29(11):1282-91.

Q

Quaresima B, Alifano P, Tassone P, Avvedimento EV, Costanzo FS, Venuta S. Human mismatch-repair protein MutL homologue 1 (*MLH1*) interacts with *Escherichia coli* MutL and MutS in vivo and in vitro: a simple genetic system to assay *MLH1* function. *Biochem J.* 2003 Apr 1;371(Pt 1):183-9.

R

Raedle J, Trojan J, Brieger A, Weber N, Schäfer D, Plotz G. Bethesda guidelines : relations to microsatellite instability and *MLH1* promoter methylation in patients with colorectal cancer. *Ann Intern Med* 2001;135:566-576.

Raevaara TE, Korhonen MK, Lohi H, Hampel H, Lynch E, Lonnqvist KE, et al. Functional significance and clinical phenotype of nontruncating mismatch repair variants of *MLH1*. *Gastroenterology* 2005;129:537–549.

Rahman N, Scott RH. Cancer genes associated with phenotypes in monoallelic and biallelic mutation carriers: new lessons from old players. *Hum Mol Genet* 2007;16:R60–R66

Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* 2002 Sep 1;30(17):3894-900.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, de Wind N. Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. *Hum Mutat.* 2012 Dec;33(12):1617-25.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, de Wind N. Response to: Design of a core classification process for DNA mismatch repair variations of a priori unknown functional significance. *Hum Mutat.* 2013 Jun;34(6):923-4.

Riccio A, Aaltonen LA, Godwin AK, Loukola A, Percesepe A, Salovaara R, et al. The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. *Nat Genet.* 1999 Nov;23(3):266-8.

Riegler G, Savastano A, Selvaggi F, Ciocano R, Martino R, Riccio G, et al. Prevalence of HNPCC in a series of consecutive patients on the first endoscopic diagnosis of colorectal cancer: a multicenter study. *Endoscopy* 1999;31:337-341.

Rigau V, Sebbagh N, Olschwang S, Paraf F, Mourra N, Parc Y. Microsatellite instability in colorectal carcinoma. The comparison of immunohistochemistry and molecular biology suggests a role for hMSH6 [correction of hMLH6] immunostaining. *Arch Pathol Lab Med* 2003;127:694-700.

Rijcken FE, Hollema H, Kleibeuker JH. Proximal adenomas in hereditary non-polyposis colorectal cancer are prone to rapid malignant transformation. *Gut.* 2002 Mar;50(3):382-6

Rodriguez-Bigas MA, Boland CR, Hamilton SR, Henson DE, Jass JR, Khan PM, et al. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J Natl Cancer Inst.* 1997 Dec 3;89(23):1758-62.

Rodriguez-Bigas MA, Vasen HFA, Lynch HT, Watson P, Myrhoj T, Jarvinen HJ, et al. Characteristics of small bowel carcinoma in hereditary non-polyposis colorectal carcinoma. *Cancer* 1998; 83:240-244.

Rovella V, Carrara S, Crucitti SC, Coco C, Magistrelli P, Lucci-Cordisco E, et al. Familial MSS non-polyposis colorectal cancer : incidence and characteristics in a clinic-based population. *Ann Oncol* 2001;12:813-818.

S

Salovaara R, Loukola A, Kristo P, Kaariainen H, Ahtola H, Eskelinen M, et al. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 2000;18:2193-2200.

Sankila R, Aaltonen LA, Järvinen HJ, Mecklin JP. Better survival rates in patients with MLH1-associated hereditary colorectal cancer. *Gastroenterology* 1996 Mar;110(3):682-7.

Saurin JC, Pilleul F, Soussan EB, Manière T, D'Halluin PN, Gaudric M, et al. Small-bowel capsule endoscopy diagnoses early and advanced neoplasms in asymptomatic patients with Lynch syndrome. *Endoscopy*. 2010 Dec;42(12):1057-62.

Senter L, Clendenning M, Sotamaa K, Hampel H, Green J, Potter JD, et al. The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations. *Gastroenterology*. 2008 Aug;135(2):419-28.

Shimodaira H, Filosi N, Shibata H, Suzuki T, Radice P, Kanamaru R, et al. Functional analysis of human *MLH1* mutations in *Saccharomyces cerevisiae*. *Nat Genet* 1998;19:384–389.

Shin JH, Ku JL, Shin YK et al. GST-M1 associated with cancer occurrence in Korean HNPCC families carrying the *MSH2/MLH1* mutation. *Oncol Rep* 2003;10:483-486.

Sijmons RH, Kiemeny LA, Witjes JA, Vasen HF. Urinary tract cancer and hereditary non-polyposis colorectal cancer: Risks and screening options. *J. Urol.* 1998; 160:466-470.

Soussi T, Ishioka C, Claustres M, Bérout C. Locus-specific mutation databases: pitfalls and good practice based on the p53 experience. *Nat Rev Cancer*. 2006 Jan;6(1):83-90.

Studamire B et coll. Separation-of-function mutations in *Saccharomyces cerevisiae MSH2* that confer mismatch repair defects but do not affect nonhomologous-tail removal during recombination. *Mol Cell Biol* 1999;19:7558-7567.

Sun X, Zheng L, Shen B. Functional alterations of human exonuclease 1 mutants identified in atypical hereditary nonpolyposis colorectal cancer syndrome. *Cancer Res*. 2002 Nov 1;62(21):6026-30.

Takahashi M, Shimodaira H, Andreutti-Zaugg C, Iggo R, Kolodner RD, Ishioka C. Functional analysis of human *MLH1* variants using yeast and in vitro mismatch repair assays. *Cancer Res*. 2007 May 15;67(10):4595-604

Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, Seruca R, Iacopetta B, Hamelin R. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 2002;123:1804-1811.

T

Teraoka SN, Telatar M, Becker-Catania S, Liang T, Onengüt S, Tolun A, et al. Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *Am J Hum Genet.* 1999 Jun;64(6):1617-31.

Terdiman JP, Levin TR, Allen BA, Gum JR Jr, Fishbach A, Conrad PG, et al. Hereditary nonpolyposis colorectal cancer in young colorectal cancer patients: high-risk clinic versus population-based registry. *Gastroenterology* 2002;122:940-947.

Thompson D, Easton DF, Goldgar DE. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet.* 2003 Sep;73(3):652-5.

Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, et al. A large fraction of unclassified variants of the mismatch repair genes *MLH1* and *MSH2* is associated with splicing defects. *Hum Mutat.* 2008 Dec;29(12):1412-24.

Trojan J, Zeuzem S, Randolph A, Hemmerle C, Brieger A, Raedle J, et al. Functional analysis of *hMLH1* variants and HNPCC-related mutations using a human expression system. *Gastroenterology* 2002;122:211-219.

U

Umar A, Boland CR, Terdiman JP et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 2004;96(4):261-268

V

Vasen H, Wijnen JT, Menko FH. Cancer risk in families of patients with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology* 1996;110:1020-1027.

Vasen HF, Mecklin JP, Khan PM, Lynch HT. The international collaborative group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). *Dis Colon Rectum* 1991;34(5):424-425. 25.

Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the international collaborative group on HNPCC. *Gastroenterology* 1999;116(6):1453-1456).

Viel A. Lack of PMS2 gene-truncating mutations in patients with hereditary colorectal cancer. *Int J Oncol* 1998;13:565-569.

W

Wahlberg SS, Schmeits J, Thomas G, Loda M, Garber J, Syngal S, et al. Evaluation of microsatellite instability and immunohistochemistry for the prediction of germ-line *MSH2* and *MLH1* mutations in hereditary nonpolyposis colon cancer families. *Cancer Res* 2002;62:3485-3492.

Wang Q, Lasset C, Desseigne F, Frappaz D, Bergeron C, Navarro C, et al. 1999. Neurofibromatosis and early onset of cancers in h*MLH1*-deficient children. *Cancer Res* 59: 294-297

Warthin AS Heredity with reference to carcinoma as shown by the study of the cases examined in the Pathological Laboratory of the University of Michigan, 1895-1912. *Arch Int Med* 1913;12:546-555

Warthin AS. The further study of a cancer family. *J Cancer Res* 1925;9:279-286

Watson P and Lynch H. Extracolonic cancer in hereditary nonpolyposis colorectal cancer. *Cancer* 1993;71:677-685.

Watson P, Butzow R, Lynch H, Mecklin J, Jarvinen H, Vasen H, et al. The clinical features of ovarian cancer in hereditary nonpolyposis colorectal cancer. *Gynecol Oncol* 2001; 82:223-228.

Wautot, V. et al. Germline mutation profile of MEN1 in multiple endocrine neoplasia type 1: search for correlation between phenotype and the functional domains of the MEN1 protein. *Hum. Mutat.* 2002;20, 35–47.

Wijnen JT, Vasen HF, Khan PM, Zwinderman AH, van der Klift H, Mulder A, et al. Clinical findings with implications for genetic testing in families with clustering of colorectal cancer. *N Engl J Med* 1998;20:511-518.

Win AK, Young JP, Lindor NM, Tucker KM, Ahnen DJ, Young GP, et al. Colorectal and other cancer risks for carriers and noncarriers from families with a DNA mismatch repair gene mutation: a prospective cohort study. *J Clin Oncol.* 2012 Mar 20;30(9):958-64.

Wu Y, Berends MJ, Post JG, Mensink RG, Verlind E, Van Der Sluis T, Kempinga C, et al. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. *Gastroenterology.* 2001 Jun;120(7):1580-7.

Y

Young J, Simms L, Biden K, Wynter C, Whitehall V, Karamatic R, et al. Features of colorectal cancers with MSI-H occurring in familial and sporadic settings. *Am J Pathol* 2001;159:2107-2116. Yuan ZQ. Polymorphisms and HNPCC: PMS2-*MLH1* protein interactions diminished by single nucleotide polymorphisms. *Hum Mut* 2002;19:108-112.

ANNEXES

Annexe 1: critères de Bethesda modifiés (Umar et al., 2004)

Table 1. Recommendations for the evaluation of MSI-H and MSI-L*

The original National Cancer Institute (NCI) microsatellite panel included BAT25, BAT26, D2S123, D5S346 and D17S250 (7); however, the following caveats may apply:

1. If only dinucleotide repeats are mutated, test a secondary panel of microsatellite markers with mononucleotide repeats (e.g., BAT40 and/or MYCL) to exclude MSI-L.
2. Dinucleotide repeats are less sensitive than mononucleotide repeats for MSI-H; however, they provide an internal control for the prevention of sample mix-up.
3. A pentaplex panel of five quasimonomorphic mononucleotide repeats may be more sensitive for MSI-H tumors than other microsatellite markers and may obviate the need for normal tissue for comparison; this approach requires three or more mutant alleles to indicate MSI-H (25).

Table 2. The Revised Bethesda Guidelines for testing colorectal tumors for microsatellite instability (MSI)

Tumors from individuals should be tested for MSI in the following situations:

1. Colorectal cancer diagnosed in a patient who is less than 50 years of age.
2. Presence of synchronous, metachronous colorectal, or other HNPCC-associated tumors,* regardless of age.
3. Colorectal cancer with the MSI-H[†] histology[‡] diagnosed in a patient who is less than 60 years of age.§
4. Colorectal cancer diagnosed in one or more first-degree relatives with an HNPCC-related tumor, with one of the cancers being diagnosed under age 50 years.
5. Colorectal cancer diagnosed in two or more first- or second-degree relatives with HNPCC-related tumors, regardless of age.

Table 3. Recommendations for the process of molecular evaluation of patients identified as being at risk, based on meeting the Bethesda Guidelines*

Process of molecular evaluation:

1. The optimal approach to evaluation is microsatellite instability (MSI) or immunohistochemical (IHC) analysis[†] of tumors, followed by germline MSH2/MLH1 testing in patients with MSI-H tumors or tumors with a loss of expression of one of the mismatch repair genes.[‡]
2. After the mutation is identified, at-risk relatives should be referred for genetic counseling and tested if they wish.
3. An alternative approach, if tissue testing is not feasible, is to proceed directly to germline analysis of the MSH2/MLH1 genes.
4. If no mismatch repair gene mutation is found in a proband with an MSI-H tumor and/or a clinical history of hereditary nonpolyposis colorectal cancer (HNPCC), the genetic test result is non-informative. The patients and the at-risk individuals (i.e., relatives) should be counseled as if HNPCC was confirmed and high-risk surveillance should be undertaken.
5. There is a need to assure patients of confidentiality to allay fears related to discrimination based on genetic status.

Annexe 2 : Système MMR

MISMATCH REPAIR SYSTEM

Normal DNA

5' TGGTCATGCTAGCTGGATGTGTGTGTGTGTGTATCGTACAAAAAAAAAAAA-C
3' ACCAGTACGATCGACCTACACACACACACACATAGCATGTTTTTTTTTTTTTTTG

Replication error

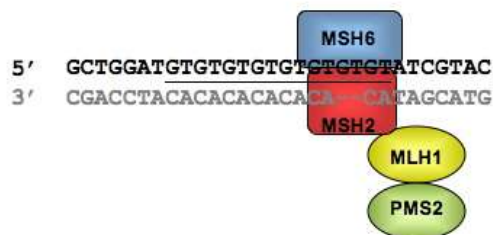
5' TGGTCATGCTAGCTGGATGTGTGTGTGTGTGTATCGTACAAAAAAAAAAAA-C
3' ACCAGTAAGATCGACCTACACACACACACA--TAGCATGTTTTTTTTTTTTTTTG

substitution deletion Insertion

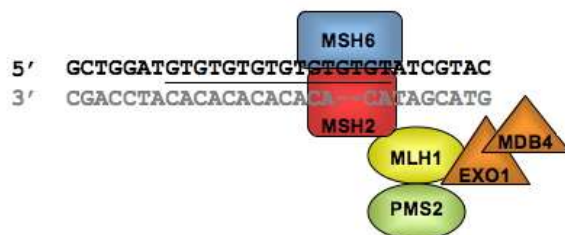
Mismatch recognition by MutSα



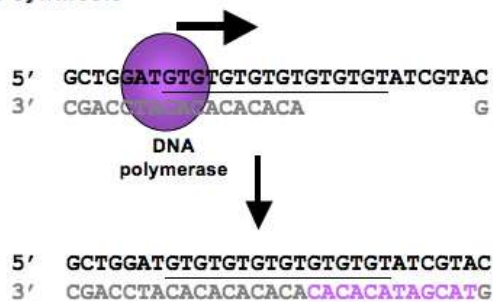
Recruitment of MutL complex



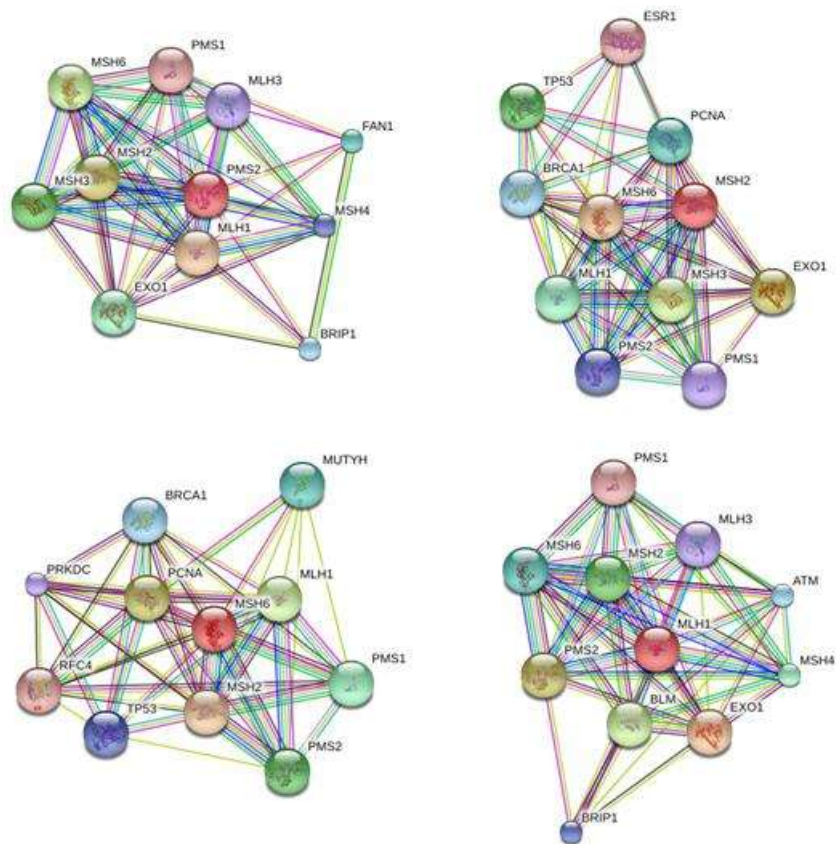
Recruitment of endo- and exonucleases



Excision and synthesis

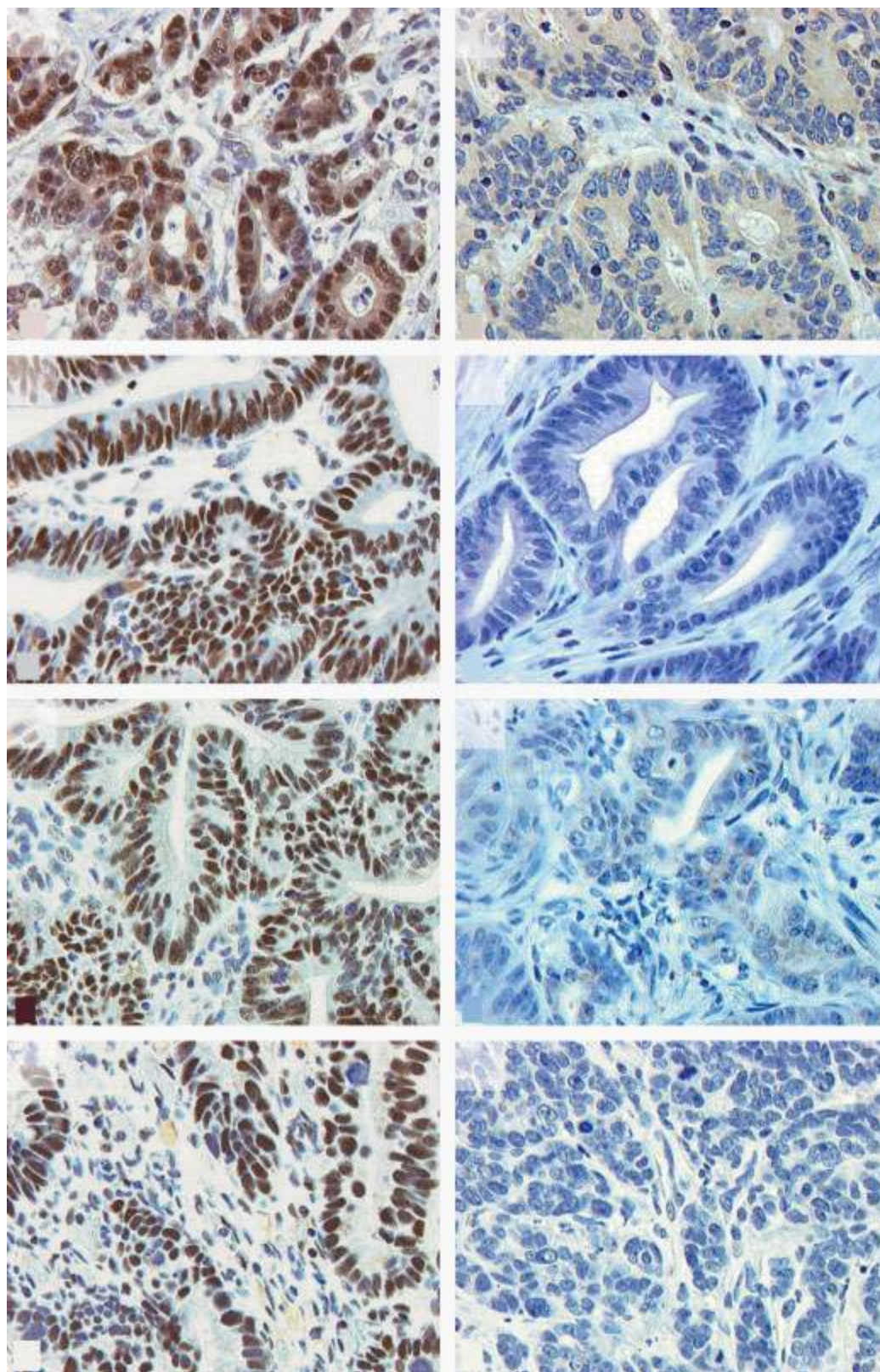


Annexe 3 : interactions entre les protéines du système MMR (<http://string-db.org/>)



Annexe 4: immuno histochimie

De haut en bas et de droite à gauche, expression et extinction de *MLH1*, *MSH2*, *MSH6* et *PMS2*
(D'après HAMPEL H et al. NEJM 2005;352 :1851-60)



Annexe 5 : SIFT

Predictions for positions 1 through 100

Threshold for intolerance is 0.05.

Amino acid color code: nonpolar, **uncharged** **polar**, **basic**, **acidic**.

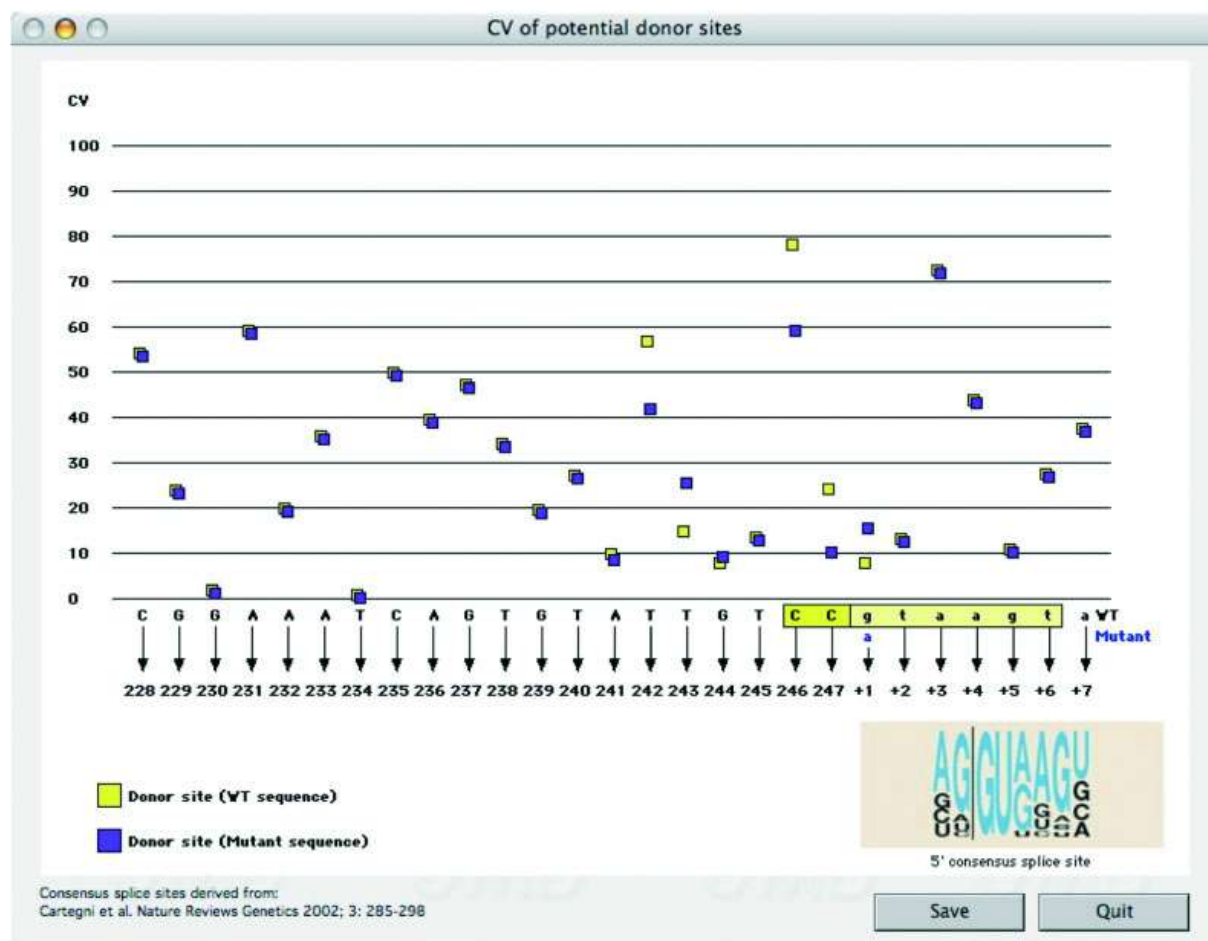
Capital letters indicate amino acids appearing in the alignment, lower case letters result from prediction.

'Seq Rep' is the fraction of sequences that contain one of the basic amino acids. A low fraction indicates the position is either severely gapped or unalignable and has little information. Expect poor prediction at these positions.

Predict Not Tolerated	Position	Seq Rep	Predict Tolerated
y w v t s r q n m l k i h g f e d c a 1P	0.85	P	
c w f m i y v l h r g t n s p a k d 2Q	0.95	E Q	
d h n e c p g q r k s w t a 3I	0.95	m Y V L F I	
w f y h m c r e l i q g k d v 4T	0.95	n a P S T	
y w v t s r q p n m k i h g f e d c a 5L	0.95	L	
h m c q i r e k v n d l p t y g f a 6W	0.95	S W	
c w d f m i y v g p s h n a l t e 7Q	0.95	K Q R	
y w v t s q p n m l k i h g f e d c a 8R	1.00	R	
y w v t s r q n m l k i h g f e d c a 9P	1.00	P	

Annexe 6: score prédictif pour potentiel site donneur d'épissage (UMD Predictor®)

Calcul d'un score consensus permettant de prédire si une séquence donnée, éventuellement mutée, constitue un site potentiel d'épissage (Frederic et al., 2009)



Annexe 7: Guidelines HGVS pour les soumissions de nouvelles mutations à un base LSDB

A. SUBMITTED BY

Family name: _____

Given name: _____

Address (List-Dept.,City, State, Country, Postcode-fields to fill in): _____

Email: _____

Fax: _____

URL of institution: _____

Date (auto recorded): _____

Unpublished (Y/N) OR Citation (e.g. *Nature* 352:77, 1999): _____, Medline No.: _____

B. GENERAL INFORMATION-

Species:

Human

Other _____

Gene/Locus name: _____

HUGO-approved Gene symbol (Link to site explaining what these are): _____

Chromosome Region (_____+ unknown)

Category:

mutation associated with disease

polymorphism not causing disease

don't know

Phenotype (if any): _____ OMIM No. _____

C. DNA-

Reference Sequence Accession No.:

DDBJ/EMBL/ GenBank: _____

Tissue analysed: _____ (blood, muscle, tumour, etc.)

Type of DNA analysed

genomic

mitochondrial

cDNA

Systematic name of variation (to follow nomenclature guide by Antonarakis et al 1988 {link to this}) e.g.

1425G->A : _____

Aliases/Published name: _____

Unique ID provided by database: _____

Base change type:

base substitution

base(s) deletion

base(s) insertion

duplication

unknown

complex

Other list _____

Location

exon

intron

5' UTR

3' UTR

Not within a known transcription unit

Exon/Intron no: _____

1st Affected Nucleotide (or starting nt): _____

Last Affected Nucleotide: _____

No. bases inserted: _____

No. bases deleted: _____

25 or more nucleotides 5' of variation site: _____
 Reference subsequence: _____
 Variant subsequence: _____
 25 or more nucleotides 3' of variation site: _____
 Detection method
 Direct Sequencing
 SSCP
 PTT
 DGGE
 CCM
 EMC
 DHPLC
 Dideoxy Fingerprinting
 FISH
 Heteroduplex analysis
 other _____
 Detection conditions (List-sequence of primers, PCR and electrophoresis conditions, etc.): _____
 Extent of DNA analyzed: (% coding seq., codon 2, 5'UTR, etc.): _____
 D. RNA EFFECT
 RNA nucleotide change is identical to DNA mutation Yes/No if No then
 Sequence range affected _____ to _____

 Reference mRNA _____, or No Reference _____
 1st Affected Nucleotide (or starting nt): _____
 Last Affected Nucleotide: _____
 Original codon: _____
 New codon: _____
 Effect on RNA
 promoter activity
 splicing change
 RNA stability
 poly A addition
 no effect
 unknown
 E. PROTEIN EFFECT
 Reference protein _____ or No reference _____
NAME:
 Trivial (protein) name of mutation (to follow nomenclature guide by Antonarakis et al 1988 {link to this})
 e.g. P405G: _____
Effect on Protein (if any)
 missense (aa sub)
 nonsense (stop codon)
 truncating del
 in-frame del
 truncating ins
 in-frame ins
 no effect
 unknown
 two changes
 Wild type amino acid(s): _____
 Mutated amino acid(s) (including new extension seq(s)): _____
 1st Residue No. affected: _____
 Last Residue No. affected: _____
 Position of any new termination codon: _____
 Activity change in protein measured Y/N, Result _____
 F. QUALITY ASSURANCE CHECKLIST FOR MUTATION/POLYMORPHISM
 (LIST-Yes-Y/No-N/Not checked-NC/Not applicable)
 Mutation/polymorphism found on repeat PCR sample (not an artifact) Y/N/NC/NA: _____
 Mutation segregates with trait Y/N/NC/NA: _____

Trait : (____ + unknown + NA)

Parents, grandparents, siblings, or other family members checked for carrier status

Y/N/NC/NA: ____ Result if tested

Other mutation found on same allele (in cis) Y/N/NC/NA: ____

Amino acid in human seq. is conserved in ____ out of ____ other species for which seq is known (date)

Expression analysis confirms phenotypic effect Y/N/NC/NA: ____

50 (or other ____) normal chromosomes tested Y/N/NC/NA: ____

Result: e.g. NIL for disease causing & ratio of alleles (____) for harmless polymorphisms

G. INFORMATION IN OTHER DATABASES

Mutation in OMIM (link to this): No/Yes Accession No. ____

Mutation in HGMD (link to this): No/Yes Accession No. ____

Mutation in dbSNP (link to this): No/Yes Accession No. ____

Mutation in GDB (link to this): No/Yes Accession No. ____

Mutation in other database: No/Yes Accession No. ____ Name of database ____

H. DIAGNOSTIC DETECTION

Diagnostic method developed (after mutation found): Y/N ____

Diagnostic method technology (ASO, RFLP etc.): (____ or NA)

Ability to perform tests for others: (list tests- Y/N/NA): ____

I. OTHER ANCILLARY DATA

Haplotype association: (____ + unknown)

Population association: Freq. Mutation/Wild type alleles ____

Geographic Location: (____ + unknown)

Ethnic association: (____ + unknown)

Linguistic association: (____ + unknown)

Religious association: (____ + unknown)

J. PATIENT/SUBJECT AND PHENOTYPE

Coded Patient/Subject Identifier (if applicable): ____

Sex:

unknown

male

female

NA

Phenotype Occurrence

Familial

Sporadic

NA

Phenotype (multiple selection allowed) (List possible phenotypes & Other): ____

Inheritance-

Unknown OR

Somatic OR

Germinal

Autosomal

X-linked

Mitochondrial

Gonadal mosaicism

Genotype at other loci if known & relevant: ____

K. COMMENTS

Add as much descriptive information as possible ____

L. REFERENCES CITED

Annexe 8 : UMD (Universal Mutation Database)

Liste des critères intégrés dans la base UMD pour chaque gène (Beroud et al., 2000)

TABLE 2. Fields Available in 8 LSDBs

	APC	FBN1	LDLR	MEN1	p53	SUR1	VHL	WT1
AA position, nucleotide, sequence search	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WT codon, * WT AA*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mutant codon, mutant AA*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exon*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mutational event*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Type*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CpG*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Py-Py doublet*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Structure*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Insertion*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Highly conserved domain*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polymorphism or mutation*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample name	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Patient	—	—	—	—	Yes	—	—	—
Origin	Yes	Yes	—	Yes	Yes	Yes	Yes	Yes
Pathology	—	Yes	—	—	—	Yes	—	—
Age of onset	—	Yes	—	—	—	Yes	—	—
Mutation number	—	—	—	—	—	Yes	—	—
Tumor	—	—	—	Yes	—	—	—	—
Cancer	Yes	—	—	—	Yes	—	Yes	Yes
Organ	—	—	—	—	Yes	—	—	—
Histology	Yes	—	—	—	Yes	—	Yes	Yes
Stage	Yes	—	—	—	Yes	—	Yes	Yes
LOH	Yes	—	—	—	Yes	—	Yes	Yes
Mutation type*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Smoking	Yes	—	—	Yes	Yes	—	Yes	Yes
Aflatoxin	—	—	—	—	Yes	—	—	—
Diazoxide status	—	—	—	—	—	Yes	—	—
Skin tumor or Internal tumor	—	—	—	Yes	Yes	—	Yes	Yes
Leukemia/Lymphoma or Solid tumor	—	—	—	Yes	Yes	—	Yes	Yes
Reference	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clinical data	Yes	Yes	Yes	Yes	—	Yes	Yes	Yes
Comments	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*These fields are automatically filled in by the software.

Annexe 9 : Listes des laboratoires français impliqués dans le syndrome de Lynch :

1. CLCC Bergonié, Bordeaux
2. CHU Hôtel Dieu, Nantes
3. CHU Brabois, Nancy
4. CHU Georges Pompidou, Paris
5. CHU Pitié-Salpêtrière, Paris
6. CLCC Gustave Roussy, Villejuif
7. CHRU Lille
8. Hôpitaux Civils, Strasbourg
9. CLCC Paoli-Calmettes, Marseille
10. CHU Timone, Marseille
11. CHU Cochin, Paris
12. CHU Arnaud de Villeneuve, Montpellier
13. CLCC René Huguenin, Saint-Cloud
14. Hôpital Charles Nicolle, Rouen
15. CLCC Claudius Regaud, Toulouse
16. CLCC Jean Perrin, Clermont-Ferrand