Génétique du cancer colorectal : polyposes adénomateuses non liées à APC et cancers de survenue précoce

Jérémie Lefèvre

▶ **To cite this version:**

Jérémie Lefèvre. Génétique du cancer colorectal : polyposes adénomateuses non liées à APC et cancers de survenue précoce. Cancer. Université Pierre et Marie Curie - Paris VI, 2012. Français. NNT : 2012PAO66237 . tel-00833304

HAL Id: tel-00833304

https://theses.hal.science/tel-00833304

Submitted on 12 Jun 2013

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

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

THESE DE DOCTORAT DE

L'UNIVERSITE PIERRE ET MARIE CURIE

PARIS VI

Doctorat des Sciences de la Vie

Physiologie et Physiopathologie (ED 394)

Présentée par

Monsieur Jérémie LEFEVRE

Le 5 septembre 2012

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

GENETIQUE DU CANCER COLORECTAL :

Polyposes adénomateuses non liées à APC et cancers de survenue précoce

JURY

Monsieur le Professeur Florent SOUBRIER	**Président**
Monsieur le Professeur Alex KARTHEUSER	**Rapporteur**
Monsieur le Professeur Christophe PENNA	**Rapporteur**
Monsieur le Professeur Yves PANIS	**Examinateur**
Monsieur le Professeur Christophe MARIETTE	**Examinateur**
Monsieur le Docteur Alex DUVAL	**Examinateur**
Monsieur le Professeur Yann PARC	**Directeur**

RESUME

Le cancer colorectal (CCR) est le troisième cancer dans le monde et devenu un véritable enjeu de santé publique. Environ 5% des CCR sont associés à une forme familiale de transmission autosomique dominante : la polypose adénomateuse familiale (PAF) liée à une mutation sur le gène APC et le Syndrome de Lynch (SL) lié à des mutations sur les gène du système MMR notamment MLH1 et MSH2.

En 2002, un nouveau syndrome de transmission récessive a été découvert : la MAP (*MUTYH-Associated Polyposis*) liée à des mutation biallélique sur le gène MUTYH appartenant au système de réparation de l'ADN BER (*Base Excision Repair*).

1. Implication du syndrome MAP dans les polyposes adénomateuses : Parmi 31 patients avec une polypose sans mutation retrouvée sur APC, 6 (20%) présentaient une mutation biallélique sur MUTYH. Des transversions étaient identifiées sur Kras ou APC chez 5 d'entre eux (83%).
2. Interaction entre MUTYH et MLH1 : Une famille présentait une mutation biallélique sur MUTYH (c.1185_1186dup, p.Glu396GlyfsX43) associée à une perte d'expression somatique de MLH1 et un phénotype MSI. Aucune mutation germinale ou méthylation somatique n'expliquait l'inactivation de MHL1. Une transversion somatique sur l'intron 7 +5 (c.588 +5 G>T) sur MLH1 était identifiée, confirmant pour la première fois que des mutations sur MUTYH pouvait donc inactiver APC et la voie Wnt ou MLH1 et le système MMR.
3. Corrélation génotype phénotype dans les polyposes adénomateuses : 384 patients avec une polypose ont été analysés. 17 étaient mutés sur MUTYH, 322 sur APC et 32 n'avait pas d'explication pour leur polypose. Deux groupes de mutation sur APC ont été identifiés. Le nombre de polypes, leur localisation, l'âge de réalisation de la chirurgie colique ne différaient pas selon la mutation sur APC. Les principales différences étaient observées entre les patients mutés sur APC et ceux mutés sur MUTYH. Cette étude infirmait une corrélation nette entre la mutation sur APC et l'expression phénotypique de la polypose.
4. Fréquence de la mutation c.1185_1186dup dans les polyposes liées à MUTYH : Au sein d'un groupe de 36 familles mutées sur MUTYH, 11 avaient une mutation biallélique homozygote c.1185_1186dup. Cette mutation était significativement plus fréquemment observée chez les patients provenant d'Afrique du Nord (79% vs. 5%, $p<0,0001$). La recherche d'un haplotype commun en utilisant 10 microsatellites a identifié un segment de 1,3 cM présent chez tous les patients avec la mutation c.1185_1186dup.

Une autre explication des polyposes ou des CCR de survenue précoce (avant 50 ans) n'appartenant à un syndrome connu est l'implication des variants rares (VR). Ils correspondent à des anomalies génétiques dont la fréquence est comprise entre 0,1 et 1% et qui confèrent une susceptibilité à une pathologie. L'accumulation de plusieurs VR sur des gènes différents pourrait expliquer l'apparition du CCR ou des polypes.

5. Variants rares de la cycline D1 : La comparaison des fréquences alléliques des VR de la cycline D1 fut réalisée entre les cas (112 patients avec une polypose indéterminée et 44 avec un CCR précoce) et 866 témoins sains. Les VR étaient plus fréquemment observés dans le groupe de malades. En combinant les VR, une augmentation du risque était retrouvée pour le sous groupe de patients avec une polypose indéterminée : (OR= 2,2); 95%IC, 1.1–4.4; P=0,03). L'étude in silico du rôle des VR confirmait pour la plupart un effet fonctionnel.
6. Rôle des variants rares : 70 variants provenant de 17 gènes ont été examinés au sein de la même population. 21 était des VR (fréquence <1%) et 4 étaient plus fréquemment observés chez les cas (EXO1-12, MLH1-1, CTNNB1-1 and BRCA2-37, $p<0,05$). En combinant tous les VR avec une fréquence allélique <0,5%, un sur risque de 3,2 était observé (95%CI=1,1-9,5; p=0,04).

REMERCIEMENTS

Au Professeur Florent Soubrier,

Vous me faites l'honneur de présider cette soutenance de Thèse de Sciences.

J'ai eu la chance de travailler dans votre laboratoire et la majorité des travaux présentés n'auraient pas existé sans votre aide.

Depuis quelques années, j'ai le plaisir de co-encadrer des internes dans votre laboratoire et j'ai apprécié votre sens clinique et votre culture scientifique.

Veuillez trouver, dans ce travail, l'expression de mon respect et de mon admiration.

« D'une certaine façon, la génétique n'est qu'une mémoire. Celle de notre évolution, incrustée dans notre chair. »

de Jean-Christophe Grangé

Extrait du Le Concile de pierre

Au Professeur Alex Kartheuser,

Vous avez accepté d'être un rapporteur de cette thèse et je vous en remercie.

Votre expérience sur les polyposes rendait votre présence évidente.

C'est un honneur de vous compter dans ce jury.

« L'Angleterre s'écroule dans l'ordre, et la France se relève dans le désordre. »
de Winston Churchill

Au Professeur Christophe Penna,

Cher Christophe, merci encore d'avoir accepté d'être rapporteur. Depuis que je suis jeune interne tu as été présent lors des étapes qui jalonnent la vie professionnel d'un futur hospitalo-universitaire (DEA, thèse de médecine, comité de mi-thèse, CNU).

Vois dans cette thèse le résultat de ton investissement dans la recherche scientifique en tant que directeur des DEA qui m'avait guidé ainsi que la marque de ma reconnaissance.

« Le docteur a dit tout ça est normal
Il a trop peur de casser mon moral
Tu vois je vais bien je suis entouré

En septembre
Un soleil d'été
Dans une chambre
A Ambroise Paré »

de Vincent Delerm

Au Professeur Yves Panis,

Merci d'avoir accepté de juger ce travail scientifique.

Vous avez vécu les dernières retouches de cette thèse dans votre service…

Veuillez trouver dans ce travail, l'expression renouvelée de mon admiration pour votre force de travail et de mes remerciements pour cette année de clinicat passée à vos cotés.

« Les écrivains tombaient comme des mouches, remplacés par les "auteurs", ces espèces de techniciens de surface du supplément d'âme préfabriqué [...].
Avec Céline, l'outrage avait commencé à devenir récit. L'offense conduisait le bal. La complexité de l'humanité se réorganisait dans la trame radieuse d'une tapisserie d'injures. On pouvait le continuer, les motifs ne manquaient pas.
Ils manquent moins que jamais aujourd'hui." »

Philippe Muray

Au Professeur Christophe Mariette,

Même si je n'ai pu travailler dans ton service, j'ai depuis la fin de mon internat eu la chance de participer à certaines de tes études ou la rédaction des RPC sur les métastases hépatiques.

J'ai toujours admiré ton énergie, ta disponibilité qui rendent ses lettres de noblesse au métier de Professeur de Chirurgie.

C'est un honneur de t'avoir dans ce jury.

« Un mauvais chirurgien ne fait de mal qu'à une personne à la fois. Un mauvais professeur à 130. »

Ernest Boyer

Au Docteur Alex Duval,

Merci d'être présent à cette thèse de sciences après avoir participé au comité de mi-thèse.

Depuis quelques années, nous travaillons ensemble et j'ai été impressionné par la richesse et l'inventivité de tes recherches et pu profiter de ta simplicité dans le travail.

Je suis honoré de ta présence et de faire partie de ton équipe INSERM.

J'espère que cette thèse sera le début d'une collaboration plus que fructueuse.

« Un véritable savant, qui travaille dans son laboratoire, n'écrit point science avec un grand S. »

de Charles Péguy

Extrait des Cahiers de la quinzaine

Au Professeur Yann Parc,

Cette thèse de sciences a été portée par votre énergie et votre détermination. Depuis maintenant 10 ans j'ai la chance de vous avoir pour mentor.

Voyez dans ce travail, mes remerciements pour tout ce que je n'aurais pas pu réussir sans votre aide.

« Le professeur a réussi au moment où son élève devient original. »
de Lane Cooper

A ceux sans qui ces travaux n'auraient jamais abouti :

L'Ecole Doctorale De Physiologie et Physiopathologie du Professeur Philippe Cardot

Laboratoire de Saint-Antoine (INSERM U673)
- Le Docteur Patricia Forgez
- Le Docteur Christelle Rodrigue
- Le Docteur Christian Gespach

Laboratoire de la Pitié-Salpétrière
- Le Docteur Florence Coulet, qui m'a à peu près tout appris de la PCR et qui encadre parfaitement les internes en Master M2
- Le Docteur Chrystelle Colas, qui suis les patients et sans qui nous n'aurions pas de database
- Les techniciennes du laboratoire d'oncogénétique
- Le docteur Cécile Mongin
- Le docteur Mélanie Eyries
- Les docteurs Sophie Baert-Desurmont et le professeur Thiery Frebourg de Rouen

Laboratoire d'Oxford
- Le Professeur Walter Bodmer
- Le Docteur Cristina Bonilla
- Les Professeur David Kerr et Oliver Sieber
- Bruce Winney, Elaine Johnstone, Susan Tonks, Tammy Day, Katarzyna Hutnik, Abdelhamid Boumertit, Rachel Midgley, Shazad Asraf

Le service d'anatomopathologie de Saint-Antoine
- Le Professeur Jean-Francois Flejou
- Le Docteur Magali Svrcek
- Le Docteur Najat Mourra

Le Docteur Solen Kerneis qui m'a apporté une aide statistique indispensable.
Les Docteurs Morgan Roupret et Jean-David Zeitoun pour être d'indéfectibles supporters.

Les médecins qui m'ont entouré et formé pendant ces quatre années de thèse de sciences :

A Saint-Antoine :
- Le Professeur Emmanuel Tiret qui me guide depuis ces années
- Les Professeurs François Paye et Pierre Balladur
- Les docteurs Najim Chafai et Magali Lefrancois,
- Mes collègues Mircea Chirica, Thibault Desurmont, Valeria Peveri, Stéphane Zalinski, Séverine Noullet, Laure Kerger.

A Beaujon :
- Le Professeur Frédéric Bretagnol
- Mon premier et seul co-chef-PUPH, Mehdi Karoui
- Les docteurs Caroline Hatwell et Cécile Mongin

Et bien sur les internes….

Enfin ceux qui sont toujours là :

- Mes parents, qui ne comprendront pas tout de la thèse mais que je n'aurais pas pu rédiger sans eux.
- Mes amis, médecins ou non.

- Les deux femmes de ma vie, Marianne et Daphné.

« La vie est courte, l'art est long, l'occasion fugitive, l'expérience trompeuse, le jugement difficile. »

de Hippocrate

Extrait des Aphorismes

THESE DE DOCTORAT DE

L'UNIVERSITE PIERRE ET MARIE CURIE

PARIS VI

Doctorat des Sciences de la Vie

Physiologie et Physiopathologie (ED 394)

Présentée par

Monsieur Jérémie LEFEVRE

Le 5 septembre 2012

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

GENETIQUE DU CANCER COLORECTAL :

Polyposes adénomateuses non liées à APC et cancers de survenue précoce

JURY

Monsieur le Professeur Florent SOUBRIER	**Président**
Monsieur le Professeur Alex KARTHEUSER	**Rapporteur**
Monsieur le Professeur Christophe PENNA	**Rapporteur**
Monsieur le Professeur Yves PANIS	**Examinateur**
Monsieur le Professeur Christophe MARIETTE	**Examinateur**
Monsieur le Docteur Alex DUVAL	**Examinateur**
Monsieur le Professeur Yann PARC	**Directeur**

SOMMAIRE

LISTE DES FIGURES

LISTE DES TABLEAUX

LISTE DES ABBREVIATIONS

8oxoG : 7,8-dihydro-8-oxoguanine

BER : *Base Excision Repair*

CCR : cancer colorectal

CIMP : *CpG Island Methylator Phenotype*

CIN : instabilité chromosomique

GWAS : *Genome Wide Association Studies*

HNPCC : *Hereditary Non Polyposis colorectal cancer*

IHC : Immunohistochimie

MAP : *MUTYH Associated Polyposis*

MMR : *Mismatch Repair*

MSI : *Micro-Satellite Instability*

MSP : *Methylation Specific PCR*

PAF : Polypose Adénomateuse familiale

PCR : *Polymerase Chain Reaction*

PJ : Polypose Juvénile

SL : Syndrome de Lynch

SNP : *Single Nucleoid Polymorphism*

Les gènes sont cités en italique et les protéines sont citées en romain.

INTRODUCTION

1. Epidémiologie et histoire naturelle du cancer colorectal

1.1. Epidémiologie :

Le cancer colorectal est le 3ème cancer dans le Monde et touche plus de 36.000 personnes chaque année en France (Gellad, ZF *et al.* 2010). La mortalité imputée aux cancers colorectaux est de 11.000 décès par an en France. Sa survie actuelle est d'environ 50% à 5 ans tous stades confondus (Mitry, E *et al.* 2008). La prévalence élevée du CCR et sa gravité (17.000 décès par an) en font un véritable enjeu de santé publique. Il est donc fondamental de déterminer les facteurs de risques accessibles à une prévention primaire ou secondaire. Les facteurs acquis (nutrition, tabac, alcool) sont discutés comme facteurs de risque pour le CCR et, à l'heure actuelle, seuls les facteurs génétiques ont été clairement incriminés. L'aspirine a été retrouvé comme facteur protecteur du CCR en cas de traitement reçu pendant au moins 5 ans et à des doses supérieures à 300 mg par jour (Flossmann, E *et al.* 2007).

Le CCR est en effet, le premier des cancers héréditaires. Les études de jumeaux et de familles ont permis d'évaluer qu'environ 30% des CCR appartiennent à une forme génétique (Lichtenstein, P *et al.* 2000; Grady, WM 2003).

Une méta-analyse incluant 59 études a estimé que le risque relatif de développer un CCR en cas d'antécédent familial au premier degré est de 2,24 (95%IC : [2,06-2,43]). Ce risque passait à 3,97 si deux antécédents familiaux au premier degré existaient (Butterworth, AS *et al.* 2006). Seulement 5% de ces cas correspondent à une forme familiale cliniquement identifiée associée à une mutation. Les 25% restants n'ont toujours pas trouvé d'explication génétique. C'est pourtant, la recherche et l'indentification de gènes ou groupes de gènes responsables qui pourraient améliorer la surveillance, la prévention permettre un meilleur diagnostic et une prise en charge optimale.

1.2. Histoire naturelle du cancer colorectal:

Dans la majorité des cas, le CCR provient d'un polype adénomateux dégénéré. La séquence colon-adénome-adénocarcinome est maintenant bien caractérisée au plan histologique et génétique. Le cancer va ensuite infiltrer les différentes parois du colon ou du rectum et diffuser dans les ganglions lymphatiques puis à distance. Environ la moitié des patients avec un CCR présenteront des métastases hépatiques soit au moment du diagnostic soit à distance de la résection digestive. La classification TNM mise à jour en 2010 reste un des facteurs pronostiques majeurs.

ANATOMIC STAGE/PROGNOSTIC GROUPS					
Stage	T	N	M	Dukes*	MAC*
0	Tis	N0	M0	–	–
I	T1	N0	M0	A	A
	T2	N0	M0	A	B1
IIA	T3	N0	M0	B	B2
IIB	T4a	N0	M0	B	B2
IIC	T4b	N0	M0	B	B3
IIIA	T1–T2	N1/N1c	M0	C	C1
	T1	N2a	M0	C	C1
IIIB	T3–T4a	N1/N1c	M0	C	C2
	T2–T3	N2a	M0	C	C1/C2
	T1–T2	N2b	M0	C	C1
IIIC	T4a	N2a	M0	C	C2
	T3–T4a	N2b	M0	C	C2
	T4b	N1–N2	M0	C	C3
IVA	Any T	Any N	M1a	–	–
IVB	Any T	Any N	M1b	–	–

Tumeur primitive (T)

TIS = Carcinome in situ. Tumeur intra-épithéliale ou envahissant la lamina propria sans extension à la sous-muqueuse
T1 = Tumeur infiltrant la sous-muqueuse sans la dépasser
T2 = Tumeur infiltrant la musculeuse sans la dépasser
T3 = Tumeur envahissant toute la paroi sans la dépasser
T4 = Tumeur dépassant la séreuse

Envahissement ganglionnaire (N)

N0 = Absence de métastase ganglionnaire régionale
N1 = Métastase dans un à trois ganglions régionaux
- N1a = 1 ganglion envahi
- N1b = 2 ou 3 ganglions envahis
- N1c = présence d'emboles tumoraux dans la sous-séreuse sans métastase ganglionnaire

N2 = Métastase dans 4 ganglions régionaux ou plus
- N2a = 4 à 6 ganglions envahis
- N2b = 7 ganglions envahis ou plus

Nx = Statut ganglionnaire inconnu

Métastases (M)

M0 = Absence de métastase
M1 = Présence de métastases à distance
- M1a = limité à un site ou organe
- M1b = touchant plusieurs organes ou le péritoine

Mx = Statut métastatique inconnu

Figure 1. Classification TNM des cancers colorectaux (AJCC 7ème édition).

1.3. Cancérogenèse colorectale:

Si les CCR sont très homogènes au niveau histologique (plus de 90% d'adénocarcinome), trois mécanismes différents ont été individualisés dans la cancérogenèse colorectale :

- L'instabilité chromosomique (CIN, *Chromosomal Instability*)
- L'instabilité génétique ou microsatellitaire (MSI : *Microsatellite Instability*)
- Hyperméthylation des ilôts CpG (CIMP : *CpG Island Methylator Phenotype*).

Ces mécanismes vont modifier le fonctionnement de certaines voies de signalisation cellulaires et entrainer la cellule vers un phénotype tumoral. Au moins deux de ces mécanismes sont étayés par l'existence de syndrome de prédispositions héréditaires majeures au CCR : la polypose adénomateuse familiale (PAF) et le syndrome de Lynch.

1.3.1. Instabilité chromosomique (phénotype CIN):

Elle rend compte d'environ 80% des cancers sporadiques (Laurent-Puig, P *et al.* 2010) et est caractérisée par des pertes alléliques sur les bras courts des chromosomes 17, 8 et sur le bras long des chromosomes 18, 5 et 22. Ces anomalies sont le plus souvent associées à des mutations sur les gènes *APC* ou *TP53* qui entrainent une inactivation complète de ces gènes suppresseurs de tumeur. On retrouve fréquemment une aneuploïdie cellulaire. L'origine de cette instabilité est encore mal connue, mais il a été montré que des mutations sur le gène *APC*, aboutissant à la formation d'un codon stop, peuvent favoriser la CIN. En effet une des fonctions de la protéine APC est de maintenir la polymérisation des microtubules du noyau cellulaire (Fodde, R *et al.* 2001).

1.3.2. Instabilité génétique (phénotype MSI):

Ce mécanisme est retrouvé dans environ 15% des CCR (Laurent-Puig, P *et al.* 2010) et se caractérise par une instabilité des locus microsatellitaires liée à un défaut de réparation des mésappariements de l'ADN, tâche normalement dévolue au système MMR (*Mismatch Repair*) composé en autres gènes *hMLH1, hMLH3, hMSH2, hMHS6, hPMS1, hPMS2*… Ces séquences microsatellitaires sont très fréquentes dans l'ensemble du génome et sont particulièrement à risque d'être mal répliquées par l'ADN polymérase. Le système MMR peut être inactivé en cas de mutation germinale associée à une mutation somatique (Syndrome de Lynch) ou par méthylation du promoteur de *hMLH1* qui inactive sa transcription (forme sporadique). Le phénotype MSI est identifié par PCR en testant des séquences microsatellitaires de longueur connues et en comparant la taille des produits d'amplification entre le tissu sain et le tissu tumoral (figure 2).

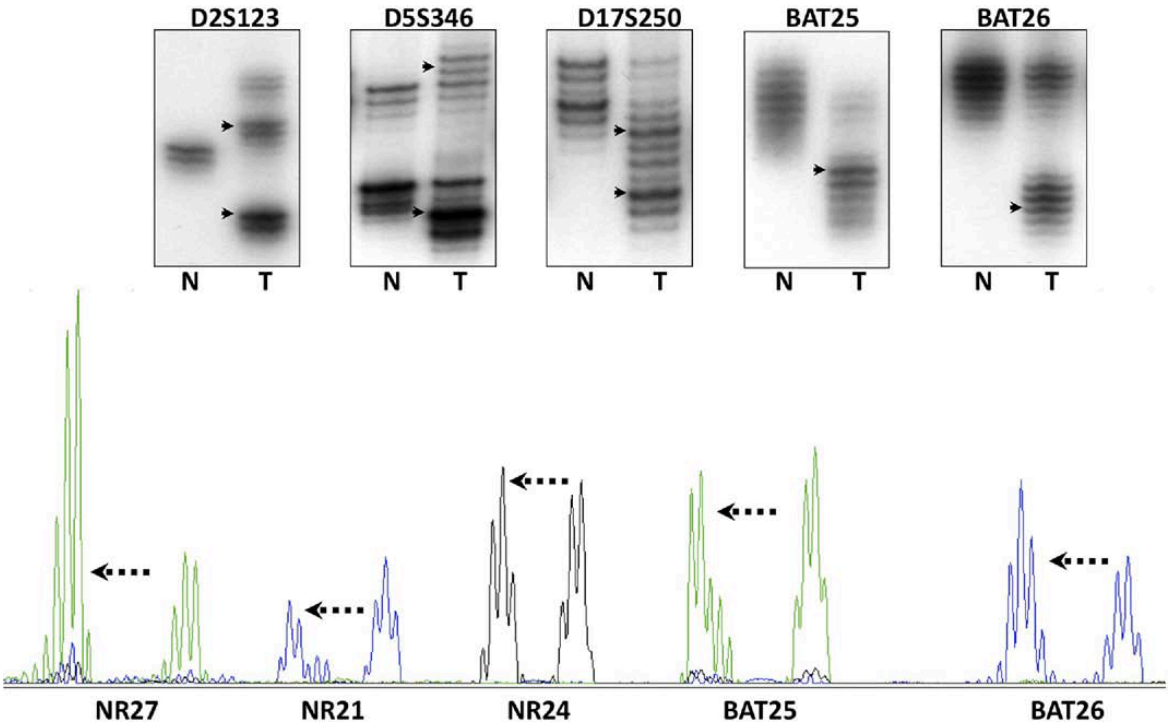


Figure 2. Phénotype MSI mis en évidence par amplification de 5 marqueurs (d'après (Boland, CR et al. 2010))

Contrairement aux tumeurs avec un phénotype CIN, les cellules MSI sont le plus souvent diploïdes et on retrouve moins fréquemment des mutations sur *APC* ou *TP53* (Olschwang, S *et al.* 1997). L'inactivation du système MMR conduit à l'accumulation de mutations secondaires qui vont inactiver de nombreux gènes aboutissant à la transformation de la cellule. Les gènes *BAX*, *TGFRII*, les facteurs de transcription TCF-4 ou E2F4 sont souvent retrouvés mutés dans les cancers MSI+ (Duval, A *et al.* 2002).

1.3.3. Hyperméthylation de l'ADN (phénotype CIMP):

Des dinucléotides CpG regroupés en ilots sont retrouvés dans les régions promotrices pour la moitié des gènes. Ces îlots CpG peuvent inactiver un gène en empêchant sa transcription si leur cytosine est méthylée (figure 3).

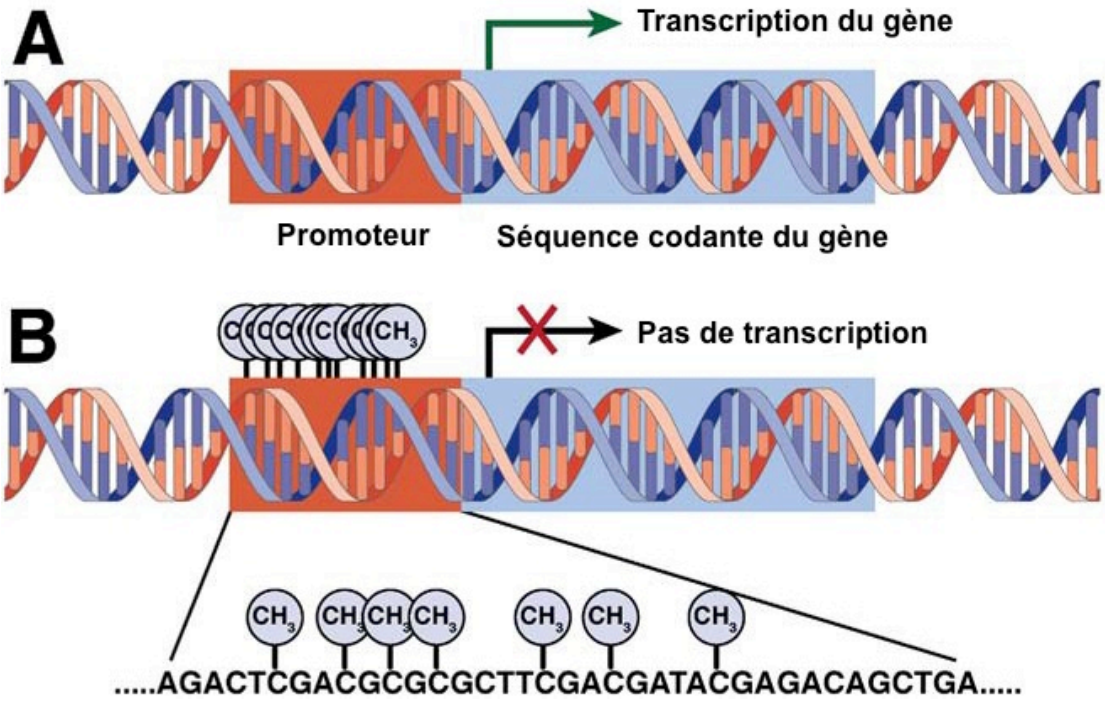


Figure 3. Mécanisme d'inactivation de la transcription (d'après (Leggett, B *et al.* 2010)).

Le mécanisme du phénotype CIMP n'est pas encore élucidé mais certains gènes semblent particulièrement sensibles pour le définir. En effet les gènes *hMLH1, RUNX3, IGF2,*

CACNA1G, NEUROG1 et *SOCS1* ont été retrouvé méthylés sur leur promoteur en cas de phénotype CIMP et pourrait devenir le panel de gènes à étudier. Actuellement il n'existe aucune définition consensuelle du phénotype CIMP (nombre de gènes à étudier, nombre de promoteurs méthylés).

Dans le CCR, plusieurs gènes suppresseurs de tumeurs peuvent ainsi être inactivés conduisant à ce phénotype CIMP (*CpG Island Methylation Phenotype*) (Issa, JP 2004). La méthylation du gène *hMLH1* responsable d'une inactivation du système MMR et donc d'un phénotype MSI rentre également dans ce cadre expliquant la possibilité de tumeur MSI+/CIMP+. La plupart des CCR sporadiques présentant un phénotype MSI surviennent chez des individus âgés, sont associés à des mutations de *BRAF* et à un phénotype CIMP+. Ces cas dérivent le plus souvent de la voie des adénomes festonnés.

1.4. Importance des mécanismes :

Ces mécanismes ne sont pas exclusifs comme le démontre la figure 4.

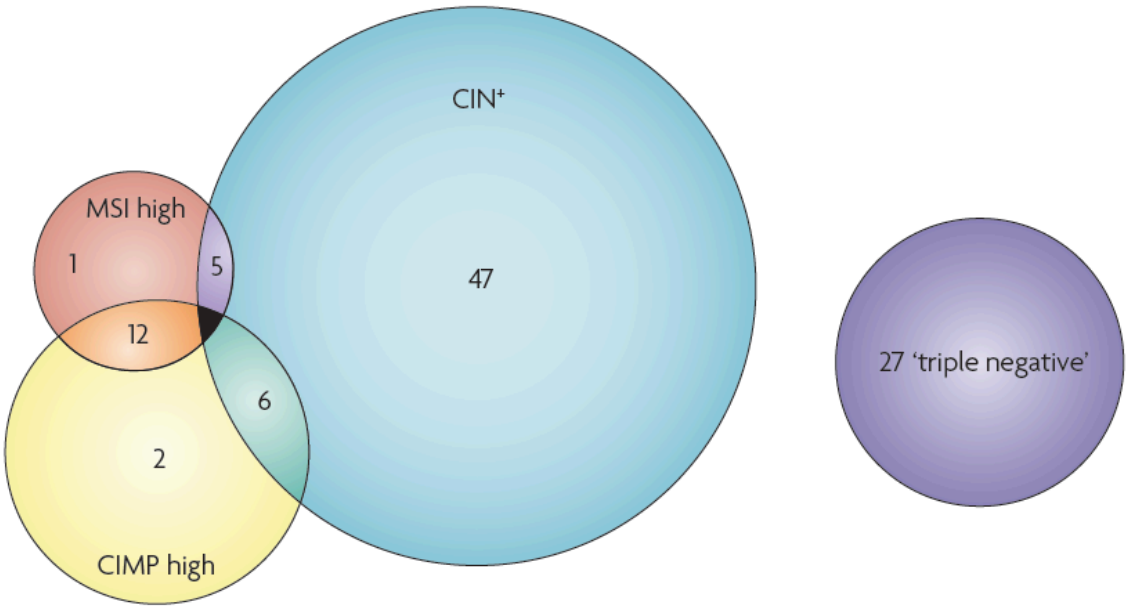


Figure 4. Intrication des différentes voies de la cancérogenèse. Les valeurs sont des pourcentages. (D'après (Walther, A et al. 2009))

La connaissance des mécanismes permet d'améliorer la compréhension de la cancérogenèse colorectale. On commence à voir poindre des voies de la cancérogenèse du CCR en fonction du type histologique du polype à l'origine du cancer (Figure 5). La voie du CIMP est responsable de CCR à partir d'adénomes festonnés alors que les adénomes tubuleux sont le plus souvent liés au mécanisme d'instabilité chromosomique.

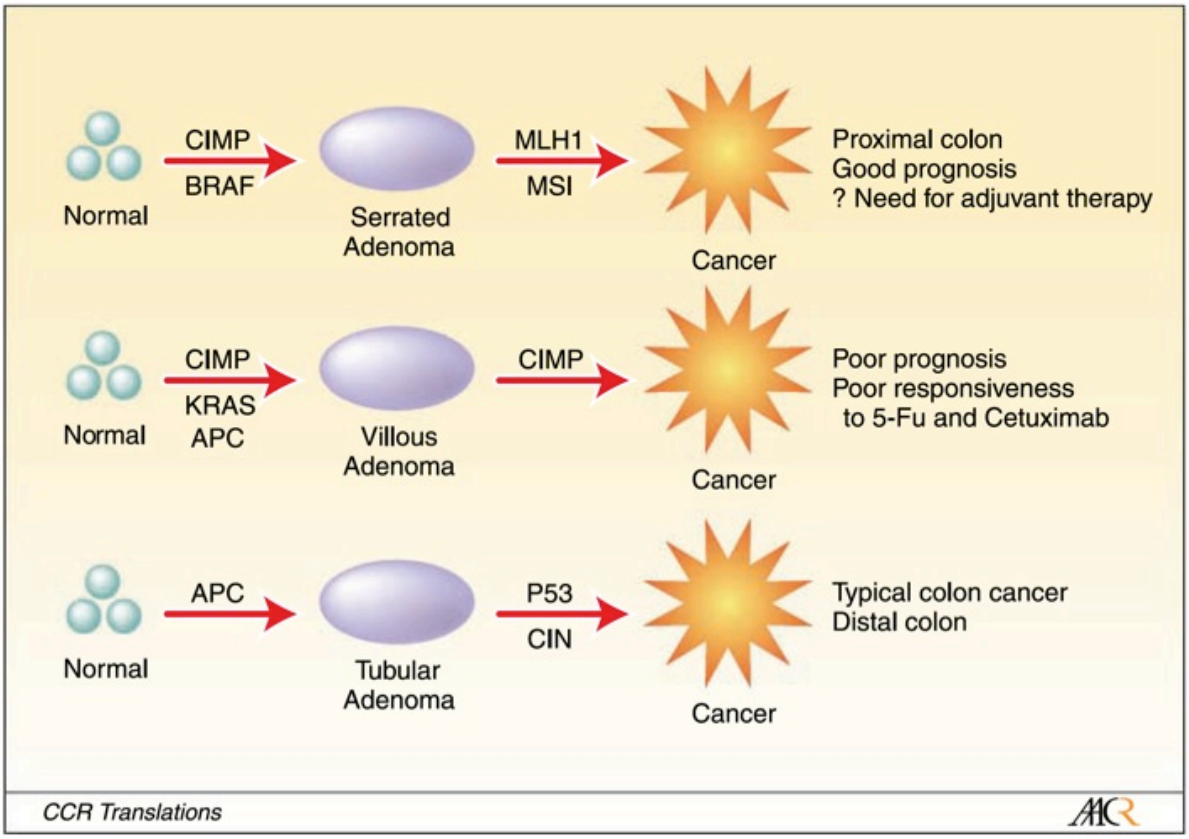


Figure 5. Voies de la cancérogenèse. D'après (Issa, JP 2008)

Ces mécanismes ont enfin également un impact sur le pronostic. Les CCR MSI+ ont un meilleur pronostic que les CCR CIN+ et les patients présentent moins fréquemment des métastases à distance. Deux méta-analyses ont observé un hazard-ratio de décès de 1,45 pour les tumeurs CIN+ et de 0,65 pour les MSI (Walther, A *et al.* 2008). Certains mécanismes sont également associés à des résistances à la chimiothérapie ou à certaines biothérapies (figure 5).

1.5. Voies de signalisation cellulaire:

La cancérogenèse est un processus composé de multiples étapes nécessitant l'accumulation d'anomalie génétiques héritées et acquises. Ces anomalies vont donner à la cellule des capacités nouvelles et un avantage de croissance par rapport aux cellules adjacentes. Six caractéristiques vont être acquises : l'insensibilité aux signaux inhibiteurs de la croissance cellulaire, perte de l'apoptose, multiplication cellulaire infinie, acquisition de l'angiogenèse, acquisition de la possibilité d'envahir les tissus avoisinant et possibilité de métastaser à distance. Certains oncogènes ou suppresseur de tumeur ont été identifiés comme favorisant le passage d'une cellule normale en une cellule cancéreuse. Même si ce sont les gènes qui sont étudiés et décris, ce sont bien sur les protéines codées par ces gènes qui agissent sur la mécanique cellulaire.

En fonction du mécanisme de progression prépondérant au sein d'un CCR (CIN, MSI ou CIMP), une voie de signalisation est préférentiellement affectée. Les 4 voies de signalisation habituellement impliquées dans la cancérogenèse du CCR sont la voie Wnt (ou *APC*/B-caténine), la voie du TGFB, la voie RAS/MAPK et la voie p53.

1.5.1. Voie de signalisation Wnt :

La voie Wnt est une des signalisations biochimiques les plus indispensables en biologie du développement et en cancérogénèse. On peut résumer la voie Wnt ainsi (Klaus, A *et al.* 2008) (Figure 6) :

- En l'absence du ligand Wnt, la β-caténine est recrutée dans un complexe associant la protéine APC et l'Axine. Dans ce complexe, la β-caténine va être phosphorylée par les enzymes GSK3-β et CK1α et être adressée vers une dégradation via le protéasome. Le niveau de β-caténine cytoplasmique reste faible. L'expression des gènes cibles de la voie Wnt est alors réprimée par l'association de Groucho aux facteurs de transcription LEF et TCF.

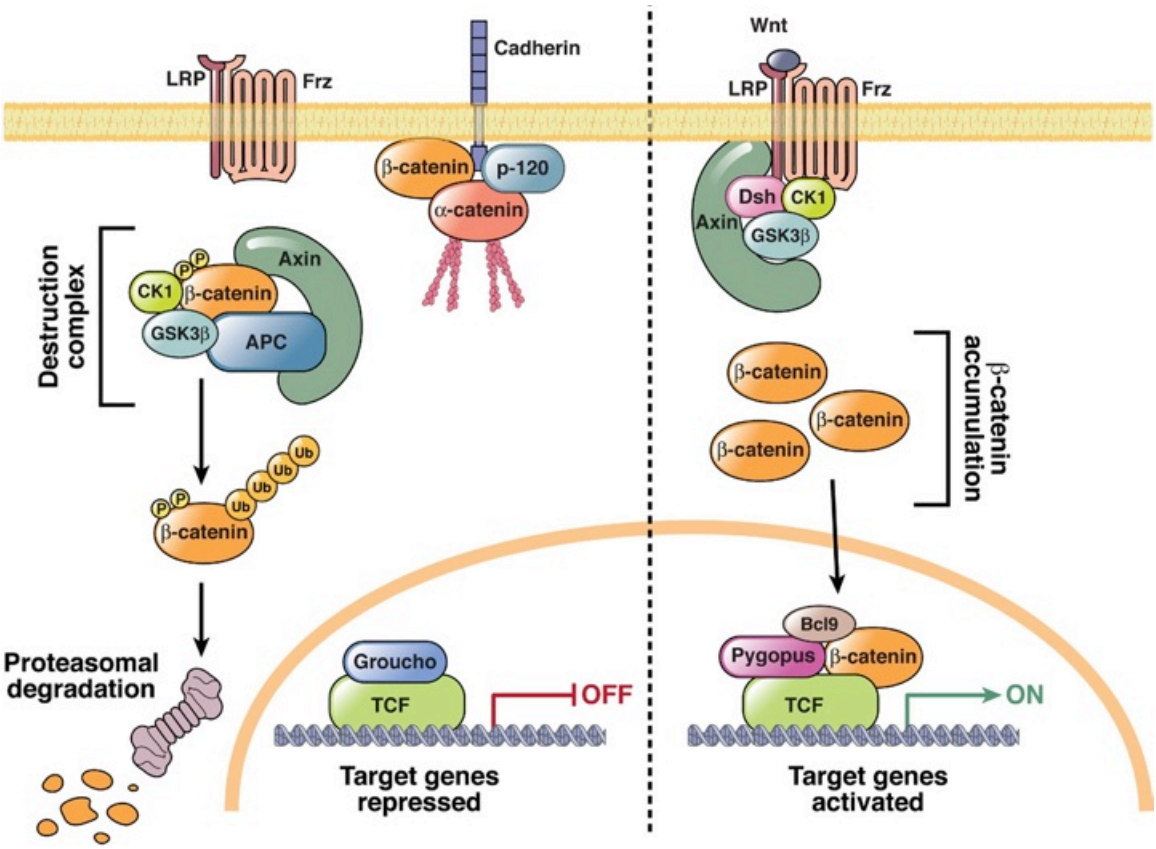


Figure 6. Voie de signalisation Wnt en l'absence du ligand Wnt / en présence du ligand Wnt d'après (Pino, MS *et al.* 2010).

- En présence du ligand Wnt, il va se lier à son récepteur Frizzled et ses co-récepteurs (LRP5/LRP6). L'Axine va interrompre la dégradation de la β-caténine via DVL (*dishevelled*). La β-caténine libre dans le cytoplasme, va se transloquer dans le noyau pour former un complexe d'activation transcriptionnel avec LEF et TCF en déplaçant Groucho et en recrutant des coactivateurs (Bcl9, Pygopus et CBP). Les gènes cibles alors exprimés sont nombreux : *MYC, FGF, VEGF, CyclinD1, E-Cadhérine, Matrisyline*... Plusieurs études ont montré que la voie Wnt était activée par des mutations du gène *APC* dans la polypose adénomateuse familiale et dans environ 80% des formes sporadiques de cancers colorectaux, (tumeurs avec instabilité chromosomique) (Kinzler, KW *et al.* 1991; Korinek, V *et al.* 1997). Dans les tumeurs avec instabilité des microsatellites, les mutations d'*APC* sont plus rares et la voie

Wnt est généralement activée par des mutations du gène de la β-caténine et/ou par des mutations du gène de l'Axine (Narayan, S *et al.* 2003).

1.5.2. Voie de signalisation du TGFβ:

La voie du TGFB est impliquée dans de nombreux types de cancer humains. Cette vois de signalisation peut être schématisée ainsi : le TGFB se fixe à deux récepteurs transmembranaires (TGFBR1 et TGFBR2). La fixation du ligand provoque la phosphorylation du TGRBR1 par le TGFBR2. Le signal est ensuite transmis en intra-cellulaire via les voies SMAD et non-SMAD en activant notamment SMAD2, SMAD3. Une fois activée, ces protéines se fixent à SMAD4 et sont alors transloquées dans le noyau. Ce complexe SMAD va alors réguler l'expression de certains gènes (figure 7).

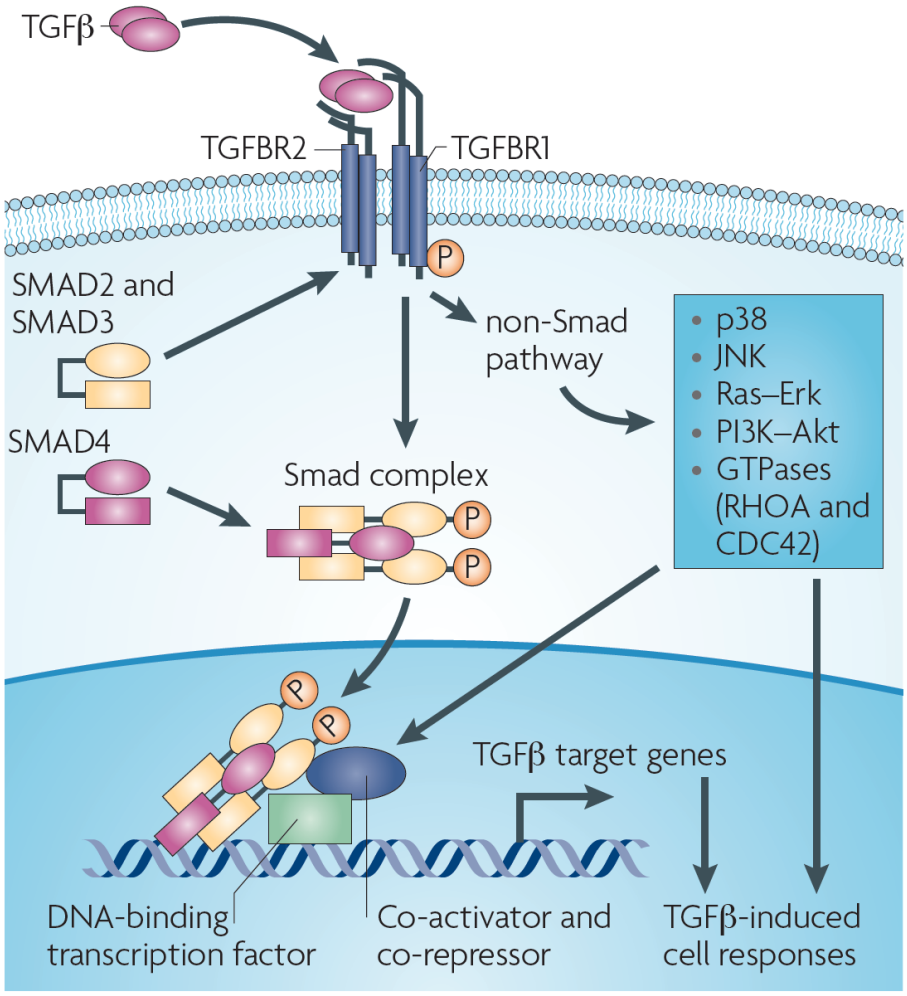


Figure 7. Voie de signalisation du TGFB (D'après (Ikushima, H *et al.* 2010)).

Cette voie de signalisation est impliquée dans les cancers colorectaux par plusieurs mécanismes : le *TGFBR2* est inactivé en cas de phénotype MSI dans 60 à 90% des cas (Markowitz, S *et al.* 1995), des mutations germinales sur *SMAD4* ou *BMPR1A* prédisposent aux polyposes juvéniles, *SMAD4* est également inactivé par perte d'hétérozygotie dans 20 à 30% des cas , enfin des variants communs de *TGFBR1* ont été décrits comme prédisposant au CCR (Houlston, RS *et al.* 2008) .

1.5.3. Voie de signalisation Ras/Raf/MAPKinase :

Elle appartient avec la voie de la PI3K à la voie de l'EGFR (Epidermal growth factor receptor) qui est impliquée dans le contrôle de l'apoptose et la prolifération cellulaire (figure 8).

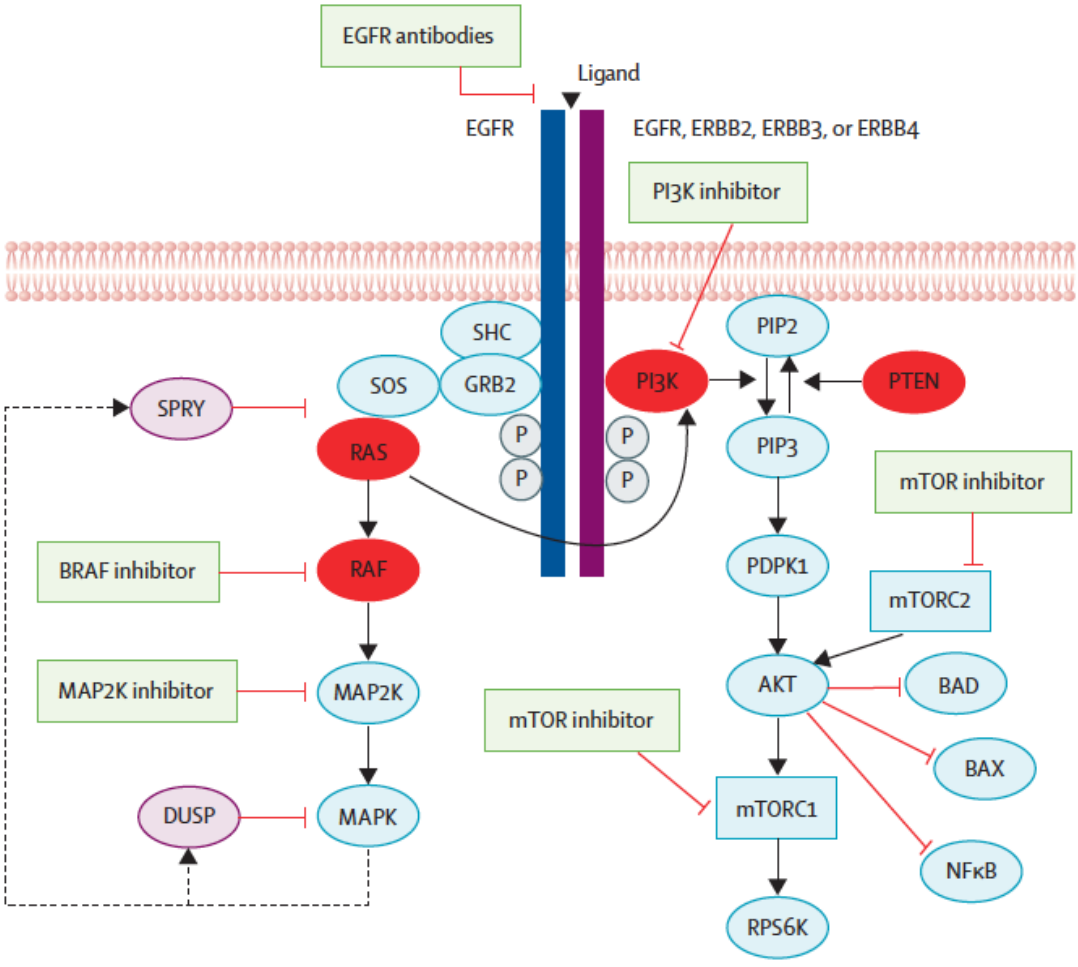


Figure 8. Représentation schématique de la voie de l'EGFR (kras et PI3K), d'après (De Roock, W *et al.* 2010)

Kras appartient à la famille des protéines Ras avec HRAS et NRAS. Elle codée par le gène du même nom et active la voie MAPK et PI3K. Environ 40% des CCR présentent une mutation sur *kras* située dans la majorité des cas sur les codons 12 ou 13 (Allegra, CJ *et al.* 2009). Les mutations de *Kras* sont essentiellement trouvées dans les tumeurs avec instabilité chromosomique, MSS et CIMP- . (Figure 8)

BRAF appartient à la famille Raf avec ARAF1 et RAF1 et est une protéine activée en aval de Kras. Environ 15% des CCR présentent des mutations sur BRAF, elles sont mutuellement exclusives avec les mutations de *kras* (De Roock, W *et al.* 2010). La mutation la plus fréquemment reportée sur *BRAF* est V600E au sein du domaine kinase de la protéine et est observée principalement sur les CCR sporadiques avec phénotype MSI et CIMP+. En revanche, elle n'est jamais observée en cas de mutation germinale sur un gène MMR dans le cadre d'un syndrome de Lynch (Kambara, T *et al.* 2004).

Le proto-oncogène *k-ras* sur le chromosome 12 est retrouvé muté chez environ 50% des CCR (Robbins, DH *et al.* 2002).

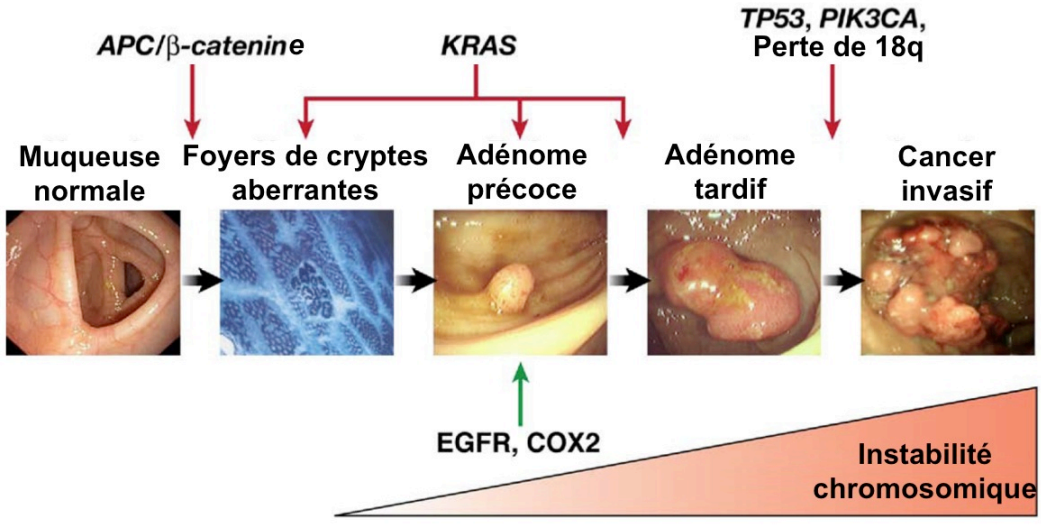


Figure 9. Séquence adénome-cancer, instabilité chromosomique.

2. Formes héréditaires de cancer colorectal :

Les syndromes prédisposant au CCR ont tout d'abord été définis par leur présentation clinique avant que les gènes responsables ne soient identifiés. On peut distinguer principalement les polyposes et le syndrome de Lynch (OMIM : 120435 et 609310). Chaque syndrome a son mécanisme propre : la polypose adénomateuse familiale est liée à une instabilité chromosomique, le syndrome de Lynch est caractéristique du phénotype MSI. La polypose liée à *MUTYH* est pour le moment la moins bien caractérisée.

2.1. Syndrome de Lynch :

Il est responsable d'environ 2-4% des CCR (Hampel, H *et al.* 2008). Les principales affections sont le CCR et le cancer de l'endomètre avec un risque au cours de l'existence compris entre 50-80% et 40-60% respectivement (Stoffel, E *et al.* 2009). Les polypes coliques sont rares et le cancer survient préférentiellement à un âge jeune. Le syndrome de Lynch (SL) expose également les patients à d'autres cancers mais avec une fréquence beaucoup plus faible : estomac (15%, ovaires 10%, hépatobiliaire 5%, tractus urinaire supérieur 5%, pancréas 3%, intestin grêle 3% et glioblastome 2%).

2.1.1. Diagnostic d'un syndrome de Lynch :

Devant l'absence d'un phénotype clairement identifiable (comme la présence de centaines d'adénomes en cas de polypose) d'un SL, plusieurs outils ont été décrits pour poser le diagnostic de SL. Des critères cliniques à la sensibilité variable ont été développés par les sociétés savantes : Amsterdam I puis II et Bethesda (Balmana, J *et al.* 2008; Jasperson, KW *et al.* 2010; Trano, G *et al.* 2010).

Deux tests sont également disponibles pour bénéficier rapidement d'arguments en faveur du diagnostic : le marquage immunohistochimique (IHC) et la recherche d'instabilité des microsatellites par PRC. Environ 90% des SL ont un phénotype MSI mais 15% des CCR

sporadiques également en raison d'une méthylation acquise du promoteur d'*hMLH1* (Cancers MSI+/CIMP+) (Jasperson, KW *et al.* 2010). L'IHC utilise 4 anticorps dirigés contre les 4 principales protéines dont l'expression peut disparaître en cas de mutation. Ce test a l'avantage par rapport au phénotype MSI de diriger la recherche sur le gène muté. En effet, le seul test de certitude reste le séquençage génétique avec identification de la mutation germinale. Actuellement les recommandations sont de proposer une consultation d'oncogénétique aux:

- 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 d'un phénotype MSI sur la pièce tumorale est indiquée :

- 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.

L'extension des indications de recherche du phénotype MSI ou d'une perte d'expression des protéines en IHC doit en effet être la plus large possible afin de ne manquer aucun patient avec un LS. Une étude de notre équipe a en effet montré que la stricte application des critères de Bethesda entrainait la non identification d'un phénotype MSI chez 37% des patients (Canard, G *et al.* 2011) (Annexe 1, page 209). Actuellement tous les CCR opérés dans notre département bénéficient d'une IHC et de la recherche du phénotype MSI par PCR.

2.1.2. Génétique du syndrome de Lynch :

C'est une forme de cancer MSI+. Le SL est un syndrome de transmission autosomique dominante et résulte de l'existence d'une mutation germinale sur un des gènes du système MMR (*MisMatch Repair*). Il est composé des gènes *hMLH1, hMSH2, hMSH6* et *hPMS2* et corrige les erreurs du code génétique survenant lors de la réplication de l'ADN. Les mutations sur *hMLH1* et *hMSH2* représentent environ 90% des SL, *hMHS6* environ 10% et *hPMS2* est rarement trouvée mutée (Peltomaki, P *et al.* 2004). En accord avec la théorie de Knudson (Knudson, AG, Jr. 1971), les mutations doivent être biallélіques afin d'inactiver ces gènes suppresseurs de tumeur. La première est héritée et la deuxième est acquise. La mutation biallelique d'un des gènes du système MMR entraîne l'accumulation de mutations somatiques dans différents gènes impliqués dans le CCR, tels : le récepteur du *TGFβ-II* (Markowitz, S *et al.* 1995) ou de l'*IGF-II* (Souza, RF *et al.* 1996) ou le gène *BAX* régulateur de l'apoptose (Rampino, N *et al.* 1997) ou même certains gènes du système MMR comme $hMSH_3$ ou $hMSH_6$.

Dans certaines familles présentant un phénotype de syndrome de Lynch typique avec un phénotype MSI et une perte d'expression de MSH2, aucune mutation germinale n'est identifiée sur un des gènes du système MMR. Récemment, certains de ces cas ont été explicités par l'existence d'une délétion germinale du gène *EpCAM* (également nommé *TACSTD1*) (Ligtenberg, MJ *et al.* 2009). La délétion provoque une fusion des transcrits EpCAM-MSH2 responsable de l'inactivation de la protéine et serait responsable d'environ 6% des SL (Niessen, RC *et al.* 2009).

Il existe enfin quelques cas décrit de méthylation germinale du promoteur de *hMLH1* responsable de SL sans mutation germinale identifiée (van Roon, EH *et al.* 2010).

2.1.3. Prise en charge d'un syndrome de Lynch :

Le diagnostic de SL est fondamental pour la prise en charge des patients et de leurs apparentés. En effet, le type de résection colique peut différer (colectomie sub-totale au lieu d'une colectomie segmentaire), une hystérectomie prophylactique peut être proposée (Schmeler, KM *et al.* 2006) et la surveillance doit être plus régulière (tous les 1-2 ans) pour tous les porteurs de la mutation. Cette surveillance permet de réduire significativement l'incidence du CCR (11% *vs*. 27%) ainsi que sa mortalité (2% *vs*. 12%) (Stupart, DA *et al.* 2009). Le choix entre colectomie totale et colectomie segmentaire n'a pas été encore clairement posé. Les recommandations de 2010 laissent le choix au chirurgien en privilégiant la colectomie totale avec anastomose iléo-rectale surtout chez les patients jeunes (Collectif. 2009). L'équipe de Parry a montré que une cohorte de 382 patients ayant un LS prouvé que le risque de CCR métachrone était réduit de 31% (95%IC : 12%-46%; p=0,002) pour chaque segment de 10 cm retiré (Parry, S *et al.* 2011).

2.2. Polypose adénomateuse familiale :

Contrairement au syndrome HNPCC, la PAF (OMIM : 175100) a une expression clinique facile à diagnostiquer : les patients présentent une multitude de polypes dans le côlon. Sa prévalence est estimée aux alentours de 1/10000 (Jasperson, KW *et al.* 2010). Sans traitement, ces polypes vont être responsables de l'apparition d'un CCR. L'âge moyen de diagnostic du cancer est de 39 ans et le cancer est très rare avant 20 ans (Penna, CP *et al.* 1992). Les autres manifestations de ce syndrome sont nombreuses et sujettes à de grandes variations interindividuelles : apparition d'une hyperplasie de l'épithélium pigmentaire de la rétine, de polypes duodénaux, de tumeurs desmoïdes, d'ostéomes, de cancers thyroïdiens…

2.2.1. Génétique de la polypose adénomateuse familiale :

La PAF est un syndrome de transmission autosomique dominante avec une pénétrance de 100%. Son apparition est consécutive à une mutation congénitale sur le gène *APC*

(GenBank ID : 324). Constitué de 15 exons avec une séquence de 8532 paires de bases codant pour une protéine de 2843 acides aminés (Macrae, F *et al.* 2009). Les mutations congénitales d'*APC* sont le plus souvent des mutations ponctuelles (Nagase, H *et al.* 1992). On estime que les mutations *de novo* d'*APC* sont responsables d'environ 20 à 25% des cas de PAF (Jasperson, KW *et al.* 2010). Les mosaïques germinales sont également une explication de PAF de novo dans 20% des cas (Hes, FJ *et al.* 2008). Une mutation germinale est identifiée chez environ 70% à 80% des cas en cas de polypose adénomateuses classique (Miyoshi, Y *et al.* 1992; (Lamlum, H *et al.* 2000). *APC* code pour une protéine impliquée dans la voie de signalisation cellulaire Wnt (Figure 6). Son inactivation par un deuxième événement inactivant le gène sur l'allèle non muté va provoquer l'apparition de polypes puis de CCR. Ces tumeurs sont typiquement MSS.

2.2.2. Polypose atténuée et corrélation génotype-phénotype :

A la suite d'une observation d'une famille présentant une forme atténuée de polypose, la PAF a été scindée en deux syndromes : classique ou atténuée (Lynch, HT *et al.* 1995). Les principales caractéristiques de la forme atténuée sont un nombre limité de polypes (moins de 100), un âge d'apparition des polypes plus tardif (vers 35 ans) et moins de manifestations extra-coliques (Lynch, HT *et al.* 1995; Knudsen, AL *et al.* 2003; Nieuwenhuis, MH *et al.* 2007). Certains auteurs ont même subdivisé les polyposes adénomateuses en trois entités : profuse, intermédiaire et atténuée (Nieuwenhuis, MH *et al.* 2007). Parallèlement à la description d'entité clinique, une corrélation entre la mutation germinale et l'expression phénotypique de la maladie a été envisagée. Le phénotype atténuée était lié à la présence d'une mutation sur certaines régions d'*APC* : l'extrémité 5', l'extrémité 3' et l'exon 9 (Spirio, L *et al.* 1993; van der Luijt, RB *et al.* 1995; Gebert, JF *et al.* 1999).

Une représentation schématique publiée dans une revue récente sur la PAF résume les rôles de la protéine APC ainsi que les principales corrélations génotype-phénotypes (figure

10). Cette apparente relation a conduit les mêmes auteurs à recommander d'adapter la prise en charge des polyposes (colectomie totale et anastomose iléo-rectale à la place d'une colo-proctectomie totale avec anastomose iléo-anale) en fonction de la localisation de la mutation (Vasen, HF *et al.* 1996; Wu, JS *et al.* 1998; Bulow, C *et al.* 2000).

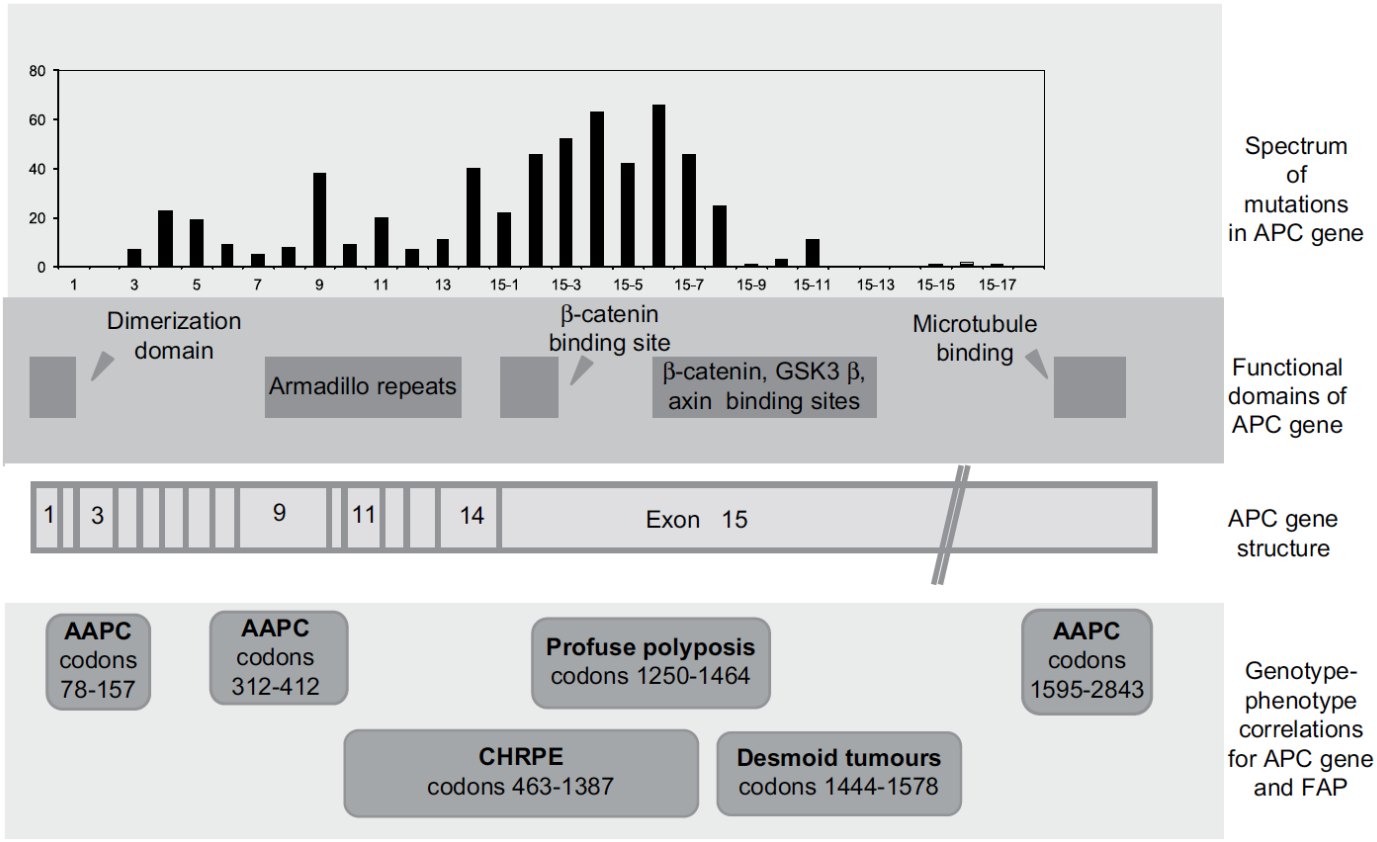


Figure 10. Rôle du gène APC, position des principales mutations germinales et corrélation génotype-phénotype. D'après (Macrae, F et al. 2009)

Cependant de nombreux arguments allaient contre cette corrélation entre l'expression colique de la maladie et la mutation : les limites des régions impliquées variaient grandement entre les publications et la grande variabilité d'expression de la maladie chez les porteurs de la même mutation (Giardiello, FM *et al.* 1994). Enfin la publication de cas clinique de famille diagnostiquée comme ayant une polypose atténuée avec diagnostic de cancer du rectum T3

chez un fils âgé de 16 ans confirmait que cette corrélation n'était pas fiable (Matsuo, S *et al.* 2001).

En revanche la corrélation génotype-phénotype semble plus significative pour d'autres affections de la PAF : les tumeurs desmoïdes sont significativement plus fréquentes lorsque la mutation est au delà du codon 1444 (Caspari, R *et al.* 1995; Lefevre, JH *et al.* 2008) (Annexe 2, page 217) et l'atteinte rétinienne lorsque la mutation est avant le codon 1444 (Caspari, R *et al.* 1995). Ces résultats ont été confirmés par une étude de cohorte internationale avec 2260 patients ayant une PAF. Les facteurs de risque identifiés étaient : la présence d'antécédent familial de tumeur desmoïdes, une mutation au delà du codon 1444 et les antécédents de chirurgie abdominale (Nieuwenhuis, MH *et al.* 2011) (Annexe 3, page 226).

2.2.3. Prise en charge d'une polypose adénomateuse familiale :

Elle fait l'objet de recommandations pour la pratique clinique en 2009 par l'INCA (Collectif. 2009) et en 2008 par le groupe de Majorque (Vasen, HF *et al.* 2008). Le dépistage des polypes colorectaux doit débuter à partir de 10-12 ans. L'âge de la colectomie et le type d'intervention doivent être guidés par les constatations endoscopiques et non pas par la mutation germinale. Habituellement, la chirurgie prophylactique (colectomie totale ou coloproctectomie totale) a lieu aux alentours de 20-25 ans et dépend du nombre de polypes rectaux. En dessous de 20 polypes rectaux, une conservation rectale est le souvent possible au prix d'une surveillance annuelle du rectum. Cette intervention est maintenant réalisée habituellement sous cœlioscopie, sans nécessité d'iléostomie de protection ce qui limite le traumatisme pariétal et la morbidité associée à ce type de stomie (Hor, T *et al.* 2012).

Après colectomie totale, le risque de décès est lié aux polypes duodénaux qui sont le lit du cancer duodénal et aux tumeurs desmoïdes. Le dépistage par une endoscopie haute à vision latérale est indispensable tous les 3 ans au moins. Le traitement des polypes dépend des constations endoscopiques et de l'examen histologique. Une exérèse endoscopique est le plus

souvent suffisante mais dans de rares cas, une chirurgie prophylactique (duodéno-pancréatectomie céphalique) est parfois nécessaire associée à une lourde morbidité (Parc, Y *et al.* 2011; Caillie, F *et al.* 2012).

Les tumeurs desmoïdes restent l'affection extra-colique la plus difficile à traiter ce qui explique qu'elles soient la première cause de mortalité après chirurgie prophylactique colorectale (Nieuwenhuis, MH *et al.* 2011). Il n'existe actuellement aucun traitement ni intervention pour prévenir leur apparition. Leur prise en charge (chimiothérapie, radiothérapie, résection chirurgicale, surveillance) dépend de la localisation (paroi musculaire abdominale, racine du mésentère, membres), la taille, le mode évolutif ainsi que du retentissement clinique.

2.3. Polypose liée à MUTYH :

2.3.1. Découverte du syndrome MAP :

En 2002, Al-Tassan et son équipe (Al-Tassan, N *et al.* 2002) rapportent le cas d'une famille anglaise N., où trois enfants sur 7 présentaient des polypes nombreux et des CCR tous découverts avant 50 ans. Les parents ne présentaient aucun antécédent de cancer. Tout le gène *APC* fut séquencé sans retrouver de mutation congénitale. La recherche d'un phénotype MSI fut également réalisée. Aucun argument pour une déficience du système MMR ne fut détecté. L'étude du gène *APC* dans les tumeurs mettait en évidence 18 mutations acquises dont 15 (83%) étaient des transversions G:C→T:A. Dans ces tumeurs les mutations bialléliques retrouvées inactivaient la protéine APC. Cette proportion était statistiquement supérieure aux 15% habituellement retrouvés parmi les mutations somatiques dans les tumeurs sporadiques. Cette découverte conduisit l'équipe à étudier les 3 gènes du système BER impliqués dans la réparation des lésions induites par le 8oxoG et conduisant à des transversions G:C→T:A. Seul le gène *MUTYH* (GeneID 4595) présentait des mutations congénitales : Tyr165Cys et Gly382Asp. Chaque enfant malade était hétérozygote composite. Tous les autres membres

sains de la famille présentaient une absence de mutation ou étaient hétérozygotes pour une des mutations.

2.3.2. Système BER :

Le syndrome MAP est la seule maladie prédisposant aux CCR de transmission autosomique récessive. Le gène *MUTYH* appartient donc au système de réparation de l'ADN BER dont le rôle est de corriger les erreurs liées au métabolisme aérobies des cellules lors de la réplication (*Base Excision Repair*) avec *hOGG1* et *hMTH1*. Le métabolisme aérobie aboutit à la formation de radicaux dérivés de l'oxygène ($O_2^{\bullet -}$, $OH^{\bullet}$ ou H_2O_2) qui peuvent induire des lésions oxydatives au sein de l'ADN. Le 7,8-dihydro-8-oxoguanine (également appelé 8oxoG) est la lésion la plus fréquente touchant les purines. Le 8oxoG est un produit stable qui possède les capacités de s'apparier à la cytosine mais également avec prédilection à l'adénosine, provoquant ainsi une transversion G:C → T:A après deux réplications (Figure 11).

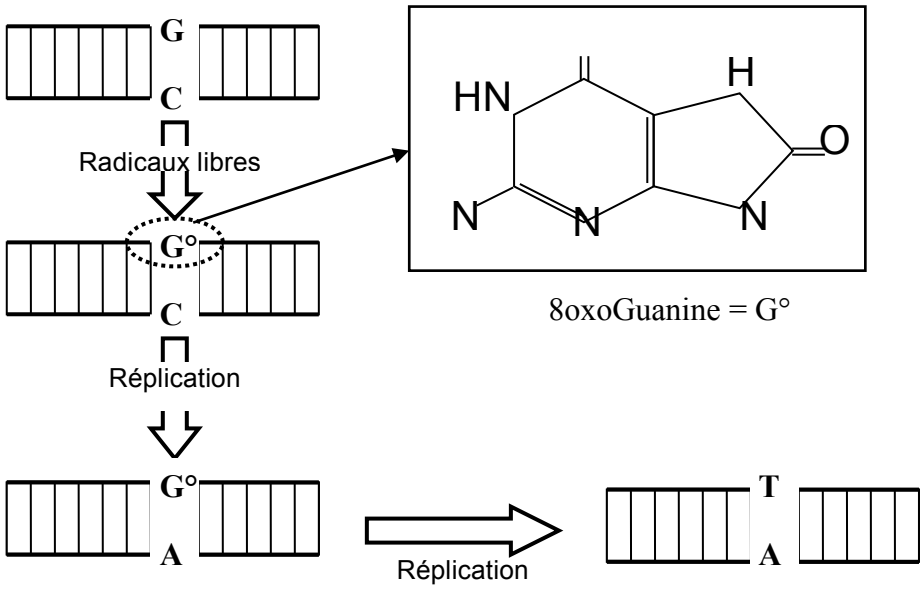


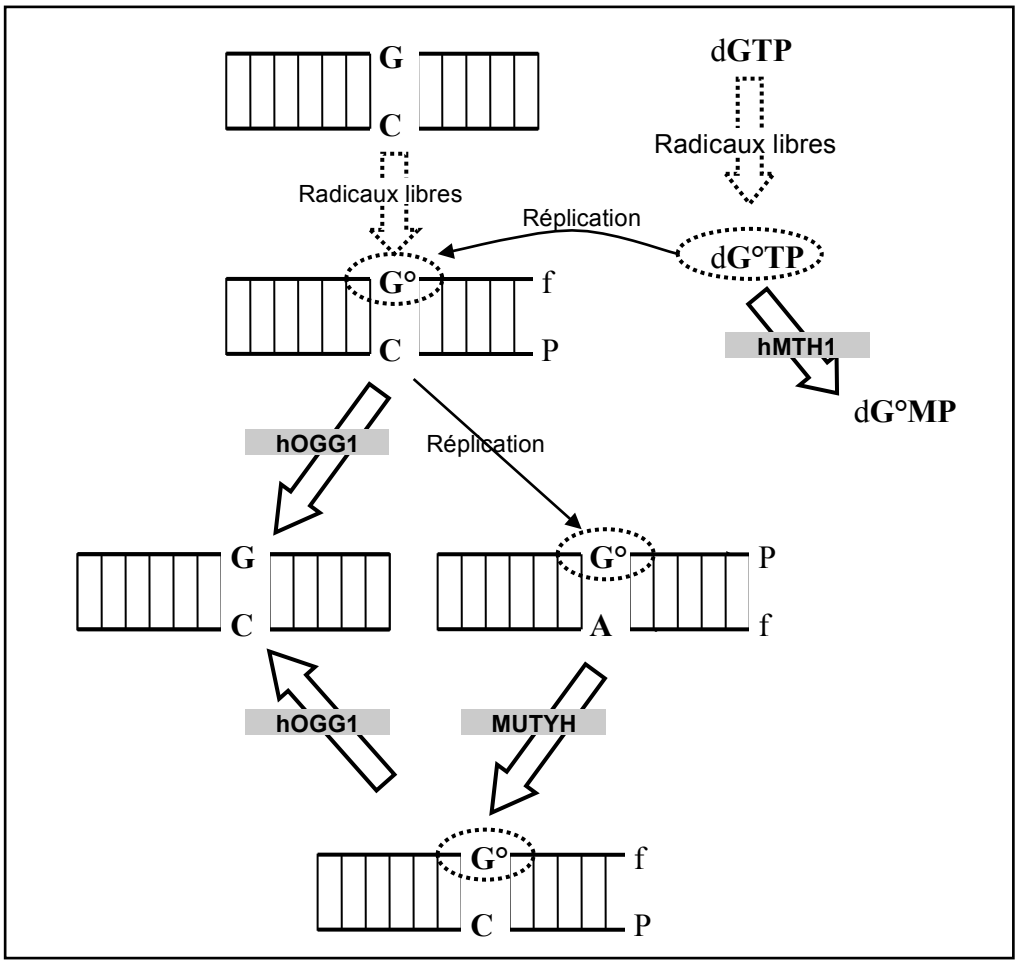
Figure 11. Acquisition d'une transversion G:C→T:A.

Le système BER constitue trois lignes de défense contre ces lésions oxydatives :

- MTH_1 constitue la première ligne de défense du système BER. Localisé dans le noyau, les mitochondries et le cytoplasme (Kang, D *et al.* 1995), il enlève du pool nucléotidique les dG°TP en les hydrolysant.
- Quand MTH_1 n'a pas hydrolysé tous les dG°TP, OGG_1 enlève les 8oxoG incorporés sur le brin fille d'ADN en face d'une adénine. Il permet ensuite l'incorporation d'une guanine. $hOGG_1$ est situé dans le noyau et les mitochondries.
- La dernière ligne de défense avant la seconde réplication qui aboutirait à la transversion G:C→T:A est MUTYH. Cette enzyme a pour substrat les adénines ou les 2-hydroxyadenine (2-OH-A) appariées avec une guanine ou un 8oxoG (Slupska, MM *et al.* 1996; (Takao, M *et al.* 1999).

L'ensemble des réactions permettant l'excision d'un 8oxoG est représenté sur la figure 12.

Le syndrome MAP est de transmission autosomique récessive avec mutation biallélique des allèles de *MUTYH* à la naissance. L'absence d'activité de *MUTYH* va entrainer l'accumulation de transversion G:C→T:A sur le gène *APC* au niveau des séquences GAA (Jones, S *et al.* 2002). *APC* est touché avec prédilection car il présente 216 sites GAA où une telle transversion peut conduire à un codon stop (Cheadle, JP *et al.* 2003; Al-Tassan, N *et al.* 2004). Ce sont ces transversions qui vont finir par inactiver somatiquement *APC* conduisant à l'apparition de polypes puis de CCR. La similitude des symptômes entre la PAF et le syndrome MAP vient de l'atteinte commune d'*APC* de manière germinale et somatique pour la PAF et purement acquise pour le syndrome MAP.



P : brin mère, f : brin fille, G° : 8oxoG

Figure 12. Mécanismes d'action du système BER.

Le mécanisme aboutissant à l'apparition du cancer en cas de mutation bi-allélique de *MUTYH* semble différent des mécanismes déjà connus et donc spécifique. Les cellules cancéreuses sont le plus souvent diploïdes (Lipton, L *et al.* 2003) et seule une étude a identifié un phénotype MSI lié à une méthylation du promoteur de *MLH1* (Colebatch, A *et al.* 2006). Les cancers liés au syndrome MAP sont donc CIN-/MSI-. Enfin, la présence d'une transversion sur *k-ras* est retrouvée plus fréquemment en cas de CCR lié à une mutation biallélique sur *MUTYH* qu'en cas de CCR sporadique.

2.3.3. Génétique du syndrome MAP, effet fondateur :

Il existe actuellement plus de 105 mutation décrites sur le site www.lovd.nl/mutyh en utilisant la séquence la plus longue NM 001128425.1 (Nielsen, M *et al.* 2011). Les mutations les plus fréquemment observées dans la littérature sont Y165C et G382D (nouvellement dénommées Y179C et G396D) (Buecher, B *et al.* 2012). Ces deux mutations ont une fréquence d'environ 1% dans les populations caucasiennes (Sieber, OM *et al.* 2003) et représentent environ 70% des patients ayant un syndrome MAP.

Cependant, des différences ethniques et géographiques ont été observes dans les quelques séries étudiant des patients non caucasiens: E480X est fréquente chez les patients indiens (Sampson, JR *et al.* 2003), Y179C et G396D n'ont jamais été identifiées chez les patients japonais (Miyaki, M *et al.* 2005). Enfin la mutation c.1227_1228dup (p.Glu396-GlyfsX43) mutation semble rare et a été décrite chez quelques patients portugais (Isidro, G *et al.* 2004). De surcroit, cette mutation n'a pas été trouvée dans une large série multicentriques avec 114 patients provenant des Etats-Unis, d'Australie ou du Canada (Cleary, SP *et al.* 2009). Il semble donc que la fréquence des mutations varie en fonction des origines géographiques des patients. La grande majorité des publications provenant de pays d'Europe de l'ouest ou d'Amérique du Nord, la prévalence des mutations au sein des autres populations reste inconnue (Sampson, JR *et al.* 2009).

2.3.4. Expression clinique de la MAP :

La principale manifestation est l'apparition d'une polypose adénomateuses de type atténuée : le plus souvent le nombre de polypes est inférieur à 100 et l'âge de découverte est plus tardif. Il existe une hétérogénéité intra et inter-familiale pour la sévérité de la polypose. Du fait du mode de transmission récessif, il n'existe pas d'agrégation sur plusieurs générations successives mais les fratries ont un risque de récurrence de 25%. Les cas isolés apparemment sporadiques sont donc fréquents principalement au sein des familles de petite

taille. Il existe comme dans la PAF des manifestations extra-coliques : principalement les polyposes duodénales (entre 10 et 20%) (Sampson, JR *et al.* 2009) et de rares cas de cancers dermatologiques (adénocarcinomes sébacés). Une expertise vient d'être publiée récemment par l'INCA sur la polypose liée à MUTYH (Collectif. 2011; Buecher, B *et al.* 2012).

La découverte de cette nouvelle polypose a permis d'ouvrir un nouveau champ de recherches sur sa prévalence au sein des polyposes non liées à APC, la véritable existence de la polypose atténuée avec la corrélation génotype-phénotype et les caractéristiques cliniques et biologiques des patients présentant une mutation biallélique sur MUTYH. Ces différents aspects font l'objet de la première partie de cette thèse de science (Articles A, B, C et D).

3. Polyposes indéterminées et cancer colorectal familial :

En marge de ces formes héréditaires bien définies cliniquement et génétiquement, on estime à 25% les formes de CCR héritées (Grady, WM 2003). Les principales formes de CCR sont représentées sur la figure 13 (Jasperson, KW *et al.* 2010).

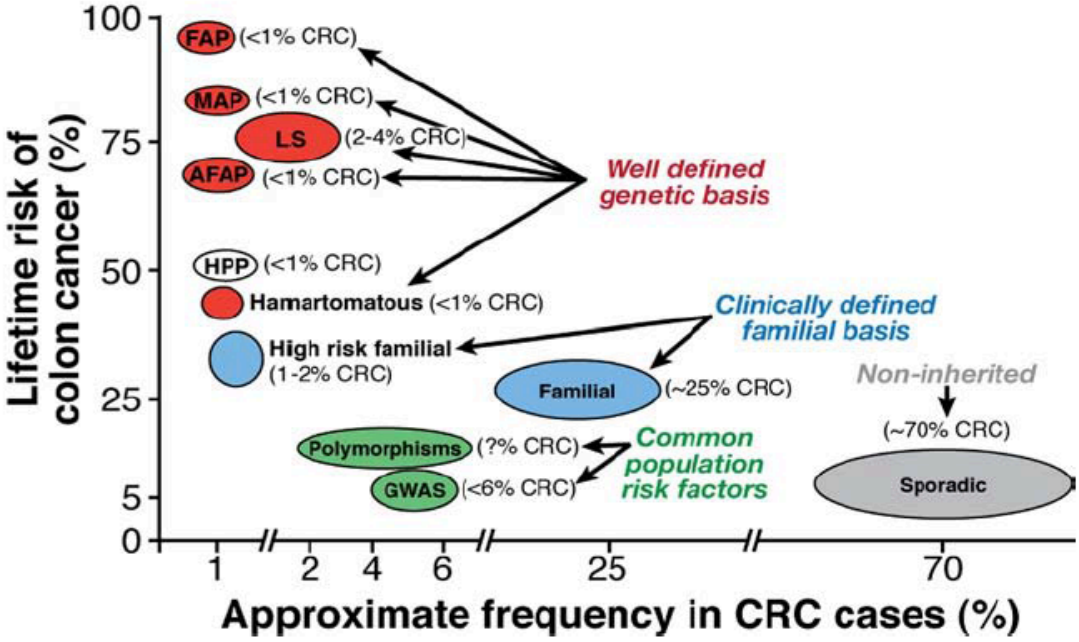


Figure 13. Répartition des cancers colorectaux selon leur origine génétique d'après (Jasperson, KW et al. 2010).

3.1. Familles à haut risque, Syndrome X :

Le syndrome HNPCC correspond à deux entités : le syndrome de Lynch (avec une mutation germinale prouvée sur un des gènes du système MMR) et le syndrome X (familles remplissant les critères d'Amsterdam mais n'ayant pas de mutation identifiée). En effet entre 1 et 2% des patients avec un CCR remplissent les critères d'Amsterdam II (Kerber, RA *et al.* 2005) (au moins 3 cas avec un cas de CCR prouvé liés par un membre au premier degré, sans polypose adénomateuses et deux générations successives touchées). Cependant au sein de ce groupe, une mutation germinale sur un gène du système MMR n'est retrouvée que dans la moitié des cas (Lindor, NM *et al.* 2005; Lindor, NM 2009). Cette forme familiale de cancer colorectal type X avec une stabilité des microsatellites entraine un risque de CCR plus faible que les syndromes de Lynch prouvés et est diagnostiquée en moyenne 10 ans plus tard. Enfin, il n'est pas encore décrit de cancers extra-coliques. De ce fait, leur surveillance pourrait être plus espacée. Aucune explication génétique n'a été pour le moment trouvée pour expliquer ce syndrome d'apparence autosomique dominant avec une pénétrance incomplète.

3.2. Formes familiales communes :

Environ 20% des cas de CCR surviennent dans des familles ayant un sur-risque : un cas de cancer diagnostiqué avant 50 ans ou un antécédent au premier degré de CCR (Kerber, RA *et al.* 2005). De plus avoir un parent ayant eu un CCR avant 45 ans ou 2 antécédents de CCR parmi les membre de sa famille augmente le risque de CCR d'un facteur compris entre 3 et 6 (Johns, LE *et al.* 2001). Plusieurs régions chromosomiques ont été identifiées dans les études familiales comme potentiellement impliquées dans ce sur-risque : 7q31, 9q22.33, 3q21-24, 11q23 (Wiesner, GL *et al.* 2003; Djureinovic, T *et al.* 2006; Kemp, Z *et al.* 2006; Neklason, DW *et al.* 2008). Ces résultats confortent l'idée que ces formes familiales sont secondaires à des mutations sur des gènes de faible pénétrance. Ces gènes intermédiaires

représentent vraisemblablement une proportion faible des CCR car ils n'ont pas été retrouvés sur les études de GWAS (*Genome Wide Association Studies*) qui détectent les mutations fréquentes (variants communs avec une fréquence > 5%) (Manolio, TA 2010).

3.3. Autres polyposes

La polypose juvénile (PJ) est un syndrome de transmission autosomique dominante comportant la présence de multiples polypes hamartomateux dans le colon et le rectum. On parle de PJ en cas de 5 ou plus polypes juvéniles dans le colon ou le rectum, en cas de polypes retrouvés tout le long du tractus digestif ou en cas de polype juvénile associé à un antécédent familial de PJ (Brosens, LA *et al.* 2011). Alors qu'un polype juvénile isolé ne semble pas associé à un risque accru de CCR, les PJ augmentent le risque de CCR et de cancer gastrique (risque compris entre 20 et 38% pour les CCR et aux alentours de 20% pour les cancers gastriques) (Brosens, LA *et al.* 2007).

La PJ est liée à des mutations du gène *SMAD4* ou du gène *BMPR1A*, deux chainons de la voie du TGFB. Une mutation germinale est identifiée dans 50-60% des cas (Aretz, S *et al.* 2007; (Brosens, LA *et al.* 2011). Les patients atteints du syndrome de Rendu-Osler, parfois lié également à une mutation sur *SMAD4*, sont à risque de développer une PJ (Gallione, CJ *et al.* 2004).

Le syndrome de Peutz-Jeghers (MIM175200) est également une affection de transmission autosomique dominante qui se caractérise par l'existence de polypes multiples avec une pigmentation cutanéo-muqueuse notamment aux pourtours des lèvres. Il est lié à une mutation germinale sur le gène *STK11* (Beggs, AD *et al.* 2010). Les polypes sont également hamartomateux mais contrairement aux polypes juvéniles, ils présentent une couche musculaire lisse au niveaux de la lamina propria (Jass, JR *et al.* 1988). Ces polypes peuvent être retrouvés tout le long du tractus digestif mais ils prédominent dans l'intestin grêle (60-

90%) et dans le colon (50-64%). La mutation causale est identifiée dans plus de 85% des cas (Volikos, E *et al.* 2006; (de Leng, WW *et al.* 2007).

3.4. Polyposes indéterminées

Après recherche d'une mutation germinale sur *APC* et d'une mutation biallélique sur *MUTYH*, environ 20% des patients présentant une polypose adénomateuse n'ont aucune explication génétique identifiée (Half, E *et al.* 2009). Il s'agit le plus souvent de polyposes modérées avec une centaine de polypes et de rares affections extra-coliques. Le mosaïcisme germinal explique quelques cas de polyposes et quelques cas de mutations sur *BMPR1A* ont été rapportés (Lipton, L *et al.* 2003).

Les PJ sans mutation identifiée (30-40%) ont fait également l'objet d'exploration. L'ensemble des gènes de la voie TGF a été séquencé et le gène *ENG* codant pour l'endogline a été trouvé muté dans deux cas (Sweet, K *et al.* 2005). Quelques cas de syndrome de Cowden avec une mutation sur le gène *PTEN* pourraient expliquer une partie de ces PJ sans cause étant donné que l'apparition de polypes juvéniles fait partie de ce syndrome.

Ces cas de polyposes indéterminées et plus généralement les cancers colorectaux précoces sont probablement les meilleurs candidats à des explorations génétiques complémentaires.

4. Variants et polymorphismes :

4.1. Variants communs :

Les SNP (*Single Nucleoid Polymorphism*) sont des variations du code génétique fréquentes dans la population générale (>5%). Elles sont les cibles d'études des GWAS et actuellement plus de 600 études de GWAS recherchant des associations entre des SNP et 150 pathologies différentes ont déjà été publiées (Manolio, TA 2010). Les odd-ratio calculés pour ces variants communs sont dans la majorité des cas compris entre 1,2 et 1,5 (Bodmer, W *et al.*

2008). Etant donné que plusieurs milliers de SNP sont étudiés lors d'une GWAS, des études de réplication des résultats sont nécessaires pour conformer l'implication potentielle d'un variant commun avec une pathologie. De plus les seuils de significativité statistiques sont très faible (environ 10^{-7}) pour éviter les faux positifs.

Quatorze variants ont été déjà identifiés grâce à des GWAS concernant le CCR. Les calculs statistiques estiment que 60 à 70 variants communs pourraient expliquer environ 17% des formes familiales de cancer colo-rectal (Peters, U *et al.* 2011). Une méta-analyse récente a combiné les résultats publiés (2906 cas vs. 3416 témoins) et a testé les polymorphismes les plus significatifs dans une analyse de réplication (combinant 10 cohortes indépendantes) sur 8161 cas et 9101 contrôles. Aucun des polymorphismes n'atteignait le seuil de significativité après ajustement des calculs multiples.

Un variant commun du récepteur du TGFB (*TGFBR1*), chainon majeur de la voie du TGF, a fait l'objet de nombreuses études et serait associé à un sur-risque de CCR (Daley, D *et al.* 2007). Ce variant commun (TGFBR1*6A/9A, rs11466445) diminuerait la capacité du récepteur à transmettre le signal du TGFB1.

4.2. Variants rares :

Ce sont également des variations du code génétique ayant une fréquence inférieure à 1% dans la population générale. Elles ne sont pas pour autant considérées comme des mutations délétères (dont la fréquence doit être inférieure à 0,1%). Les études GWAS ne peuvent pas encore étudier de tels variants. Pour les identifier, il faut comparer leur fréquence entre une population de cas et un large groupe de témoins. La difficulté de ces études est liée au choix des variants à étudier. Idéalement, il faut choisir des gènes dont l'altération conduit à une forme, le plus souvent clairement héréditaire, de la maladie étudiée ou des gènes connus pour être impliqués dans les processus biologiques, physiologiques ou voies de signalisation de la pathologie d'intérêt. Les variants du gène *APC*, I1307K présent dans les populations

Ashkénazes ou E1317Q augmentent significativement le risque de polypes coliques et de CCR (Laken, SJ *et al.* 1997; Frayling, IM *et al.* 1998; Woodage, T *et al.* 1998). I1307K provoque une substitution d'acide aminé (I-K (Isoleucine en Lysine) au sein d'une région impliquée dans la fixation protéique avec un effet dominant négatif. Le variant E1317Q affecte également la fonction d'APC et serait associée à un avantage sur la croissance tumorale (Frayling, IM *et al.* 1998).

A la suite de ces observations d'autres variants ont été testés. Les gènes candidats étaient sélectionnés le plus souvent en raison de leur implication dans les CCR héréditaires ou la survenue d'adénomes.

4.2.1. Théorie des variants rares :

L'équipe du Pr Bodmer a explicité la théorie des variants rares qui propose que certaines pathologies héréditaires ou susceptibilité à une affection soient dues à une accumulations de variants rares agissant indépendamment et qui seuls ne peuvent pas être responsables de la maladie.

4.2.2. Une combinaison de variants rares augmente le risque de polypose :

Cette équipe a d'ailleurs publié une étude comparant 124 patients avec une polypose indéterminée (moins de 100 polypes) sans argument pour un syndrome HNPPC, une PAF ou une MAP. Ils ont étudié 13 variants appartenant à 5 gènes d'intérêt (*APC, CTNNB1, AXIN1, hMLH1* et *hMSH2*) (Fearnhead, NS *et al.* 2004). Certains variants étaient significativement plus fréquents chez les cas et la combinaison de ces variants conférait un sur-risque de 2,2 (tableau 1).

Gene	Variants	Patients (%)	Controls (%)	Odds ratio	*P*
APC	*E1317Q*	3/124 (2.4)	6/480 (1.3)	2.0	0.400
CTNNB1	*N287S*	1/124 (0.8)	3/483 (0.6)	1.32	1.000
AXIN1	*P312T, R398H, L445M, D545E, G700S, R891Q*	18/124 (14.5)	37/479.3* (7.7)	2.0	0.012
hMLH1	*G22A, K618A*	5/124 (4.0)	11/482.5* (2.3)	2.0	0.175
hMSH2	*H46Q, E808X, ex4SDS*	3/124 (2.4)	0/479.3 (0.0)	11.7	0.001
	Combined	30/124 (24.9)	55/479.8*† (11.5)	2.2	0.0001

SDS, splice donor site variant.
*Average number of controls typed for all variants.
†Two individuals had two variants.

Tableau 1. Comparaison des variants rares entre 124 polyposes indéterminées et 483 contrôles (d'après Fearnhead et al.)

Depuis cette publication, plusieurs variants de gènes du système MMR (*hMLH1* et *hMSH6*) ont été identifiés comme augmentant le risque de CCR sans pour autant être responsables d'un syndrome de Lynch (Allan, JM *et al.* 2008; Tulupova, E *et al.* 2008). Dans une étude portant sur 1518 CCR, le variant MLH1 -93G>A (rs1800734) augmentait d'un facteur 1,68 (IC95% (1-2,83), p=0,05) le risque de CCR. Contrairement aux mutations germinales responsables du syndrome le Lynch, ce variant n'entrainait pas obligatoire de perte d'expression de la protéine MLH1 ou de phénotype MSI. Il était cependant significativement associé avec ces deux anomalies (Allan, JM *et al.* 2008). Dans une autre étude cas-témoins (614 CCR vs. 614 contrôles), deux polymorphismes du gène hMSH6 (-556G>T (rs3136228) et Ex1 – 145G>A (rs1042821)) modifiaient significativement le risque de CCR. Les porteurs d'une haplotype GT ou TT avaient un risque augmenté de 1,29 (IC95%=(1,02-1,62) ; p=0,04) et les porteurs de l'haplotype GA ou AA étaient au contraire protégés (OR=0,76, IC95%=(0,6-0,98) ; p=0,03) (Tulupova, E *et al.* 2008).

A coté des gènes du système MMR ou très liés au CCR, d'autres gènes ont été étudiés. Le variant 1100delC du gène *CHEK2* est considéré comme augmentant le risque de cancer du sein. Il a été trouvé avec une fréquence quatre fois plus important au sein d'une population

présentant un CCR en rapport avec le syndrome HNPCC (mutation identifiée ou critères cliniques) : 4,3% vs. 1%. Parmi les 10 familles présentant le variant et un CCR, 8 étaient MSS sans mutation germinale identifiée sur les gènes MMR (Wasielewski, M *et al.* 2008).

4.2.3. La cycline D1:

De même, la cycline D1 est une protéine codée sur le chromosome 11 au locus 11q13 et régule le cycle cellulaire de la phase G1 à la phase S (Howe, D *et al.* 2001). Une expression accrue de cycline D1 est un événement observé précocement dans la cancérogenèse colorectale mais également dans d'autres types histologique de cancers (sein, prostate, poumon…).

Le gène de la cycline D1 présente un variant commun qui a fait l'objet de nombreuses publications et méta-analyse. Ce polymorphisme G870A (P241P, rs9344) ne provoque pas de modification de la séquence aminée (Proline-Proline) mais affecte l'épissage alternatif à la fin de l'exon 4. Le transcrit alternatif (dénommé transcrit b) est déficient sur un motif PEST qui est impliqué dans la dégradation protéique (Pabalan, N *et al.* 2008). En cas de variant AA, le transcrit b a une demie-vie plus longue que le transcrit normal a. Les patients avec de nombreuses copies de CCND1-870A sont donc plus susceptibles de ponter le checkpoint G1-S ce qui contribue au développement du cancer (Solomon, DA *et al.* 2003). Une méta-analyse récente de 60 études cas-témoins a mis en évidence que les individus avec le génotype GA ou AA avaient un risque accru de cancer (quel que soit son origine) d'un facteur 1,1-1,2 par rapport aux individus avec le génotype GG (Pabalan, N *et al.* 2008). Pour les cancers colorectaux, plusieurs études ont observé chez les patients avec un SL, un risque de précocité de la maladie chez les porteurs du variant rs9344. Dans une série de 85 patients porteurs d'une mutation germinale prouvée sur *MLH1* ou *MSH2*, l'âge médian de diagnostique du CCR était de 37 ans pour le génotype AA et de 48 ans pour le génotype GG (Kong, S *et al.*

2000). La même équipe a ensuite étudié une population de 321 CCR hors PAF ou Peutz-Jeghers (54% survenus avant 50 ans, 23% avant 45 ans) (Kong, S *et al.* 2001). Ils ont comparé les 156 CCR provenant de patients caucasiens avant 60 ans à un groupe contrôle (=152). La prévalence de l'allèle A était plus élevée dans le groupe de CCR (54% vs 43%, p=0,005). Le génotype AA conférait un risque de 2,58 (IC95% : 1,48-4,52) par rapport aux génotypes GG/GA (p=0,001).

Le rôle des variants rares du gène de la cycline D1 n'ont pas encore été étudié, de même que leur impact par rapport au variant rs9344.

5. Hypothèses et but du travail :

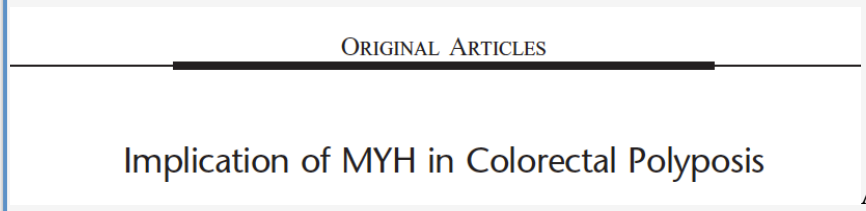
Environ 70% des polyposes adénomateuses sont liées à une mutation germinale sur le gène *APC*. Les 30% restant n'ont donc aucune explication pour leurs symptômes et ces patients ainsi que leur famille doivent se soumettre à une surveillance régulière. La découverte du syndrome MAP est une nouvelle piste de recherche pour trouver une étiologie aux polyposes indéterminées et étudier ses propriétés.

- Nous avons donc voulu étudier la fréquence des mutations sur *MUTYH* au sein d'une population de patients ayant une polypose adénomateuse sans mutation sur *APC*.
- La découverte de cette nouvelle polypose a permis ensuite de réaliser une large étude sur la corrélation génotype-phénotype au sein des polyposes adénomateuses.
- Nous avons étudié le cas d'une famille présentant une mutation germinale de *MUTYH* ainsi qu'une instabilité des microsatellites pour étudier la relation entre ces deux systèmes de réparation de l'ADN.
- Par la suite nous nous sommes intéressés à la répartition des différentes mutations de *MUTYH* au sein de notre population de patients ayant un syndrome MAP

Pour les patients n'ayant pas de mutation sur *APC* ou sur *MUTYH* et plus généralement pour les cas de cancers colorectaux survenus chez des patients jeunes (avant 50 ans), la théorie des variants rares pourrait représenter une explication de ces formes particulières de CCR.

- Nous avons donc voulu étudier la fréquence des variants rares au sein d'une cohorte de patients avec une polypose indéterminée ou un cancer colorectal diagnostiqué avant 50 ans :
    - En étudiant spécifiquement le gène de la cycline D1.
    - Puis en généralisant l'analyse à 70 variants provenant de 13 gènes différents.

IMPLICATION DE MUTYH DANS LES POLYPOSES ADENOMATEUSES


ORIGINAL ARTICLES

Implication of MYH in Colorectal Polyposis

Annexe A, page 152.

BUT

Etudier la fréquence des mutations biallélicques sur *MUTYH* au sein d'une cohorte de patients ayant une polypose adénomateuses familiale sans mutation germinale identifiée sur *APC*. Pour vérifier le caractère pathogène des mutations de *MUTYH*, une recherche de transversion sur *APC* et *k-ras* était également réalisée.

METHODES

1. Aspect médico-légal et sélection des patients

Cette étude était monocentrique, non thérapeutique, sans bénéfice individuel direct et le matériel utilisé était de l'ADN tumoral et germinal issu de pièces de résection colorectale. Elle ne relevait donc pas de la loi Huriet sur la recherche biomédicale. En revanche, elle nécessitait une autorisation de création d'un fichier de recherche médicale par la Commission Nationale de l'Informatique et des Libertés (CNIL). Cette démarche était tributaire d'une demande préalable d'avis auprès du Comité Consultatif sur le Traitement de l'Information en Matière de Recherche dans le Domaine de la Santé. Après avis écrit de ce comité, nous avons déposé une demande et obtenu un avis favorable de la CNIL. Fort de cet avis, le Conseil d'Administration de la Ville de Paris a autorisé la création d'un fichier sur le « traitement automatisé d'informations nominatives relatif à l'étude de l'existence d'altération du gène *MUTYH* chez les patients présentant une polypose colorectale atténuée, dans le service de chirurgie générale et digestive de l'hôpital Saint-Antoine ».

2. Sélections des patients :

L'analyse des dossiers s'est déroulée en trois phases :

- Le listing des patients ayant été opéré depuis 1995 dans le service de chirurgie digestive de l'hôpital Saint-Antoine. Les mots-clés utilisés ont été : colectomie totale, polypose, polypose atténuée, iléo-rectale, iléo-anale.
- La relecture des comptes-rendus opératoires de ces patients. Les dossiers des patients pouvant être inclus dans cette étude ont été alors récupérés.
- L'étude des dossiers a permis d'éliminer de l'étude :
    - Les polyposes adénomateuses familiales prouvées avec mutation congénitale du gène *APC*,
    - Les polyposes adénomateuses de transmission autosomique dominante,
    - Les syndromes cliniques typiques d'HNPCC, les patients ayant perdu en immuno-histochimie l'expression de la protéine hMLH1 ou hMSH2 (analyse réalisée systématiquement dans le service d'Anatomopathologie).
    - Les patients décédés,
    - Les patients pour lesquels aucune pièce anatomopathologique n'était disponible.

Les patients sélectionnés ont été contactés quand cela était possible. Tous les patients avaient un entretien téléphonique ou à l'hôpital (lors d'une consultation) au cours duquel leurs étaient expliqués, les buts, les modalités de l'étude et le bénéfice attendu. Après avoir répondu à leurs éventuelles interrogations et en cas d'accord, un formulaire d'information et de consentement éclairé en 3 exemplaires leur était remis pour signature.

Pour chaque patient ayant donné son accord écrit, les données suivantes étaient enregistrées dans le fichier MUTYH : le sexe, l'âge au moment du diagnostic, les antécédents personnels et familiaux de cancers, le site tumoral, le stade tumoral, le type

anatomopathologique, le nombre de polypes, l'existence d'autres localisations tumorales, le geste effectué. Pour le site tumoral, un cancer du colon droit était défini comme siégeant sur le caecum, le colon droit ou le colon transverse ; un cancer du colon gauche était défini comme siégeant sur le colon gauche, le colon sigmoïde ou le rectum.

3. Préparation des échantillons et extraction de l'ADN :

De l'ADN tumoral et non tumoral étaient extraits de tissus congelés à -80°C ou fixés en paraffine à partir des pièces de résection colique ou rectale. Pour chaque type de tissus, des coupes de 30 microns étaient réalisées. Avec l'aide d'un anatomopathologiste, des champs présentant une absence de cellule tumorale ou contenant plus de 70% de cellules tumorales étaient identifiés puis microdisséqués, permettant d'obtenir des échantillons de tissu non tumoral et tumoral (respectivement) d'environ 25 mg.

Les échantillons fixés dans le formol, et inclus en paraffine, étaient tout d'abord déparaffinés. Les échantillons étaient chauffés pendant 10 minutes à 60°C. Ils étaient ensuite placés dans un bain de 1200 mL de xylène suivie d'une centrifugation de 5 minutes à 13000 t/m. Cette étape était réalisée 3 fois permettant d'éliminer le surnageant chargé en paraffine. Puis le culot était lavé 5 fois dans 1200µL d'éthanol absolu et centrifugé 5 minutes à 13000 t/m afin d'éliminer les résidus de xylène. Enfin l'éthanol résiduel était évaporé en laissant les tubes ouverts à 37°C pendant 5 minutes.

L'ADN des 2 types d'échantillons était extrait à l'aide du kit DNeasy® (Qiagen, Inc, Santa Clarita, CA, USA) selon le protocole établit par le fabricant. La concentration de l'ADN élué à partir de chaque échantillon a été quantifiée à l'aide d'un spectrophotomètre (BioSpec-Mini®, Shimadzu), puis la qualité de l'ADN a été évaluée par migration sur gel à 0,8% d'agarose (Invitrogen Life Technologies®) contenant 5% de Bromure d'Ethidium (BET) (volume/volume) : Figure 14. L'ADN extrait était ensuite conservé à 4°C.

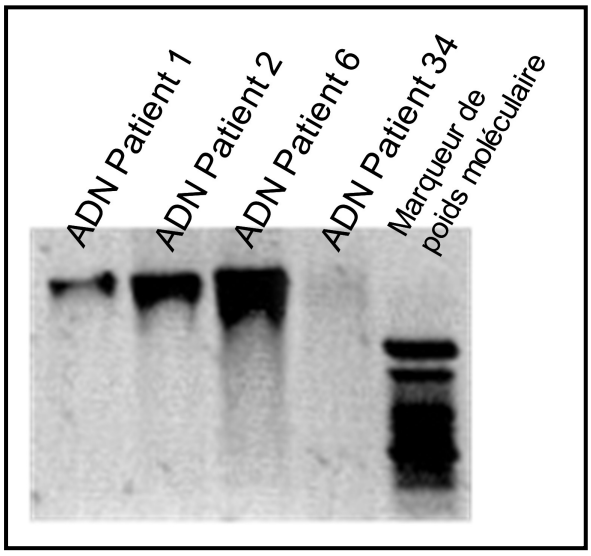


Figure 14. Dépôt des ADN après extraction (n°1-6: congelé, n°34: paraffine).

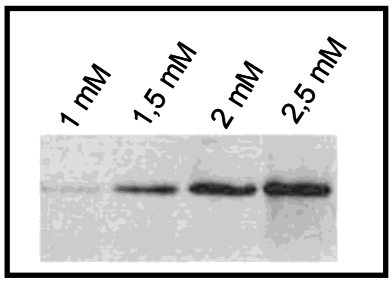
4. Amplification et séquençage des exons de MUTYH, K-ras et APC :

4.1. Préparation des PCR :

A partir des séquences publiées pour *MUTYH* (GenBank ID : 4595) par Al-Tassan et al. (Al-Tassan, N *et al.* 2002), des séquences publiées de *K-ras* (GenBank ID : 3845) et d'APC (GenBank ID : 324), nous avons fait synthétiser par la société Proligo® les amorces nécessaires à l'amplification des 16 exons de *MUTYH* (Tableau n°1), de l'exon 1 de *k-ras* (tableau 3) et de l'exon 15 d'*APC* (tableau 4). L'amplification de *MUTYH* était réalisée à partir de l'ADN issus de cellules non tumorales afin de repérer les mutations congénitales. L'amplification de *k-ras* et *APC* était réalisée sur de l'ADN tumoral et non tumoral afin de mettre en évidence les mutations somatiques.

Les conditions d'amplification optimales ont été déterminées. Pour cela, nous avons testés pour chaque couple d'amorces 4 concentrations de $MgCl_2$, 1 ; 1,5 ; 2 et 2,5 mM, et différentes températures comprises entre 4 et 8°C en dessous du Tm calculé (Tm= 4n+2y où n est le nombre de G et C et y le nombre A et T). La condition retenue correspond à celle permettant une amplification maximale sans amplicon aspécifique (Figure 15).

Figure 15. Tests d'amplification de l'exon 7 (133 pb).



4.2. Conditions d'amplification :

Le tableau 2 détaille les amorces et les conditions d'amplification du gène *MUTYH*.

Exon	Séquence	Fragment (bp)	T°C	$MgCl_2$ (mM)
MYH1	5'-GAAGCTGCGGGAGCTGAAA-3' 5'-ATCCCCGACTGCCTGAACC-3'	133	58	2
MYH 2	5'-CTGCATTTGGCTGGGTCTTT-3' 5'-CGCACCTGGCCCTTAGTAAG-3'	263	56	2
MYH 3	5'-AGCCTGTGCAGGGATGATTG-3' 5'-CAACCCCAGATGAGGAGTTAGG-3'	272	57	2
MYH 4	5'-CTCATCTGGGGTTGCATTGA-3' 5'-GGGTTGGCATGAGGACACTG-3'	167	52	2
MYH 5	5'-GGGCAGGTCAGCAGTGTC-3' 5'-TACACCCACCCCAAAGTAGA-3'	189	52	2
MYH 6	5'-TACTTTGGGGTGGGTGTAGA-3' 5'-AAGAGATCACCCGTCAGTCC-3'	185	54	2
MYH 7	5'-GGGACTGACGGGTGATCTCT-3' 5'-TTGGAGTGCAAGACTCAAGATT-3'	186	54	2
MYH 8	5'-CCAGGAGTCTTGGGTGTCTT-3' 5'-AGAGGGGCCAAAGAGTTAGC-3'	240	58	2
MYH 9	5'-AACTCTTTGGCCCCTCTGTG-3' 5'-GAAGGGAACACTGCTGTGAAG-3'	196	54	2
MYH 10	5'-GTGCTTCAGGGGTGTCTGC-3' 5'-TGTCATAGGGCAGAGTCACTCC-3'	262	60	2
MYH 11	5'-TAAGGAGTGACTCTGCCCTATG-3' 5'-GCCAAGAGGGGCTTTAGG-3'	248	54	2
MYH 12	5'-AGCCCCTCTTGGCTTGAGTA-3' 5'-TGCCGATTCCCTCCATTCT-3'	298	60	2
MYH 13	5'-AGGGCAGTGGCATGAGTAAC-3' 5'-GGCTATTCCGCTGCTCACTT-3'	242	54	2
MYH 14	5'-TTGGCTTTTGAGGCTATATCC-3' 5'-CATGTAGGAAACACAAGGAAGTA-3'	256	58	2
MYH 15	5'-TGAAGTTAAGGGCAGAACACC-3' 5'-GTTCACCCAGACATTCGTTAGT-3'	205	58	2
MYH 16	5'-AGGACAAGGAGAGGATTCTCTG-3' 5'-GGAATGGGGGCTTTCAGA-3'	224	60	2

Tableau 2. Amorces et conditions d'amplification de *MUTYH*.

Les amorces pour l'amplification du gène *k-ras* et de l'exon d'*APC* sont détaillées dans les tableaux 3 et 4.

Exon	Séquence	Fragment (bp)	T°C	$MgCl_2$ (mM)
Kras1	5'-GTACTGGTGGAGTATTTGAT -3' 5'-ACTCATGAAAATGGTCAG-3'	290	57	2

Tableau 3. Amorces et conditions d'amplification de l'exon 1 de *k-ras*.

Fragment	Séquence	Taille (pb)	T°C	$MgCl_2$ (mM)
APC15-1 : A1	5'-ACTGCATACATTGTGACC-3' 5'-CAGAACGAGAACTATCTAAGC-3'	631	52	2,5
APC 15-2 : A2	5'-CAACATGACTGTCCTTTCAC-3' 5'-CAGAGTTCAACTGCTCATCT-3'	668	55	2
APC 15-3 : A3	5'-ATGATGGAGAACTAGATACAC-3' 5'-CGTATTCTCACTGCTTGAAG-3'	608	52	2,5
APC 15-4 : A4	5'-TGCCACAGATATTCCTTCATC-3' 5'-GCCACTTACCATTCCACTGC-3'	675	57	2
APC 15-5 : A5	5'-CTTCTGTCAGTTCACTTGATAG-3' 5'-GCTTTACGTGATGACTTTGTTG-3'	707	57	2
APC 15-6 : A6	5'-TGATTCAGATGATGATGATATTG-3' 5'-GTGCAGAAGAAGACGCAGAT-3'	604	57	2
APC 15-7 : A7	5'-GTCACAAGCCTTTCCGTGTG-3' 5'-GGATTGTTTCTGAAGTATGGG-3'	586	55	2
APC 15-8 : A8	5'-ACCAACAATCAGCTAATAAGAC-3' 5'-GCTTGTCTAGATAAACATGCAG-3'	669	57	2
APC 15-9 : A9	5'-CAGATTCAGAACATGGTCTATC-3' 5'-GGCAACAGGGCTTAATTCTG-3'	604	57	2
APC15-10 : A10	5'-CTCCTAGAGGAGCCAAGCCATC-3' 5'-GTTTGTGCCTGGGACCTAGTG-3'	655	60	2
APC15-11 : A11	5'-CCATCATCTAGACCAGCTTC-3' 5'-GAAACACTGTCAATCACCGG-3'	662	57	2
APC15-12 : A12	5'-CAATCCTAGATCTGGAAGATCTC-3' 5'-CCCTCTAACAAGAATCAAACC-3'	657	57	2,5

Tableau 4. Amorces et conditions d'amplification de l'exon 15 de l'*APC*.

Les amplifications étaient réalisées dans le milieu réactionnel détaillé dans le tableau 5.

	2mM $MgCl_2$	2,5mM $MgCl_2$
Tampon Taq Gold 10X (Applied Biosystems®)	5 µL	5 µL
$MgCl_2$ 25 mM (Applied Biosystems®)	4 µL	5 µL
dNTP 20 mM (Invitrogen Life Technologies ®)	0,5 µL	0,5 µL
Amorce sens (50 pmol/mL)	1 µL	1 µL
Amorce anti-sens (50 pmol/mL)	1 µL	1 µL
Taq Gold 5 U/mL (Applied Biosystems®)	0,25 µL	0,25 µL
H_2O	37,25 µL	36,25 µL
ADN	1 µL	1 µL
Volume final	50 µL	50 µL

Tableau 5. Composition du milieu réactionnel des PCR.

Le programme était le suivant :

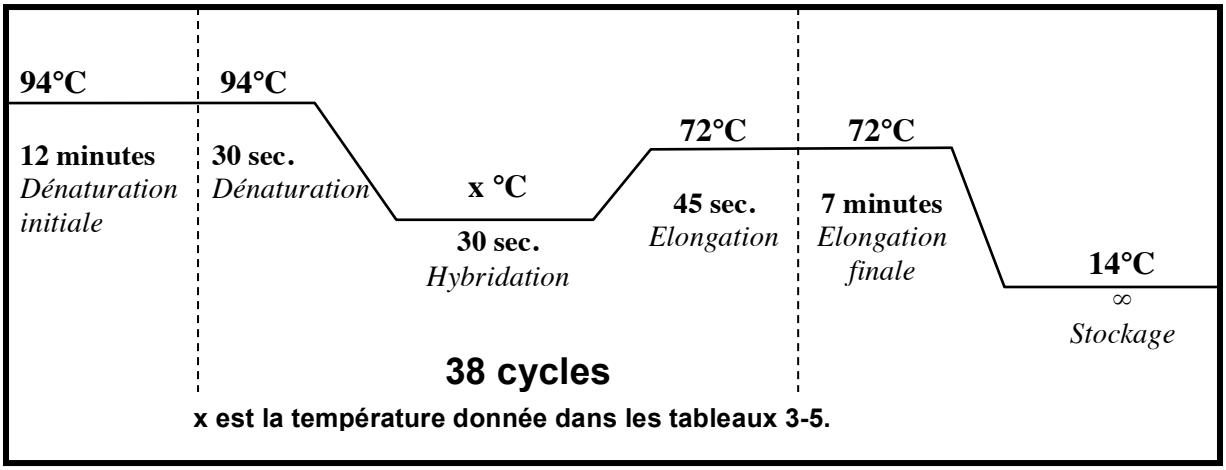


Figure 16. Détail des cycles de PCR.

Cinq µL de chaque produit d'amplification par PCR étaient déposés sur un gel d'agarose standard (BET 1,5 %). L'électrophorèse était réalisée à 120 Volts pendant 45 minutes dans du tampon TBE 1X (Invitrogen Life Technologies®). L'ADN amplifié était visualisé sous les rayons UV.

4.3. Purification des bandes d'intérêt :

L'ADN tumoral et non tumoral des pièces conservées dans de la paraffine étant de qualité moyenne (exemple patient 34, figure 14), il nous a fallu purifier les amplicons afin de permettre un séquençage optimal. Pour ce faire, nous avons déposé la totalité du produit d'amplification (45 mL) sur un gel d'agarose (BET 1,5%) et découpé les bandes d'intérêt sous les rayons UV en utilisant les protections adéquates (gants, masque). Chaque bande était placée dans un tube de 1,5 mL puis pesée.

La purification des bandes était réalisée à l'aide du kit QIAquick® (Qiagen, Inc, Santa Clarita, ÇA). On déposait 3 volumes de tampon QG (100 mg ≈ 100 mL). L'ensemble était incubé à 50°C pendant 10 minutes afin de dissoudre l'agarose. L'échantillon était vortexé toutes les 3 minutes. On rajoutait à cette solution un volume d'isopropanol et le tout était vortexé et déposé sur une colonne QIAquick® pour une centrifugation d'une minute à 13000 t/m. L'éluat était éliminé et on déposait 500 µL de tampon QG pour éliminer toute trace d'agarose sur la colonne. On centrifugeait à nouveau la préparation pendant 1 minute à 13000 t/m. La colonne était lavée par 750 µL de tampon PE et centrifugée pendant 1 minute à 13000 t/m deux fois de suite pour bien éliminer toute trace d'éthanol qui aurait pu interférer avec les réactions nécessaires au séquençage.

Pour éluer l'ADN de la colonne, on déposait 40 µL de tampon EB (10 mM Tris-HCl, pH 8,5) et on centrifugeait pendant 1 minute à 13000 t/m. Pour vérifier la pureté de l'ADN, 4 µL de l'ADN purifié étaient à nouveau déposés sur un gel d'agarose (BET 1,5 %) et visualisés sous les rayons UV (figure 17).

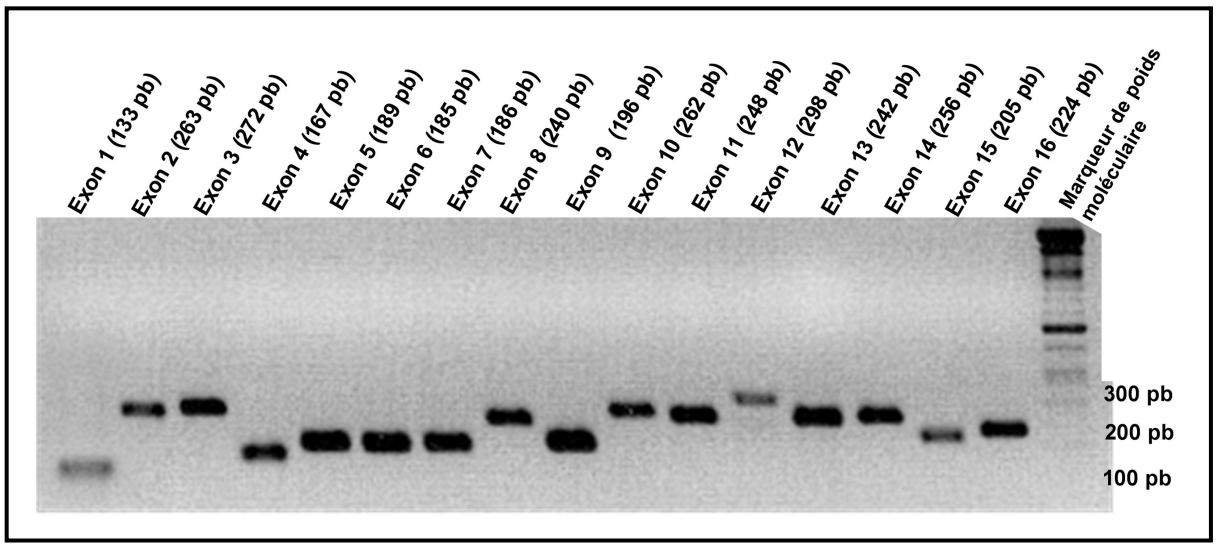


Figure 17. Exemple de profil de migration des produits d'amplification obtenus pour les 16 exons de *MUTYH* (patient n°2)

4.4. Séquençage et analyse :

Le séquençage double brin des ADN fut réalisé par la société MilleGen Biotechnologies® (Labège, France). Les séquences obtenues étaient analysées et comparées à la séquence normale en format texte grâce au logiciel FASTA version 3.4. Les chromatogrammes étaient lus avec le logiciel Chromas 2.3 (Conor McCarthy, 1996-2004, Griffith University, Australia).

RESULTATS :

1. Sélection des patients :

Le listing nous a permis de présélectionner 422 comptes-rendus opératoires. Après lecture de l'histoire de la maladie, des antécédents et du geste opératoire, 199 dossiers ont été sélectionnés. Nous avons éliminé de cette liste :

- 86 mutations congénitales prouvées du gène APC

- 38 phénotypes typiques PAF (transmission dominante) sans recherche du génotype sur APC (recherche non acceptée)
- 6 patients décédés
- 4 polyposes non adénomateuses (hamartomes, polyposes juvéniles (n=2), polypes avec prolifération kystique)
- 14 patients sans matériel (opérés pour la plupart dans un autre centre et suivis à l'Hôpital Saint-Antoine ou opérés en urgence pour une complication)
- 3 syndromes HNPCC.

Il restait donc 43 dossiers compatibles. Tous ces patients ont été contactés à l'exception de quatre patients (n°3, n°16, n°19 et n°22). Sur les 39 patients recontactés, 2 n'ont pas signé le formulaire (patients n°7 et n°42). Trente sept patients ont donc été inclus dans cette étude. Les caractéristiques cliniques de ces patients sont détaillées dans le tableau 6.

L'âge moyen des patients au moment de l'intervention chirurgicale était de 56,1 ans (DS : 8,3 ans, médiane : 57 ans, (26-80 ans). Le nombre moyen de polypes était de 52,2 (DS : 45, médiane : 30, (3-266). Vingt-huit patients (75,7%) avaient une polypose atténuée (moins de 100 polypes). Quarante cancers ont été découverts chez 24 patients (64,9%). Chez 12 patients (32,4%) les localisations étaient multiples. Ils étaient préférentiellement situés à gauche : 25 cancers (62,5%) (dont 7 cancers rectaux).

N°	sexe	DDN	Geste	Antécédent familial	Antécédent personnel	Age du diagnostic	Nombre polypes	Cancer	Localisation du cancer	TNM
1	M	23/6/57	ACR	Frère : adénome en dysplasie de haut grade à 40 ans		47	1	1	Côlon G	T3N1M1
2	M	6/3/62	AIA	Grand père maternel : CCR		42	156	3	Côlon Dt (3 fois)	pT3NOM0
4	M	27/8/44	AIR	Père : polypes Sœur : polypes	Polypes hyperplasiques	55	60	0		
5	M	15/1/34	AIA			67	70	1	Côlon G	T2N0M0
6	M	14/1/22	AIS			80	4	2	Côlon G + côlon Dt	T2N0M0
8	F	13/7/42	AIR		5 polypes lipomateux	61	> 15	2	Côlon G	T3N0M0
9	M	30/6/41	AIR	Sœur et nièce : polypose		60	> 100	0		
10	M	13/2/48	AIS			53	15	1	Côlon Dt	T3N0
11	M	13/5/52	AIS		Polypectomie avant	51	56	0		
12	F	23/5/45	AIA	Mère : polypes		59	127	1	Côlon Dt	T3N0M0
13	M	05/07/1945	ACR->AIR			56	11	2	Côlon G	T3N1M0
14	M	14/8/34	AIR			68	11	2	Côlon G+Côlon Dt	T4N1M0
15	M	10/10/36	AIS			64	17	4	Côlon Dt	T3N1M0
17	F	28/3/68	AIA	Polypes chez le grand père et le cousin paternel		34	> 200	0		
18	M	11/4/47	AIA	Polypose chez 2 frères (dont le patient n°43)		55	266	3	Rectum	T3N0M0
20	M	12/11/41	AIA	Père : cancer de l'estomac	Plusieurs polypectomies	60	15	1	Rectum	T1N0M0
21	M	1/1/39	AIA		HCEPR	60	> 30	2	Côlon G+rectum	T2N1M1
23	F	20/5/35	AIA			64	32	1	Rectum	T2N0M0
24	M	27/1/42	AIS		Plusieurs polypectomies	59	> 20	0		
25	M	14/12/53	ACR	Oncle		49	>40	0		
26	F	18/5/76	AIS	Polypose chez le père		22	13	2	Côlon G (2 fois)	T2N0M0

N°	sexe	DDN	Geste	Antécédent familial	Antécédent personnel	Age du diagnostic	Nombre polypes	Cancer	Localisation du cancer	TNM
27	M	24/8/36	iléostomie terminale		Cancer du rectum	67	4	1	Côlon G	T3N0M0
28	F	1/1/47	AIS			57	>100	0		
29	M	15/07/1944	AIA			52	50-100	1	Côlon G	Tis
30	F	8/6/42	AIS	CCR chez la grand-mère maternelle et la tante maternelle	Cancer du col de l'utérus	61	12	2	Côlon G + côlon Dt	T2N0M0
31	M	10/10/48	ACA->AIA			55	> 20	2	Côlon Dt + Côlon G	T2N0M0
32	F	2/9/52	AIR			50	32	1	Côlon Dt	T1N0M0
33	F	9/2/39	AIR			64	24	1	Côlon G	T1N0M0
34	M	29/12/42	AIS		Exérèse de 9 polypes	59	> 25	0		
35	M	29/3/59	AIA			43	40	0		
36	M	6/9/56	AIR	père et sœur : polypose		48	> 100	0		
37	M	2/5/32	AIS			71	5	2	Côlon G + Côlon Dt	T4N1M0
38	F	12/12/53	AIS			41	30	0		
39	M	6/5/50	AIR	Frère : polypose		53	> 70	0		
40	M	2/10/43	AIS			60	> 100	0		
41	M	30/12/24	AIR		Cancer du côlon G	78	7	1	Côlon G	T2N0M0
43	M	2/1/53	AIA	Polypose chez deux frères (dont le patient n°18)		49	multiples	1	Côlon G	T3N0M0

Tableau 6. Caractéristiques clinique et endoscopique des patients inclus pour une recherche de mutation sur *MUTYH*.

Trois patients présentaient des atteintes extra-coliques :

- le patient n°21 avait plusieurs foyers bilatéraux d'hypertrophie congénitale de l'épithélium pigmentaire de la rétine
- les patients n°2 et n°5 avaient plusieurs polypes duodénaux.

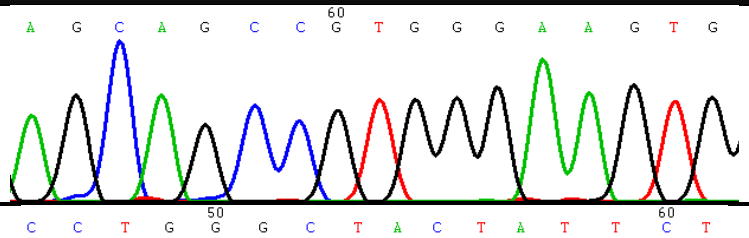
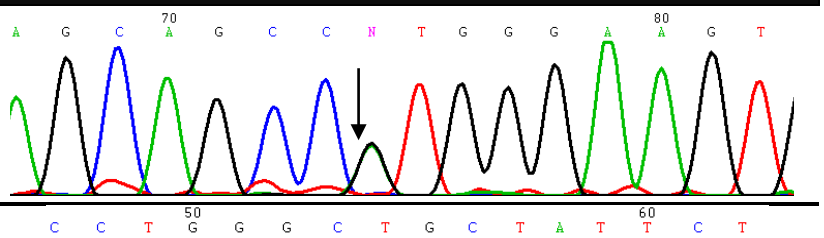
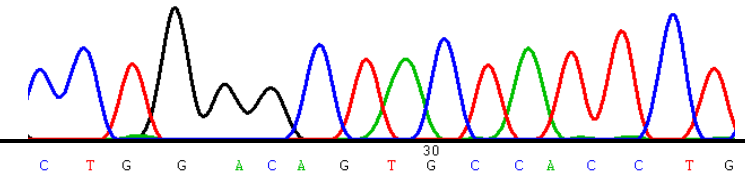
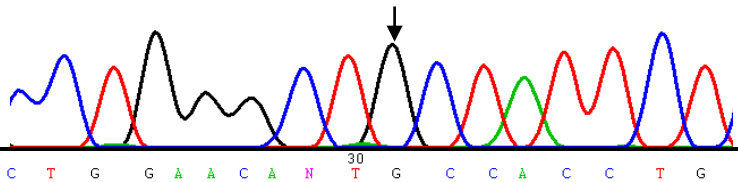
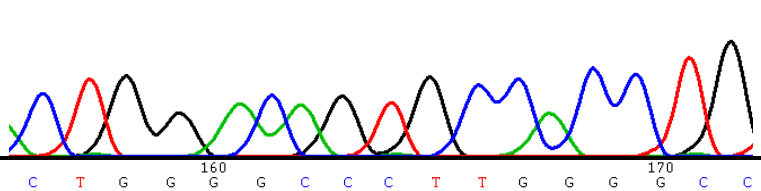
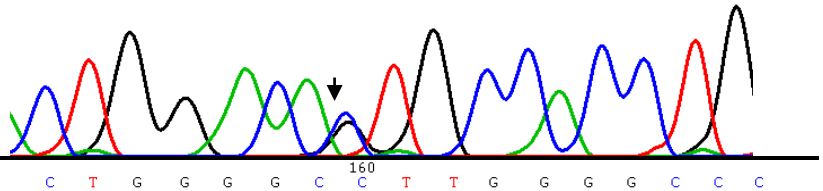
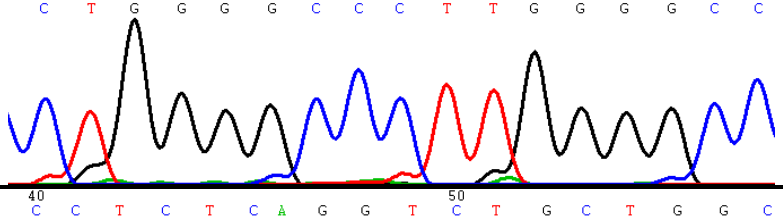
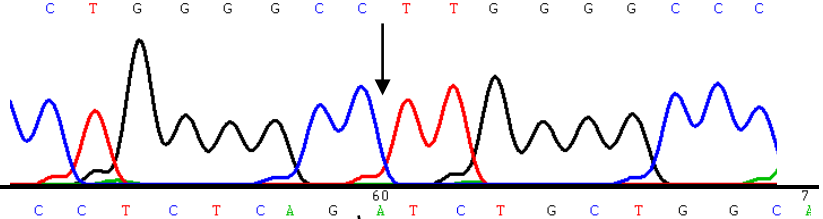
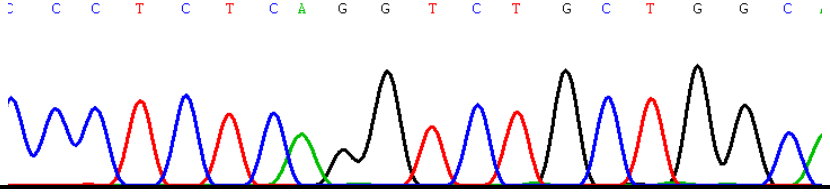
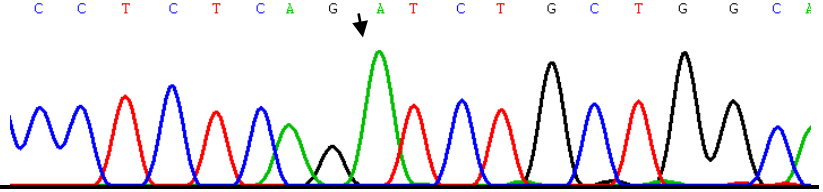
Le patient N°1 ne présentait pas de polypose mais son âge jeune, l'antécédent fraternel, les absences de mutation sur *APC* et de perte d'expression de hMLH1 et hMSH2 nous ont incités à l'inclure dans cette étude.

2. Extraction de l'ADN et réalisation des PCR :

Chez 11 patients (n°1,4, 5, 8, 9, 11, 12, 17, 27, 28 et 30), il n'a pas été possible d'obtenir un échantillon de tissu tumoral. Le service d'Anatomopathologie ne possédait que du tissu tumoral pour le patient n°29. Pour chaque patient nous avons réussi à obtenir une bande d'amplification pour le gène *MUTYH* à partir de l'ADN non tumoral (sauf pour le patient n°29) correspondant à la taille attendue. L'amplification de *K-ras* a été toujours possible. L'amplification du gène *APC* s'est révélée peu concluante. En effet nous avons obtenus de faible efficacité d'amplification pour cet exon. Ainsi, de nombreuses bandes d'amplification n'étaient pas suffisantes pour permettre un séquençage précis.

3. Etudes des mutations sur MUTYH :

Nous avons observés 26 mutations chez 18 patients (48,6%). Les mutations et leur fréquence sont détaillées dans le tableau 7. Six patients (16,2%) étaient homozygotes pour une mutation pathologique. Chez 12 patients (32,4%) nous avons retrouvé la présence d'un ou deux polymorphismes. L'ensemble des mutations est détaillé dans le tableau 8. Pour le patient n°29, la mutation a été trouvée sur de l'ADN tumoral et il n'a pas été possible d'obtenir un fragment d'ADN normal.

MUTATION	N (%)	SÉQUENCE NORMALE	SÉQUENCE MUTÉE
66 G>A V22M	1 (2,7%)		
494 A>G Y165C	4 (10,8%)		
972 G>C Q324H	11 (29,8%)		
1103 delC	2 (5,4%)		
1145 G>A G382D	2 (5,4%)		

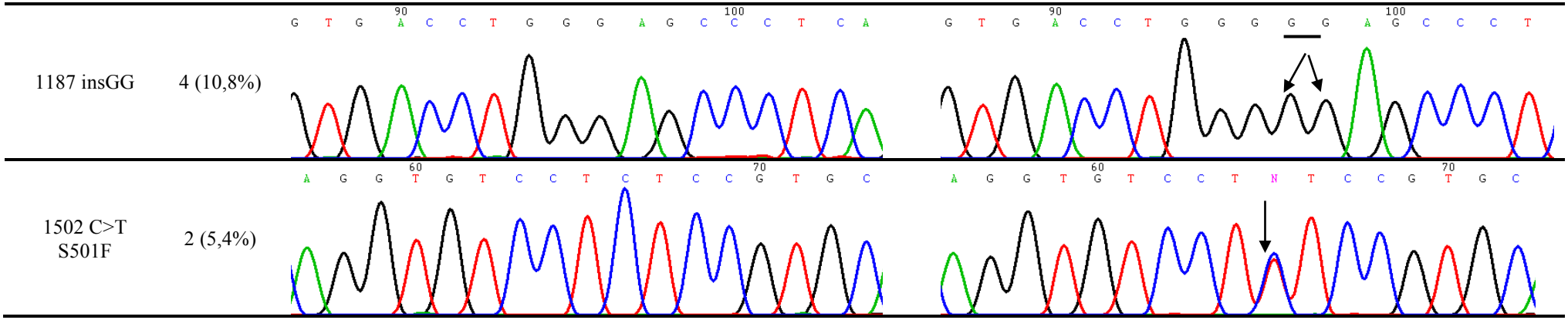


Tableau 7. Mutations identifiées sur *MUTYH*.

PATIENT	MUTATION 1	MUTATION 2
2	Y165C	Y165C
6	Q324H	
9	Q324H	
10	Q324H	
14	Q324H	
15	Q324H	S501F
18	1187 ins GG	1187 ins GG
21	1103 del C	1103 del C
23	Y165C	Y165C
24	Q324H	
28	Q324H	S501F
29	V22M	
34	Q324H	
35	G382D	G382D
38	Q324H	
40	Q324H	
41	Q324H	
43	1187 ins GG	1187 ins GG

Tableau 8. Détails des mutations biallélques et des polymorphismes sur *MUTYH*.

4. Etudes des mutations somatiques sur k-ras et APC:

Nous n'avons trouvé aucune mutation sur *K-ras* germinale sur les échantillons d'ADN non tumoral. Tous présentaient la séquence représentée sur la figure 18.

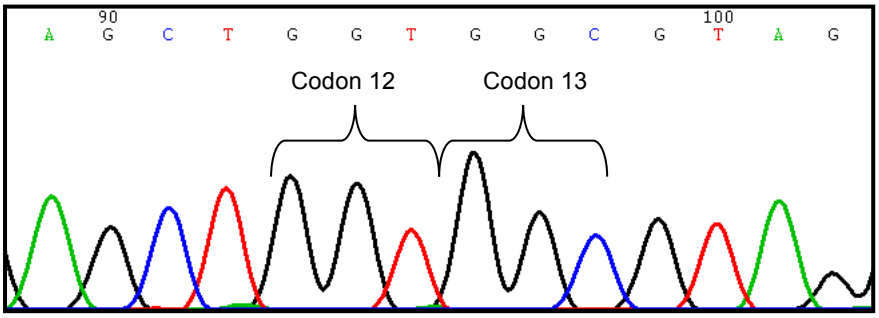


Figure 18. Chromatogramme normal de *K-ras*. (patient n°2)

Nous avons trouvé 7 mutations (26,9 %) parmi les 26 échantillons d'ADN tumoral dont nous disposions. Il s'agissait de 4 transversions G:C→T:A (50%) et 3 autres mutations.

Le tableau n°9 représente l'ensemble des patients mutés sur K-*ras* et leur statut pour *MUTYH*.

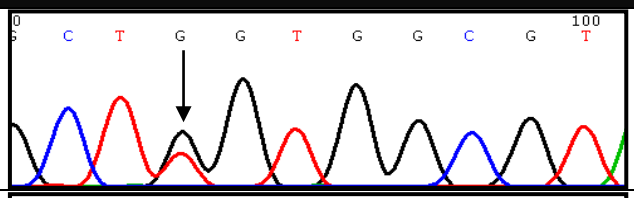
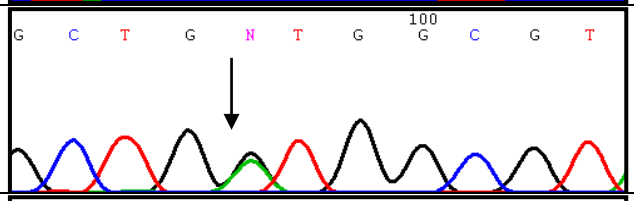
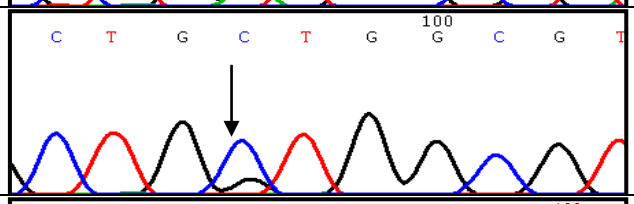
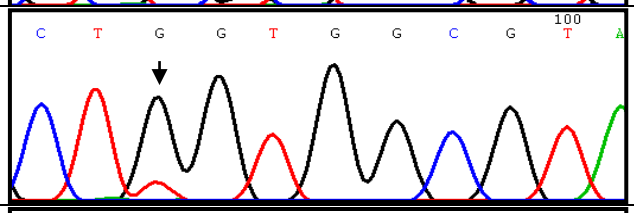
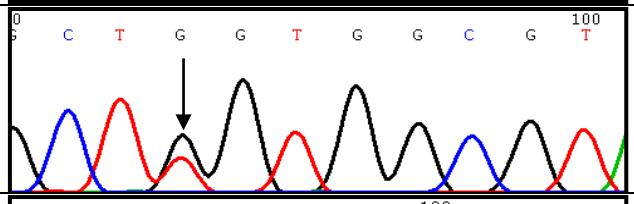
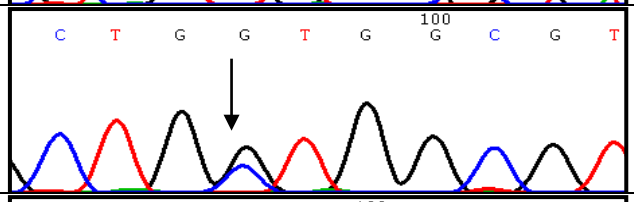
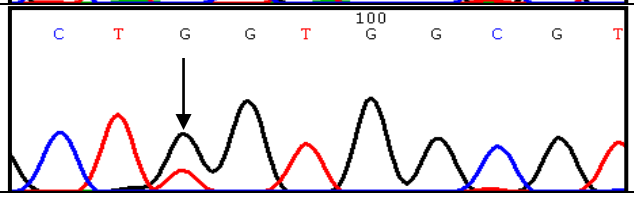
Patient	MUTYH	K-ras
2	Y165C/ Y165C	
13	Pas de mutation	
14	Q324H / normal	
18	1187insGG/ 1187insGG	
23	Y165C/ Y165C	
29	V22M/ -	
37	Pas de mutation	

Tableau 9. Mutations somatiques sur *k-ras* en fonction du statut *MUTYH*.

Nous n'avons pas retrouvé de mutation congénitale sur les ADN non tumoraux des patients. Nous avons étudié l'ensemble la totalité des séquences et notamment les séquences GAA sur *APC* comme sur le fragment A3 (Figure 19).

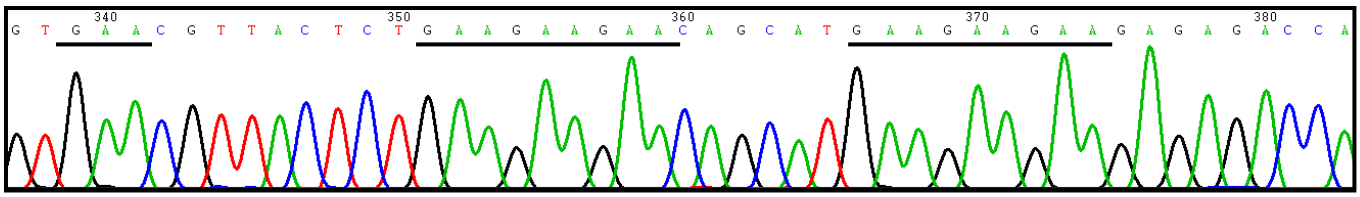


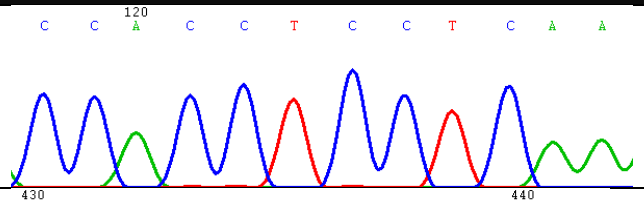
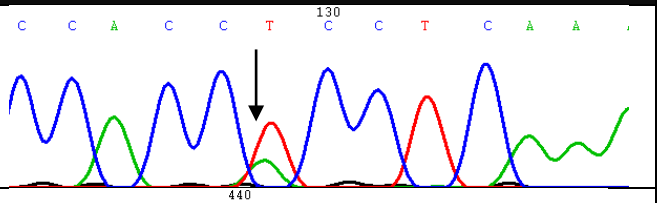
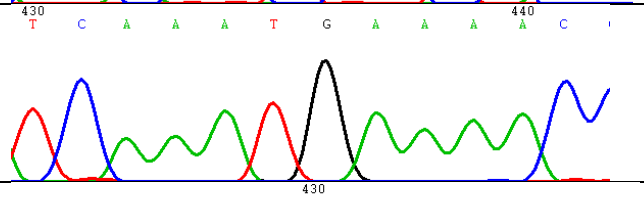
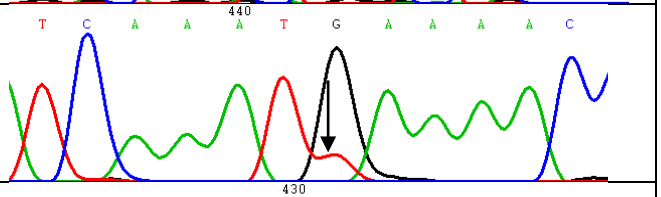
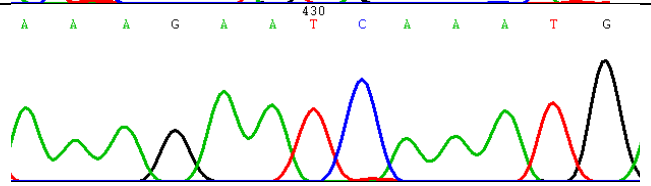
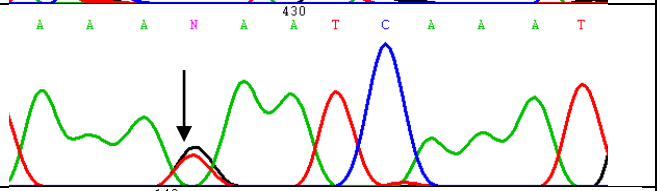
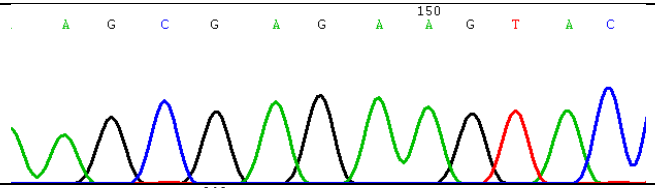
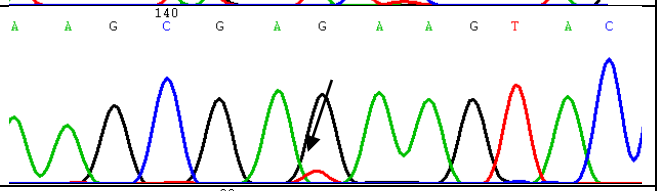
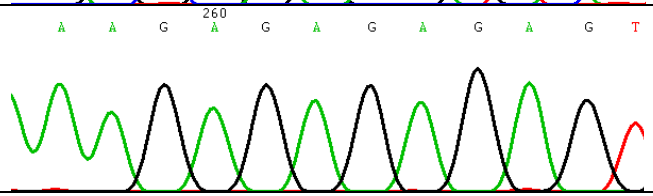
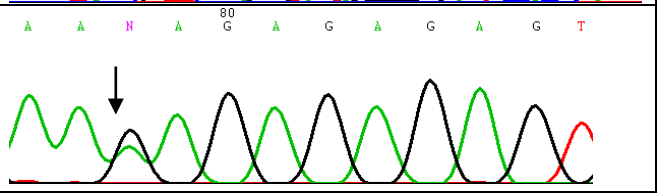
Figure 19. Séquence A3 (patient n°23) : 7 séquences GAA.

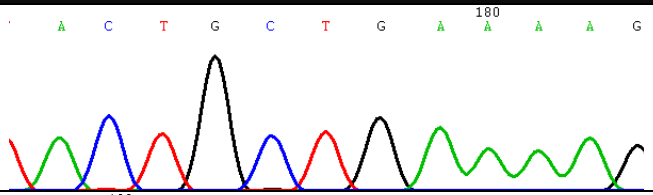
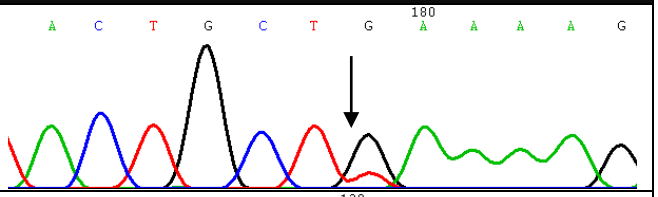
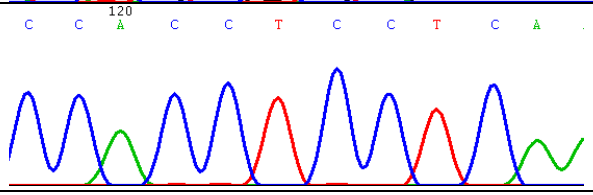
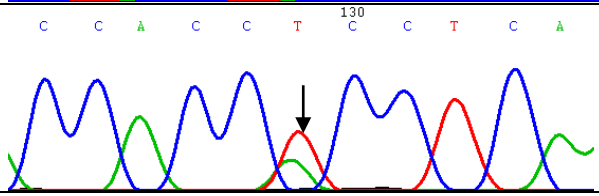
Les fragments A5 et A4 étaient ceux qui donnaient les meilleurs résultats, nous avons décidé de les amplifier en priorité avec les patients qui présentaient une mutation biallélique sur MUTYH. Aucune mutation n'a été retrouvée sur le fragment A4. Pour le fragment A5, nous avons retrouvé 7 mutations sur APC. Elles concernaient 5 des 6 patients mutés sur MUTYH (83,3%). Il s'agissait de 4 transversions G:C→T:A (57,1 %) et 3 autres mutations.

Les mutations sur APC et le statut *MUTYH* des patients sont détaillés dans le tableau 10. Cinq patients sur les 6 ayant une mutation biallélique sur *MUTYH* avaient donc une ou plusieurs transversions somatiques sur *APC* et/ou *K-ras*.

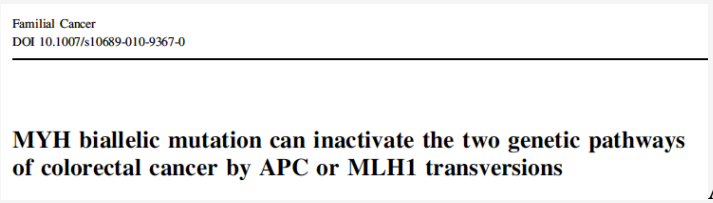
Le seul patient sans transversion identifiée était le patient 43 qui présentait en plus une perte d'expression de la protéine hMLH1 sur son tissu tumoral. A la suite de cette découverte d'association d'une mutation germinale sur *MUTYH* et d'un phénotype MSI, nous avons décidé d'étudier cette famille.

Tableau 10. Mutations sur l'exon 15 d'APC et statut MUTYH.

Patient / Fragment APC	MUTYH	Séquence normale	Séquence mutée
18 / A5	1187insGG / 1187insGG		
18 / A5	1187insGG / 1187insGG		
21 / A5	1103delC / 1103delC		
23 / A5	Y165C / Y165C		
29 / A5	V22M / -		

Patient / Fragment APC	MUTYH	Séquence normale	Séquence mutée
35 / A5	G382D / G382D		
43 / A5	1187insGG / 1187insGG		

INTERACTIONS ENTRE MUTYH ET MLH1.


Familial Cancer
DOI 10.1007/s10689-010-9367-0

MYH biallelic mutation can inactivate the two genetic pathways of colorectal cancer by APC or MLH1 transversions

Annexe B, page 159.

BUT :

Etudier une famille présentant une mutation germinale sur *MUTYH* dont l'un des membres perd l'expression de hMLH1 en immunohistochimie.

METHODES :

La sélection de cette famille provient de l'étude précédente (Lefevre, JH *et al.* 2006).

Extraction de l'ADN : méthode similaire à celle décrite précédemment.

1. Amplification des gènes MUTYH et hMLH1 :

L'amplification de *MUTYH* a été déjà décrite. L'amplification de *hMLH1* était réalisée dans un mélange similaire avec une concentration de $MgCL_2$ de 1,5 mM. Les primers utilisés sont détaillés dans le tableau 11. L'amplification a été réalisée sur un GeneAmp PCR system 9700 (Applied Biosystems, Foster city, CA, USA). Les produits de PCR ont été purifies sur une plaque MultiScreen^{TM} PCR 96 puits (MILLIPORE, Billerica, MA, USA) et séquencés grâce à l'ABI 3730 DNA sequencer (Applied Biosystems, Foster City, CA, USA). L'analyse des séquences fut faite à l'aide du logiciel SeqScape software, version 2.5 (Applied Biosystems, Foster City, CA, USA) en les alignant avec les séquences de référence d'*hMLH1* (NM_000249.1).

Exons	Amorce sens	Amorce antisens	Taille (bp)	(T°)
1	AGGCACTGAGGTGATTGGC	AGTCGTAGCCCTTAAGTGA	230	52
2	AATATGTACATTAGAGTAGTTG	CAGAGAAAGGTCCTGACTC	213	52
3	TTTACTCATCTTTTTGGTATC	GAGAAATACATACTAACAAATG	204	50
4	CCT TTC CCT TTG GTG AGG TG	TTGAGACAGGATTACTCTGAG	233	52
5	GATTTTCTCTTTTCCCCTTGGG	CAAACAAAGCTTCAACAATTTAC	190	52
6	CAAGTACTTCTATGAATTTACAAG	CACCATCTAGCTCAGCAACTG	233	52
7	ACTAAAAGGGGGCTCTGACAT	CAAACAGCAGAAGACACAAGG	214	52
8	GCCATGAGACAATAAATCCTTG	TGTGATGGAATGATAAACCAAG	200	52
9	ATGGGTGGGTGAATGGGTGAAC	CTGTGGGTGTTTCCTGTGAGTGG	364	52
10	CCTCAGGACAGTTTTGAACTGG	GAGGAGAGCCTGATAGAACATCTG	203	52
11	GGGCTTTTTCTCCCCCTCCC	AAAATCTGGGCTCTCACG	289	52
12	GTACTGCTCCATTTGGGGAC	GCAGAGAGAAGATGCAAGTGA	579	52
13	GTTCATTCACAGCTCTGTAG	CTTTCTCCATTTCCAAAACC	346	52
14	TGTCTCTAGTTCTGGTGCCTG	AAGCCTGTGCCCTCCCAAC	316	50
15	AACTGGTTGTATCTCAAGCAT	AAATAAGATATTAGTGGAGAG	206	50
16	CAT TTA TGG TTT CTC ACC TG	GAAGTATAAGAATGGCTGTC	296	52
17	CAGCATTATTTCTTGTTCCC	CTTTCCCTCCAGCACACATG	193	52
18	TAAGTAGTCTGTGATCTCCGT	GTATGAGGTCCTGTCCTAG	245	52
19	GGACACCAGTGTATGTTG	CACATCCCACAGTGCATAA	258	52

Tableau 11. Amorces pour l'amplification de *MLH1*.

2. Recherche des mutations somatiques sur BRAF:

Le séquençage de *BRAF* a été réalisé en utilisant les amorces suivantes: sens - 5'-TTGACTTCTAAGAGGAAAGATGAAGT-3' et anti-sens 5'-AGCATCTCAGGGCC-AAAAAT-3'. Les conditions de réalisation de la PCR était les suivantes : 30s à 94°C, 30 s à 60°C et 1 min à 72 °C pendant 35 cycles. Les produits de PCR étaient ensuite purifiés et séquencés en utilisant l'ABIPrism 3730 DNA sequencer (Applied Biosystems, Foster City, CA, USA).

3. Recherche de mutation somatiques sur APC et k-ras:

Cette technique a été décrite précédemment.

4. Recherche de grands réarrangements sur MLH1 et MSH2 en MLPA:

La détection de remaniements génomiques de grande taille utilise la technique MLPA. Le principe est fondé sur une réaction de ligation de 2 oligonucléotides adjacents, formant une sonde après leur hybridation à des séquences cibles spécifiques permettant d'obtenir pour

chaque locus un fragment amplifié de taille différente et de les quantifier par électrophorèse sur un séquenceur ABI-3730. Chaque fragment peut alors être visualisé sous forme d'un pic qui selon son amplitude par rapport au témoin permet la détection du nombre de copies au niveau d'un locus.

Le kit SALSA MLPA P003 MLH1/MSH2 probemix kit (MRC-Holland®, Netherlands. www.mrc-holland.com) a été utilisé selon les recommandations du fabriquant afin de rechercher des grands arrangements sur un ou plusieurs exons de *hMLH1* et *hMSH2*. Le mix contient 16 paires de sondes pour *hMSH2,* 19 pour *hMLH1* et 7 paires spécifiques de séquences à proximité de *hMLH1* et *hMSH2*. Les produits de PCR de différente taille ont été analysés avec l'ABI 3730 DNA analyser (Applied Biosystems, Foster City, CA, USA) et les résultats étudiés à l'aide du logiciel GeneMapper® software, version 4.0 (Applied Biosystems, Foster City, CA, USA). Des échantillons d'ADN de témoins sains étaient utilisés comme contrôles à chaque manipulation. La hauteur de pic de chaque fragment était comparée à celles des contrôles. Une délétion ou une duplication était suspectée en cas de variation de taille supérieure à 30%.

5. Etude de la méthylation du promoteur de hMLH1 et de hMSH2:

La méthylation du promoteur était étudiée en utilisant la technique de MSP (*Methylation Specific PCR*). Après traitement par Bisulfite de l'ADN qui remplace les cytosines non méthylées par un uracile et ne modifie pas celles qui sont méthylées à l'aide du CpGenome™ DNA Modification Kit (Millipore). La MSP était ensuite réalisée selon les recommandations du fabriquant (Serologicals Corporation): 2 PCR étaient programmées chacune avec un couple d'amorces différents (le premier spécifique du gène méthylé et l'autre du gêne sans modification épigénétique). Un contrôle de la bonne réalisation des PCR était donné grâce à un troisième set d'amorces s'hybridant sur un échantillon méthylé ou non.

Les conditions de PCR étaient les suivantes: 12 min à 94°C; 40 cycles de 94°C (30s), 56 à 62°C (30s) et 72°C (30s); avant une élongation de 10 minutes à 72°C. Les produits de PCR étaient ensuite séparés sur un gel de polyacrylamide et les bandes étaient visualisées avec de l'éthidium bromure. Toutes les PCR étaient réalisées avec des échantillons d'ADN méthylés et non méthylés.

6. Etude du statut microsatellitaire:

Une PCR avec les 5 marqueurs mononucléotidiques (BAT-25, BAT-26, NR-21, NR-22 et NR-24) a été réalisée.

7. Extraction de l'ARN et synthèse du cDNA complémentaire:

L'ARN total était extrait des cellules en utilisant du Trizol selon les recommandations du fabriquant. Après purification, le cDNA était synthétisé en utilisant 0,5 µg d'ARN total, des primers aléatoires et des dNTPS (1 mM), 200 M-MLV reverse transcriptase et 200U d'inhibiteur de la RNase. La transcription reverse était conduite pendant 50 minutes à 37°C et arrêtée en augmentant la température à 65°C pendant 5 minutes.

RESULTATS:

1. Famille H. :

Cette famille a été identifiée lors de la précédente étude sur l'implication de *MUTYH* dans la polypose adénomateuse. La recherche de mutation sur germinale sur *APC* avait été réalisée et aucune mutation n'avait été identifiée. Le séquençage complet de *MUTYH* avait mis en évidence une mutation biallélique homozygote (c.1185_1186dup, p.Glu396GlyfsX43) (cf. p. 60). Deux apparentés avaient alors bénéficié d'une coloscopie de dépistage. Un frère (II_3) et un neveu (III_{12}) présentaient des polypes coliques et la même mutation fut identifiée chez les patients. L'arbre généalogique de la famille H. est détaillé sur la figure 20.

CRC = cancer colorectal; age au moment du diagnostic

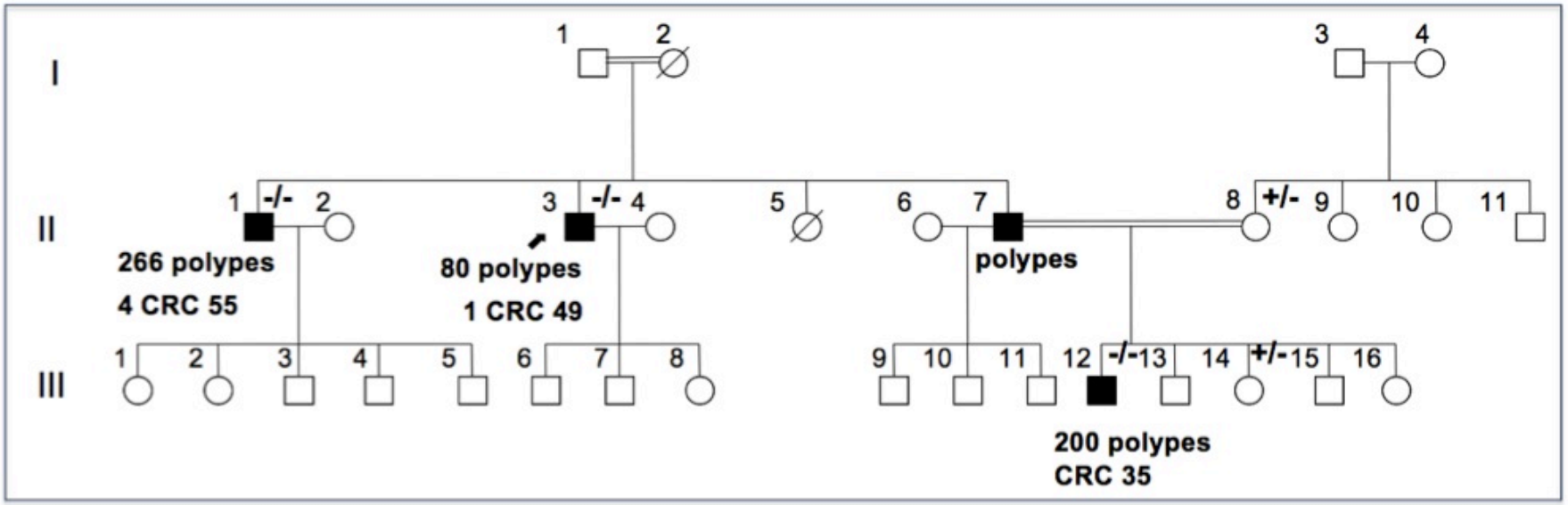


Figure 20. Arbre généalogique de la famille H.

2. Confirmation de l'instabilité microsatellitaire :

Comme tous les patients opérés dans le département pour un CCR, une recherche de perte d'expression des protéines hMLH1 et hMSH2 a été réalisée. Le patient II3 perdait l'expression de hMLH1 sur le tissu tumoral. L'instabilité des microsatellites a été confirmée en PCR chez le patient (Figure 21).

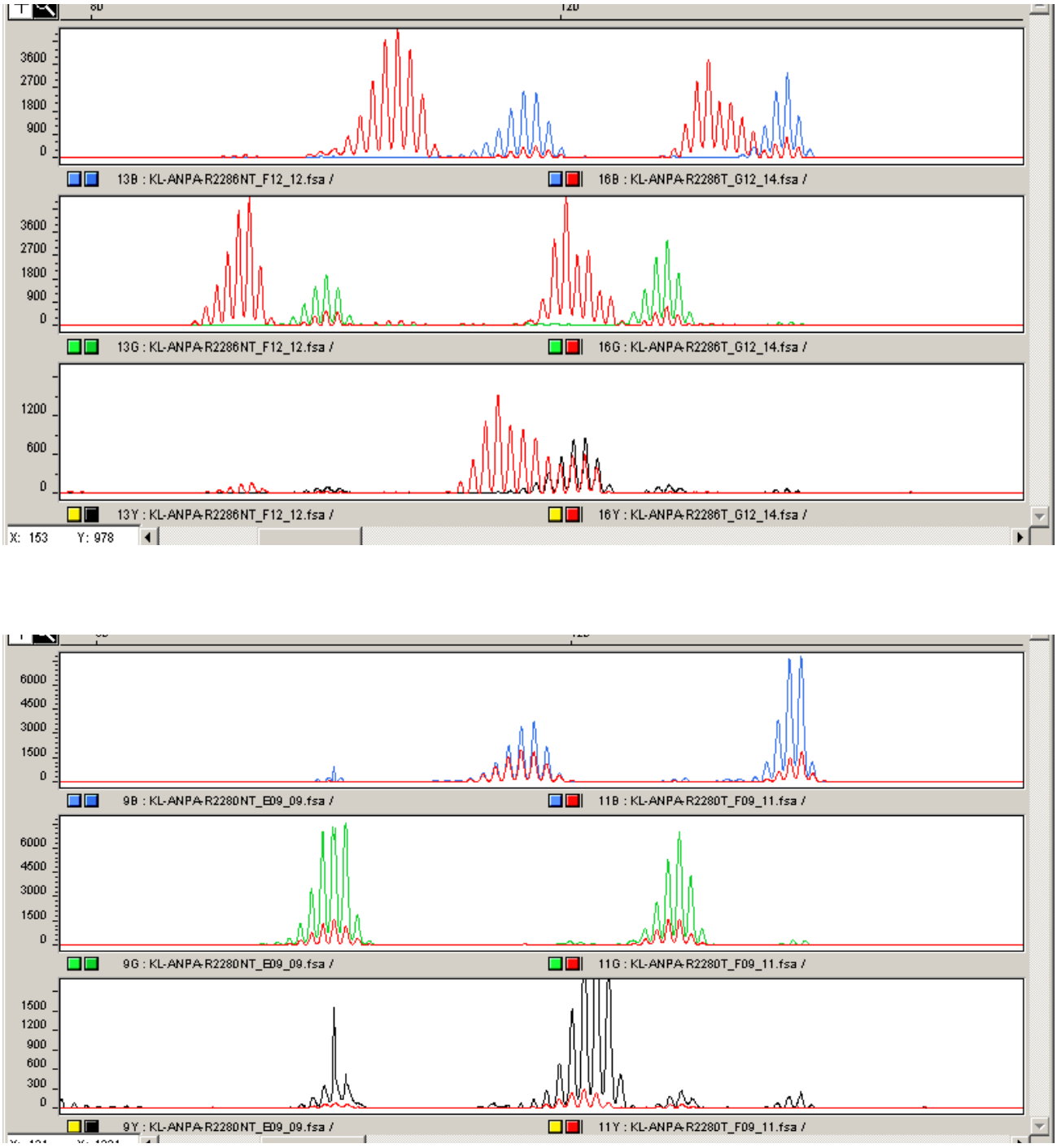


Figure 21. Résultat des PCR pour l'exploration du Phénotype MSI

3. Recherche d'une méthylation du promoteur d'hMLH1 :

La méthylation du promoteur étant responsable d'environ 15% des CCR sporadiques, nous l'avons étudiée chez le patient II3 à l'aide d'une MSP. Le promoteur n'était pas méthylé. Ce résultat fut confirmé par l'étude du gène *BRAF* qui est muté sur le tissu tumoral en cas de méthylation du promoteur (Deng, G *et al.* 2004) : il présentait deux allèles sauvages.

4. Recherche de mutation germinale sur MLH1:

La perte d'expression de MLH1 n'étant pas liée à une méthylation, le séquençage complet de *MLH1* fut entrepris et aucune mutation germinale ne fut identifiée. Une MLPA a la recherche de grandes délétions sur *MLH1* et *MSH2* fut alors faite. Aucune délétion ou duplication ne fut identifiée (extrait de la MLPA représenté sur la figure 22).

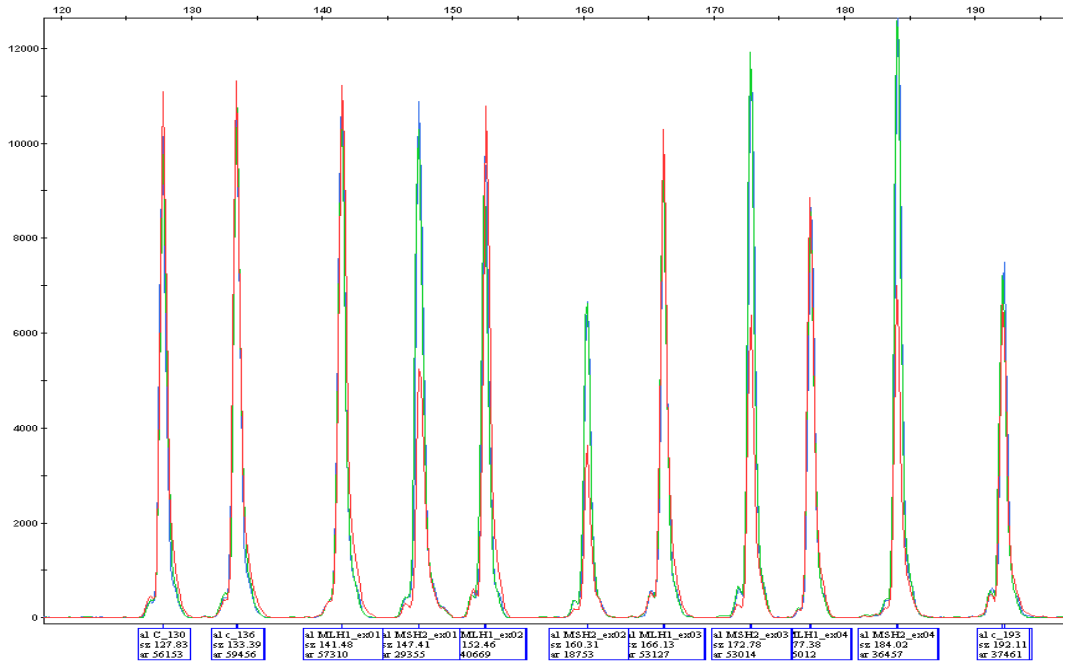


Figure 22. MLPA du gène *MLH1*.

5. Recherche de mutation somatique sur hMLH1 :

L'absence de mutation germinale sur *MLH1* nous a conduit ensuite à recherche une mutation somatique sur *MLH1*. L'analyse de la séquence publiée de *MLH1* révèle que 54

triplets GAA étaient présents parmi les 756 codons de la séquence codante. Trente quatre codaient pour un acide glutamique et pouvaient devenir un codon stop (TAA) en cas de transversion G>T par inactivation biallélique de *MUTYH*. Toute la séquence codante de *MLH1* provenant du tissu tumoral fut séquencée. Une transversion fut identifiée à la jonction intron-exon de l'intron 7 : +5 (c.588+5G>T) (Figure 23). Cette transversion touchait une séquence AAGTT. Après extraction d'ARN et RT-PCR, l'analyse du cDNA montra que l'exon 7 était délété. Ces résultats confirmaient le caractère délétère de cette transversion sur *MLH1* (Figure 24).

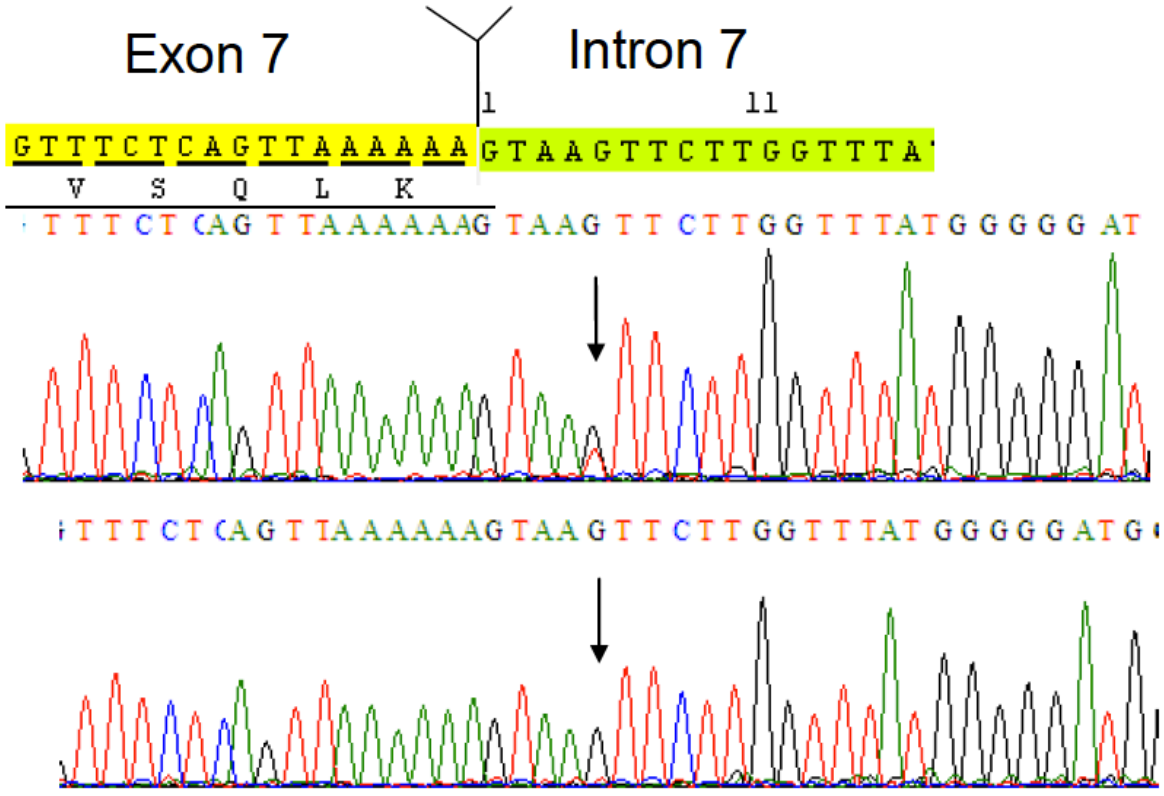


Figure 23. Transversion sur le gène *MLH1* +5 (c.588+5G>T).

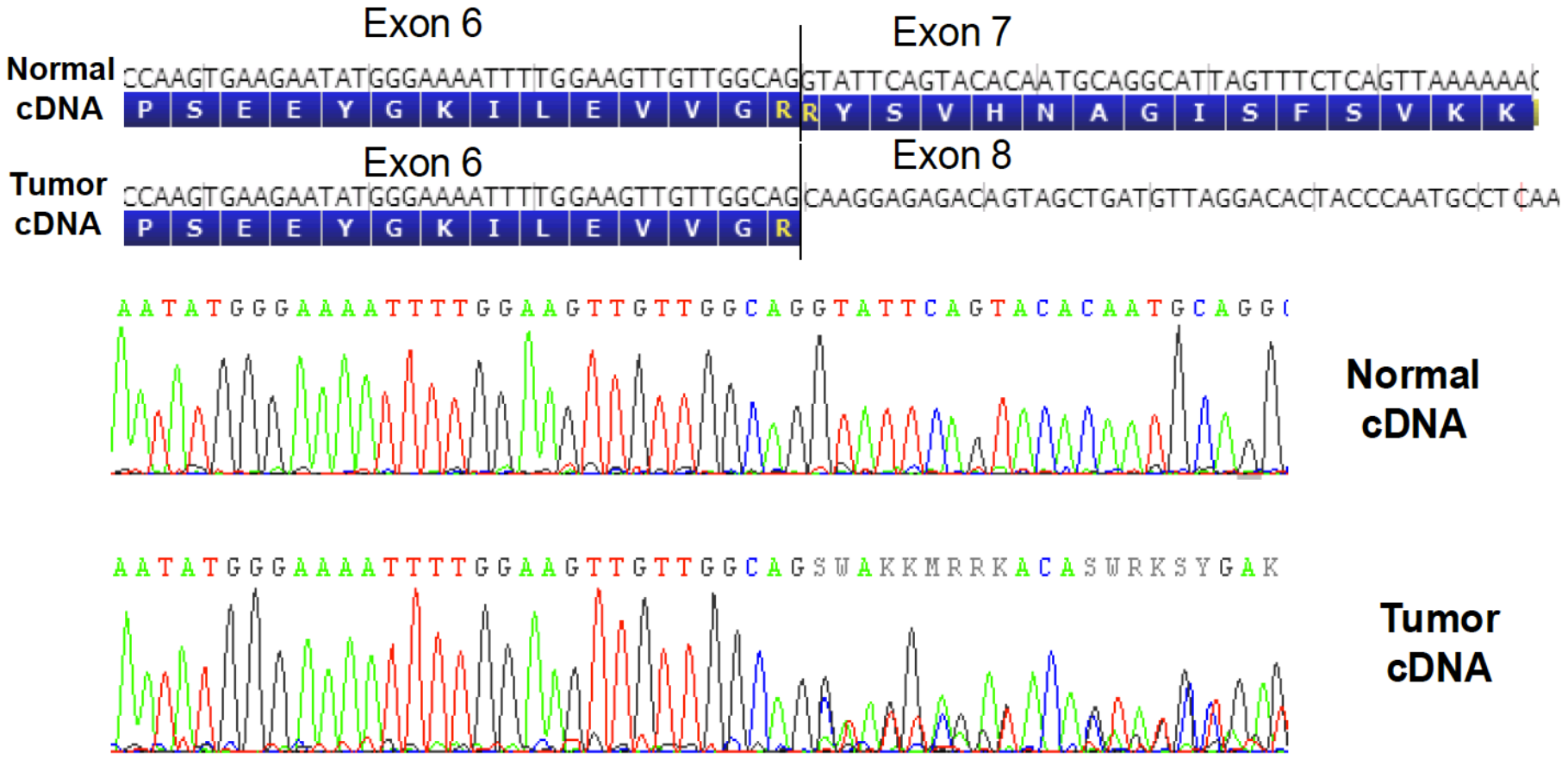


Figure 24. Séquence du cDNA

CORRELATION GENOTYPE – PHENOTYPE DANS LES POLYPOSES

ORIGINAL ARTICLE – GASTROINTESTINAL ONCOLOGY

APC, MYH, and the Correlation Genotype-Phenotype in Colorectal Polyposis

Annexe C, page 166.

BUT :

Etudier la corrélation génotype-phénotype au sein d'une cohorte de patients présentant une polypose adénomateuse liée à *APC*, *MUTYH* ou sans mutation identifiée.

METHODES :

1. Patients:

Au sein d'une cohorte de 515 patients pris en charge à l'Hôpital Saint-Antoine, 384 patients ont pu avoir un séquençage génétique sur le gène *APC* comme décrit précédemment (Lefevre, JH *et al.* 2006). Les patients sans mutation germinale identifiée sur *APC*, une recherche de mutation biallélique sur *MUTYH* a été réalisée (cf. étude précédente).

Cinq groupes ont été constitués en fonction de la mutation germinale identifiée :

- mutation sur *APC* dans les zones responsables d'une polypose atténuée (avant le codon 163 et après le codon 1596), groupe SAFAP
- mutation sur *APC* dans les zones responsables d'une polypose atténuée en prenant la définition la plus large (avant le codon 175, dans l'exon 9 et après le codon 1403), groupe LAFAP
- mutation sur *APC* dans les autres zones, groupe CFAP
- mutation sur le gène *MUTYH*
- aucune mutation identifiée

Les données cliniques, endoscopiques et chirurgicales ont été récupérées et analysées. La définition utilisée pour parler de polypose atténuée était : âge au moment de la chirurgie prophylactique supérieur ou égal à 35 ans et nombre de polypes inférieur à 100.

2. Analyse statistique:

Les variables quantitatives étaient analysées à l'aide du teste du Chi-2 et les variables continues par le test de Student. Une valeur de p inférieure à 0,05 était jugée comme significative. Les calculs statistiques ont été réalisés à l'aide du logiciel StatView (Version 5, 1992-1998, SAS).

RESULTATS :

1. Population et groupes de patients:

Selon la position de la mutation, 13 patients étaient dans le groupe SAFAP et 41 dans le groupe LAFAP. Il y avait donc 322 ou 294 patients dans le groupe CFAP selon le groupe respectivement. Parmi les 49 patients sans mutation sur le gène *APC*, une mutation biallélique sur *MUTYH* a été identifiée chez 17. Il restait donc 32 patients sans mutation expliquant leur polypose colorectale. La distribution des patients est détaillée sur la figure 25.

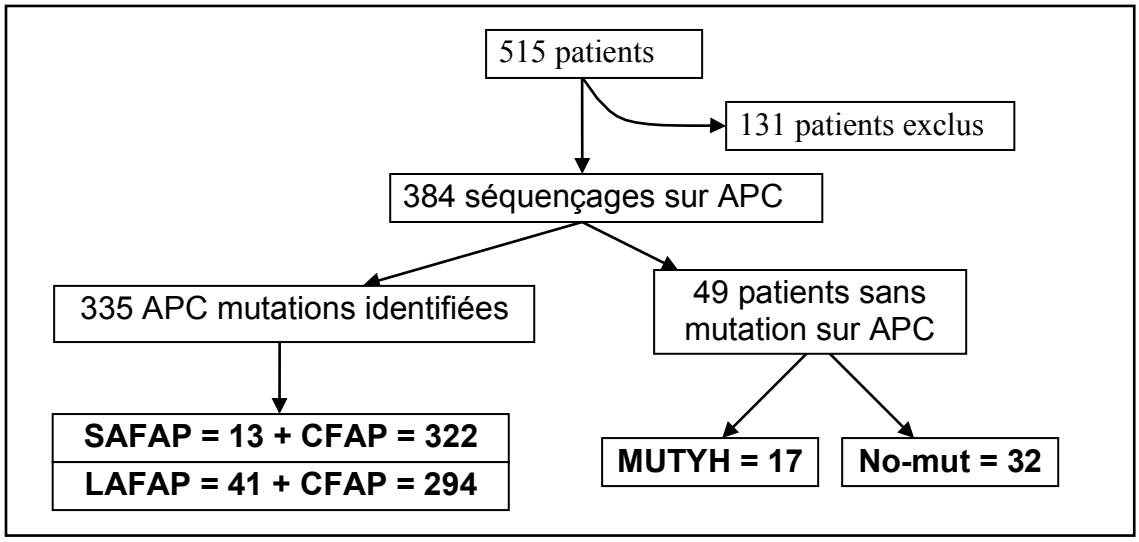


Figure 25. Répartition des patients selon la mutation identifiée.

2. Analyse des mutations et histoire familiale:

Les 13 patients du groupe SAFAP provenaient de 2 familles (n=6 patients) et les 7 restants étaient des cas individuels. Sept mutations différentes avaient été retrouvées chez ces 13 patients. Les 322 patients correspondant au groupe CFAP comprenaient 180 cas de 65 familles et 142 cas individuels. Dans le groupe LAFAP (n=41), 20 patients provenaient de 8 familles différentes et les 22 cas restants étaient des cas sporadiques. Les 294 CFAP correspondant provenaient de 59 familles (166 patients) et 128 cas individuels.

Parmi les 335 mutations identifiées sur *APC*, 272 (83,4%) étaient des mutations ponctuelles, 36 des délétions larges, 14 des altérations de l'épissage et 4 des mutations sur le promoteur. Les deux mutations les plus fréquentes étaient situées sur le codon 1309 (n=48, 17,3%) et sur le codon 1062 (n=20, 7,2%). Les 17 patients mutés sur *MUTYH* appartenaient à 15 familles différentes. Les origines ethniques des patients n'étaient pas connues.

3. Caractéristiques cliniques des patients:

Elles sont détaillées dans le tableau 12. Les comparaisons entre les groupes CFAP et SAFAP ou LAFAP sont données dans le tableau 13.

4. Nombres de polypes colorectaux:

Le nombre exact de polypes était connu chez 177 patients (46,1%). En fonction de la définition habituelle de la polypose atténuée, le seuil de 100 polypes était utilise comme cut-off. Aucune différence statistiquement significative était observe entre les patients ayant une mutation dans les zones atténuées SAFAP ou LAFAP et les patients avec une mutation classique.

La localisation majoritaire des polypes était connue pour 158 patients (41,1%) et aucune différence n'était également observée entre les groupe avec une mutation située dans les zones atténuées et une mutation classique.

Tableau 12. Caractéristiques cliniques et endoscopiques selon la mutation.

	Cohorte	Intervalle étroit		Intervalle large		MUTYH	Pas de mutation
		CFAP mutation	AFAP mutation	CFAP mutation	AFAP mutation		
n	384	322	13	294	41	17	32
Sexe (Masculin)	177 (49)	159 (49.3)	5 (38.5)	147 (50)	17 (41.4)	12 (70.6)	15 (46.9)
Age lors de la première colectomie	29.4 (±11.6)	27.9 (±10.7)	32.5 (±8.7)	27.7 (±10.9)	31.0 (±8.7)	48.0 (±8.6)	34.0 (±12.0)
Nb de polypes > 100 [a]		90[1] (68.7)	8 (61.5)	79[2](68.7)	19 [3](65.5)	6 (37.5) [4]	7[5] (41.2)
Localisation des polypes		17 / 54 / 50	1 / 7 /1	15 / 51 / 42	3 / 10 / 9	7/ 0 / 3	4 / 5 / 7
Droit / Gauche / Diffus		(14.0/44.6/41.3)	(11.1/88.8/11.1)	(15.9/47.2/38.9)	(13.6/45.4/40.9)	(70/0/30)	(25/31.2/43.7)
CRC	87 (22.6)	67 (21.1)	2 (15)	57 (19.4)	12 (29.3)	13 (76.5)	7 (21.9)
Présence de tumeur desmoïde	58 (15.1)	55 (17.5)	3 (23.1)	42 (14.3)	16 (39.0)	0 (0)	0 (0)
Polypes glandulaires gastriques	108 (28.1)	84 (34.7) [6]	6 (46.1)	67 (30.3) [7]	23 (65.7)[8]	2 (14.3)[9]	6 (19.3) [10]
Polypes duodénaux	145 (37.8)	129 (53.5)[11]	4 (31)	115 (52.2) [12]	18 (52.9) [13]	3 (21.4)[9]	9 (29.0) [10]
IRA	96 (25.1) [14]	76 (23.7) [14]	5 (38.5)	71 (24.3) [14]	10 (24.4)	7 (41.1)	8 (25)
IRA-> AIA	58 (60.4)	54 (71)	1 (20)	54 (76.0)	1 (10)	1 (14.3)	1 (12.5)

1 : connu pour 131 patients, 2 : connu pour 115 patients; 3: connu pour 29 patients; 4: connu pour 16 patients ; 5: connu pour 17 patients, 6: connu pour 243 patients ; 7: connu pour 221 patients; 8: connu pour 35 patients; 9: connu pour 14 patients; 10: connu pour 31 patients; 11: connu pour 241 patients; 12: connu pour 220 patients; 13: connu pour 34 patients; 14: 2 patients avaient eu une intervention sans colectomie.
IRA: anastomose iléo-rectale; AIA: anastomose iléo-anale

Tableau 13. Comparaison entre les patients avec une mutation atténuée ou classique sur APC.

	FAP mutation vs. AFAP mutation	
	Intervalle étroit	Intervalle large
Age lors de la première colectomie	p=0.1339	p=0.066
Présence de tumeur desmoïde	p=0.6029	p<0.0001
Nb de polypes (<100 / > 100)	p=0.5972	p=0.7429
Localisation des polypes (droit / gauche / diffus)	p=0.1379	p=0.9841
Cancer	p=0.6209	p=0.1561
Polypes glandulaires gastriques	p=0.3940	p<0.0001
Polypes duodénaux	p=0.1095	p=0.9421
AIR	p=0.2255	p=0.9916
AIR→ AIA	p=0.0179	p<0.0001

AIR : anastomose iléo-rectale ; AIA : anastomose iléo-anale.

5. Comparaison entre les patients SAFAP/LAFAP et les patients avec une mutation classique sur APC:

L'âge de la colectomie ne différait pas en fonction de la mutation de même que le pourcentage de patients avec un cancer colorectal au moment de l'intervention.

Les polypes glandulo-kystiques étaient plus fréquemment observés chez les patients ayant une mutation dans les régions LAFAP (65,7% vs. 30,3%; p<0,0001). De même, ce groupe de patients présentaient plus fréquemment des tumeurs desmoïdes que les patients avec une mutation du groupe CFAP (39% vs. 14,3%; p<0,0001).

Le taux d'anastomose iléo-rectale était similaire entre le groupe. En revanche, les patients avec une mutation dans les zones atténuées étaient moins fréquemment réopérés pour une proctectomie complémentaire avec confection d'une anastomose iléo-anale.

6. Patients mutés sur MUTYH :

Les 17 patients avec une mutation biallélique sur *MUTYH* présentaient des différences significatives avec les patients porteur d'une mutation sur *APC* : un âge plus tardif au moment de la colectomie (48,0 ± 8,6 *vs*. 28,2 ± 10,7; p<0,0001), moins de polypes (p=0,0151), plus

souvent localisés à droite (p<0,0001), aucune tumeur desmoïde (p<0,0001), des cancers colorectaux plus fréquents (p<0,0001) et moins de polypes duodénaux (p=0,0148).

7. Corrélation phénotype – génotype :

Parmi les 177 patients ayant un nombre connu de polypes, 31 ont été opérés pour leur polypose après 35 ans et correspondaient donc au phénotype de polypose atténué. Sur ces 31 patients, 5 présentaient une mutation située dans les zones atténuées d'*APC*, 10 une mutation biallélique sur *MUTYH*. Une représentation schématique des patients est donnée sur la figure 26.

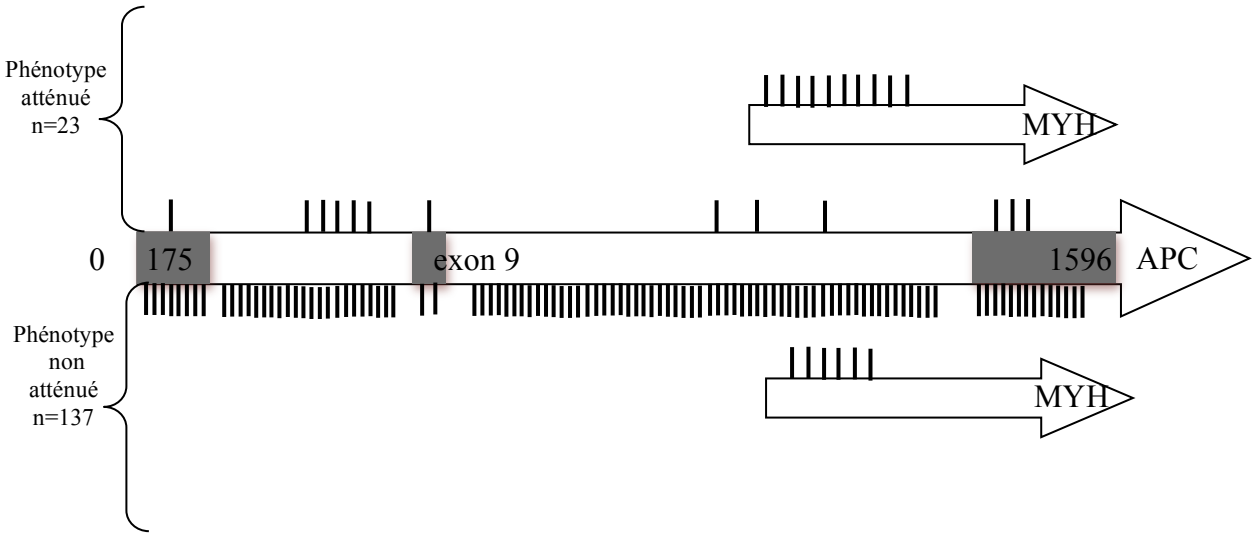
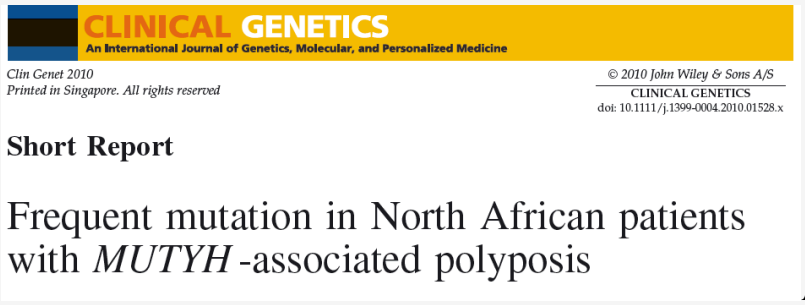


Figure 26. Répartition des mutations selon le phénotype observé.

Sur les 13 patients avec une mutation SAFAP, 3 présentaient un phénotype de polypose atténuée et parmi les 29 patients du groupe LAFAP, 4 avaient également un phénotype atténué. Dans le groupe des 17 patients avec une mutation biallélique sur *MUTYH*, 10 avaient un phénotype de polypose atténuée.

FREQUENCE DE LA MUTATION c.1227_1228dup DANS LES POLYPOSES LIEES A MUTYH, RECHERCHE D'UN EFFET FONDATEUR.


CLINICAL GENETICS
An International Journal of Genetics, Molecular, and Personalized Medicine

Clin Genet 2010
Printed in Singapore. All rights reserved

© 2010 John Wiley & Sons A/S
CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2010.01528.x

Short Report

Frequent mutation in North African patients with *MUTYH*-associated polyposis

Annexe D, page 174.

BUT:

Le but de cette étude était d'étudier la prévalence de la mutation c.1227_1228dup au sein d'une population de patients mutés sur *MUTYH* provenant d'Afrique du Nord et de rechercher l'existence d'un effet fondateur.

METHODES:

1. Sélection des patients

Tous les patients avec une polypose adénomateuse secondaire à une mutation biallélique suivi dans le centre d'Oncogénétique de la Pitié-Salpétrière ont été inclus dans cette étude. Ils avaient tous bénéficié d'une consultation d'oncogénétique au cours de laquelle une information avait été donnée et un consentement écrit recueilli.

2. Extraction de l'ADN et amplification du gène MUTYH:

Ces étapes ont déjà été détaillées précédemment.

3. Recherche d'un haplotype commun:

Pour chaque membre d'une famille ayant la mutation c.1227_1228dup, 10 marqueurs microsatellitaires entourant le gène *MUTYH* ont été utilisés. Ils se répartissaient de l'extrémité télomérique au centromère ainsi : D1S447, D1S211, D1S2733, D1S2713, D1S2802, D1S451, D1S2677, MUTYH, D1S322, D1S2797, D1S2720 (Figure 27).

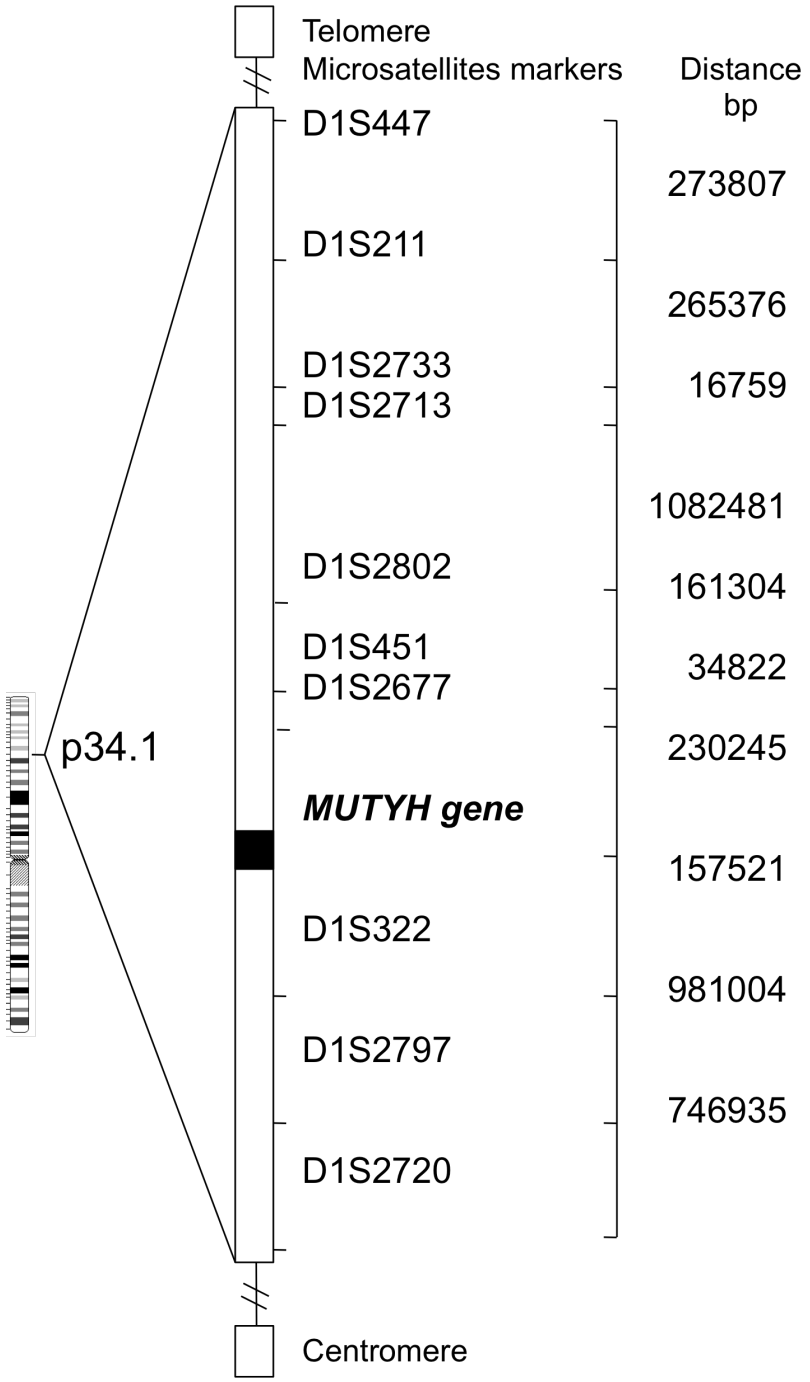


Figure 27. Répartition des microsatellites encadrant MUTYH.

Chaque marqueur a été amplifié en utilisant des amorces marquées (tableau 14). Les produits de PCR ont été analysés par électrophorèse en utilisant l'ABIPrism 3730 Analyzer (Applied Biosystems) et le logiciel Genemapper version 4.1 software (Applied Biosystems).

Microsatellite	Amorces
D1S447	5'-TTAGTCTGAGTTTGTGGGGG-3' 5'-GTTTTAACTTCATGGCTGCC-3'
D1S211	5'-AGCTACATGGCAGGATCAGA-3' 5'-GGATTCCTTGCTCTGGAAAG-3'
D1S2733	5'-TGCGGCGAGACAGACATC-3' 5'-AGGACCAGCGTGTGCG-3'
D1S2713	5'-CAGCCCCCAACACATAC-3' 5'-CCTTAGGAGTCTACAGACGCC-3'
D1S2802	5'-CAAACTAATTATCAAACAAAGGAAC-3' 5'-AGTGGCCTGTTATGTTGAAC-3'
D1S451	5'-ATCATGGGACTTTGCAGC-3' 5'-ATGGCCTGGCATAGTGTG-3'
D1S2677	5'-GTTCTGGAGTTTGCCTCTAA-3' 5'-AGCTGTGATTGTNCCACT-3'
D1S322	5'-GTCCCTAGCCTCCCAGCAT-3' 5'-CAAGAGCGAAAGTCCGTCTC-3'
D1S2797	5'-ATCACATCACACACAATGACTGTGG-3' 5'-TGTCCATTCAAAGGATTGGTCTC-3'
D1S2720	5'-AGCTACAAAGTGCTTTACTGACA-3' 5'-AATGGTCCAGGCAAAGT-3'

Tableau 14. Amorces de PCR pour les 10 marqueurs microsatellitaires.

RESULTATS:

1. Patients et familles:

Trente-six familles ayant une mutation biallélique sur *MUTYH* ont été incluses dans cette étude. Parmi ces familles, 13 (36%) provenaient d'Afrique du Nord : Algérie (n=7), Tunisie (n=5) et Maroc (n=1)). Les 23 familles restantes provenaient de pays Européens: France (n=11), Portugal (n=5), Belgique (n=1), Italie (n=1), and Turquie (n=1)) ou avaient des

origines mixtes : (France/Algérie (n=1), Allemagne/France (n=1), Belgique/France (n=1), et Arménie/France (n=1)).

2. Distribution des mutations sur MUTYH selon les ethnies d'origine:

Le cas index de 11 familles présentait la mutation c.1227_1228dup (p.Glu396GlyfsX43). Ces 11 familles provenaient d'Algérie (n=5), Tunisie (n=4), Portugal (n=1) et Maroc (n=1). Un autre cas index était hétérozygote composite (c.1227_1228dup /Y179C) et avait une double origine Algérienne et Française. Parmi les 14 familles ayant des origines d'Afrique du Nord (Algérie, Tunisie, Maroc ou origine mixte), 11 cas index (78,6%) avaient la mutation homozygote ou hétérozygote composite. En revanche, seulement un cas parmi les 22 familles ayant une autre origine présentait cette mutation (4,5%) et était d'origine Portugaise. Cette différence de prévalence était statistiquement significative ($p<0,0001$, test exact de Fischer).

3. Corrélation génotype-phénotype :

La mutation c.1227_1228dup fut également trouvée chez 4 autres patients appartenant aux familles étudiées portant à 15 le nombre porteurs de la mutation. Ils furent comparés au groupe de 27 patients ayant une mutation biallélique différente de c.1227_1228dup prouvée sur MUTYH. Les mutations identifiées étaient majoritairement Y169C (n=10, 37%) et G396D (n=13, 48%). Les détails de l'analyse sont donnés dans le tableau 15. Aucune différence d'expression phénotypique ne fut identifiée entre les deux groupes de mutation sur *MUTYH*. Seule l'existence d'une consanguinité fut retrouvée statistiquement plus fréquemment dans le groupe avec la mutation c.1227_1228dup.

	c.1227_1228dup (n=15)	Autre mutation sur *MUTYH* (n=27)	p
Homme	9/15 (60%)	28/27 (66.7%)	0,187
Consanguinité			
Non	2/13 (15%)	16/27 (59%)	
Oui	10/13 (77%)	4/27 (15%)	0,0006
Possible	1/13 (8%)	7/27 (26%)	
Plus de 30 polypes	11/15 (73%)	23/26 (88%)	0,215
Colorectal cancer	12/15 (80%)	19/27 (70%)	0,496
Polypes duodénaux	3/9 (33%)	4/17 (24%)	0,592
Antécédent familial de polypose	9/14 (64%)	14/27 (52%)	0,447
Age au moment du diagnostic	44,3 ± 8,9 [30-62]	44,2 ± 8,8 [24-60]	0,969
Nombre de polypes	75,5 ± 80,2 [7-266]	75,2 ± 40,0 [15-156]	0,855

Tableau 15. Corrélation génotype-Phénotype selon la mutation sur *MUTYH*.

4. Recherche d'un haplotype commun:

Les 10 microsatellites ont été amplifiés de par et d'autre du gène *MUTYH* chez chaque cas index des familles portant la mutation c.1227_1228dup. Le résultat est détaillé dans le tableau 16. Chaque famille présentait une similitude de taille pour 5 microsatellites : l'haplotype compris entre D1S2802 et D1S2797 était commun mesurant une taille d'au moins 1,3 cM.

		Taille théorique	Famille 1	Famille 2	Famille 3	Famille 4	Famille 5	Famille 6	Famille 7	Famille 8	Famille 9	Famille 10	Famille 11
MS1	D1S447	123-141	126/126	128/130	124/124	130/130	130/130	130/130	123/130	130/130	123/130	130/130	130/130
MS2	D1S211	170-198	185/185	166/181	183/183	166/166	166/166	166/166	166/181	166/166	166/181	166/166	166/166
MS3	D1S2733	107-121	113/113	105/111	105/105	111/111	111/111	111/111	111/111	111/111	105/111	105/111	111/111
MS4	D1S2713	227-279	262/262	252/264	260/264	264/264	264/264	264/264	266/274	264/268	264/264	264/264	264/264
MS5	D1S2802	192-202	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193
MS6	D1S451	174-188	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175
MS7	D1S2677	135-153	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139
MS8	D1S322	90	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106
MS9	D1S2797	144-180	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162
MS10	D1S2720	235-245	234/234	232/238	232/232	234/234	232/234	236/236	234/236	234/234	238/242	237/237	238/238

Tableau 16. Résultat du séquençage des microstatelllites chez les 11 familles.

ETUDE DU ROLE DE LA CYCLINE D1 DANS LES POLYPOSES INDERMINEES ET LES CANCERS COLORECTAUX PRECOCES.

Journal of Human Genetics (2011) 56, 58–63
© 2011 The Japan Society of Human Genetics All rights reserved 1434-5161/11 $32.00
www.nature.com/jhg

ORIGINAL ARTICLE

Cyclin D1 rare variants in UK multiple adenoma and early-onset colorectal cancer patients

Annexe E, page 180.

OBJECTIF:

Etudier le rôle des variants de la cycline D1 ayant une fréquence inférieure à 1% ou comprise entre 1 et 5% dans la cancérogenèse colique ainsi que l'impact du polymorphisme rs9344.

METHODES :

1. Sélection des patients :

La population d'étude était constituée de deux groupes de patients provenant de France et de Grande-Bretagne : 112 patients anglais avec entre 3 et 100 polypes adénomateux (Fearnhead, NS *et al.* 2004) (synchrones ou métachrones) et 44 ayant un CCR diagnostiqué avant 50 ans. Ces derniers provenait d'un essai clinique de phase II comparant l'utilisé du rofecoxib (Pendlebury, S *et al.* 2003) en adjuvant après colectomie pour CCR pour 38 d'entre eux. Les 6 patients restants avaient été opérés au *John Radcliffe's Hospital* dans le service de chirurgie Colorectale. Tous les patients de ce groupe UK, à l'exception d'un indien et d'un patient noir provenant des Caraïbes, étaient blancs d'origine Anglaise. Aucun de ces patients ne présentaient d'argument pour une PAF, une MAP ou un syndrome HNPCC et certains avaient déjà été testés pour rechercher des mutations germinales sur *APC* ou *MUTYH*.

Le groupe de patients Français provenait de notre département et était constitué de 75 patients avec une polypose indéterminée et 56 avec un CCR diagnostiqué avant 50 ans. Ces patients avaient été opérés dans le service et tous avaient bénéficié d'une recherche de phénotype MSI et de perte d'expression de hMLH1 ou hMSH2. Tous les patients avec une polypose avaient eu une consultation avec une onco-généticienne et un séquençage avait exclut une mutation sur *APC* ou *MUTYH*. Les origines ethniques de ces patients n'étaient pas disponibles.

Tous les patients avec une polypose avaient eu une confirmation diagnostique de son caractère adénomateux et le nombre précis de polypes était connu (sauf pour 24 patients anglais et 14 patients français qui avaient une polypose classée comme multiple).

Le groupe contrôle était constitué de 866 individus sains dont l'ADN avait été prélevé dans le cadre de l'étude PoBI (Winney, B *et al.* 2012) visant à étudier les effets fondateurs dans 10 différentes régions de Grande-Bretagne. La collecte des échantillons ainsi que le recueil des données étaient faite après accord des comités d'éthiques locaux et information donnée aux patients.

2. Extraction de l'ADN :

Cette étape a déjà été détaillée précédemment pour les patients. Pour le groupe contrôle, les échantillons de sang étaient rapportés au laboratoire à température ambiante et laissé sur alicots pendant deux jours pour séparer la phase lymphocytaire. Dix mL de sang étaient ensuite prélevés de l'échantillon en utilisant une colonne (Qiagen, Valencia, CA, USA) ou des beads magnétiques (GeneCatcher^{TM}, Invitrogen, Carlsbad, CA, USA). La concentration d'ADN était mesurée en utilisant du PicoGreen (Ahn, SJ *et al.* 1996) et normalisée ensuite pour obtenir une concentration de 25 ng/uL.

3. PCR :

Tous les exons, les extrémités 5' et 3', les jonctions intron-exons et environ 1,5 Kb du promoteur de *CCND1* ont été amplifiées en utilisant 21 PCR. Les amorces et la taille des amplicons sont détaillées dans le tableau 17.

	Amorce sens 5'-3'	Amorce antisens 5'-3'	position	taille
promoteur 5	TTGGCTCATGCTAATTCCAG	CCCACCTTTCAGGGTGAAT	-1976/-1462	514
promoteur 4	CCCAGGACCCGGAATATTAG	GCCGGCATAATTCAGAACAA	-1539/-1067	472
promoteur 3	TGGGGACCCTCTCATGTAAC	GCCAGCATTTCCTTCATCTT	-1168/-717	451
promoteur 2	CCGACTGGTCAAGGTAGGAA	ATTTAGGGGGTGAGGTGGAG	-789/-277	512
promoteur 1	GTCACCCCTTCTCGTGGTCT	ACTCTGCTGCTCGCTGCTAC	-373/88	461
exon 1	CGGGCTTTGATCTTTGCTTA	GCAACAAGTTGCAGGGAAGT	-33/455	488
exon 2	CTCGGCGCCCTCCAGACCT	CCACCGCCTGCACCTTCC	-111/296	407
exon 3	GCCGCTCACCCTGTGTTC	ACCCAGGTGGAGAGCAAGA	-22/246	268
exon 4	ATGTGGAGCCTCAGATACCG	AGGTGTCTCCCCCTGTAAGC	-139/205	344
exon 5.1	TTATAAAGGCTTCCGGGTCA	CGCTCAGGGTTATGCAAATA	-101/421	522
exon 5.2	CCGGAGCATTTTGATACCAG	CATGCATATGAGCAAGCAAAG	247/743	496
exon 5.3	GCACAGCTGTAGTGGGGTTC	AGTGCTTGGAAATGGAATGG	688/1183	495
exon 5.4	GGTTCAACCCACAGCTACTTG	CCTTCCGGTGTGAAACATCT	1115/1541	426
exon 5.5	TCTCACATTGCCAGGATGAT	TGCCTGTCCAATCAGATGAC	1380/1843	463
exon 5.6	CACAAACCTTCTGCCTTTGA	TAAAGGAAGGGGCAGGGGATAA	1683/1917	234
exon 5.7	CCTAAGTTCGGTTCCGATGA	AACGGTAGCAGCGCAATAAG	1871/2137	266
exon 5.8	GGAGGAGGTGTGAGGAGGA	CAGCAACCTTTTTGGAATAGAGCA	2042/2370	328
exon 5.9	TGTGATCAATTTTGACTTAATG	TGCAACTTCCATAGCTACACG	2317/2649	332
exon 5.10	GGGGGCGTAGCATCATAGTA	CAGCCAGGATGGTTGAGGTA	2583/3057	474
exon 5.11	TCAATGAAGCCAGCTCACAG	CCAGGACTTGTGCCCTTG	2852/3133	281
exon 5.12	TTGAGGGACGCTTTGTCTGT	TTCTAGGGCACCAGGACAC	3085/3426	341

Tableau 17. Amorces utilisées pour l'amplification de *CCND1*.

Les conditions de PCR étaient similaires à celles déjà décrites. Le cycle de PCR était le suivant : dénaturation de l'ADN à 95°C pendant 10 minutes, 35 cycles avec 25 sec à 95°C, élongation pendant 35 sec puis 30 secondes à 72°C. une période d'élongation finale de 5 minutes à 72°C terminait la PCR. Les produits de PCR étaient ensuite déposés sur un gel d'agarose pour vérifier la bonne amplification de chaque fragment.

4. Analyse des mutations :

Le dépistage des mutations dans le groupe des patients anglais a été réalisé en utilisant la technique dHPLC (WAVE^{TM} DNA fragment analysis, Transgenomic, Inc., Omaha, NE,

USA). Les gradiants de température ont été déterminés en utilisant le logicel WaveMaker™ software (Transgenomic, Inc.). Les produits de PCR ayant une taille entre 234 bp et 522 bp étaient dénaturés pendant 5 min à 94°C. La température était ensuite diminuée progressivement jusqu'à 25°C pendant 30 minutes afin de permettre la formation des homo/hétéro duplexes. Les produits étaient ensuite élués dans une colonne avec une phase mobile contenant de l'acétate de triéthylammonium (TEAA). Un détecteur d'UV (260nm) mesurait l'absorption des différents fragments élués. Les hétérodimères étaient éliminés en premiers.

5. Séquençage de l'ADN :

En cas de présence d'un hétérodimère, le séquençage de l'ADN était réalisé. Les produits de PCR étaient purifiés selon le protocole EXO-SAP et séquencé au laboratoire. Les séquences étaient ensuite analysées et comparées aux séquences publiées de *CCND1* (GenBank: NM_053056 et NT_167190). Tous les variants identifiés étaient recherchés chez les cas et les contrôles. Nous avons également testé 12 polymorphismes publiés (fréquence <3%, HapMap European (http://www.ncbi.nlm.nih.gov/snp) chez les cas et les contrôles.

Les patients provenant de Saint-Antoine ont été testés pour 5 polymorphismes (2 au niveau du promoteur (CCND1-3, 7), et 3 au niveau des exons (CCND1-19, 21 et 30). Le génotypage de tous les variants était réalisé par analyse de spectrométrie de masse (MALDI-TOF).

6. Analyse statistique :

L'équilibre d'Hardy-Weinberg était vérifié au sein des groupes contrôles. La comparaison des fréquences alléliques était faite à l'aide du test exact de Fisher. Les tests étaient bilatéraux avec un risque alpha de 0,05. Les comparaisons entre les différents groupes ont été faites par le test du Chi2. L'ensemble des tests a été réalisé à l'aide des logiciels

PowerMarker version 3.25 (Liu, K *et al.* 2005) et SPSS version 16.0 (SPSS Inc., Chicago, IL, USA). L'analyse des déséquilibres de liaison était réalisée en utilisant le logiciel Haploview version 4.1 (Barrett, JC *et al.* 2005).

7. Analyse fonctionnelles in silico :

L'étude de l'influence des variants sur la fonction de CCND1 était évaluée *in silico* en simulant l'impact sur les modifications de transcriptions et de l'épissage à l'aide des logiciels AliBaba 2.1 et Human Splicing Finder (HSF) (Desmet, FO *et al.* 2009). Pour étudier les effets des variants sur les extrémités 5' et 3', la database UTRdb était consultée (utrdb.ba.itb.cnr.it). Enfin, le logiciel GenEpi toolbox (Coassin, S *et al.* 2010) était utilisé pour examiner la conservation de la séquence dans les espèces.

RESULTATS :

1. Population et fréquence allélique des variants :

Les caractéristiques des patients et des témoins sont détaillées dans le tableau 18.

	N	Age moyen (années)	Homme:femme	Nb moyen de polypes
Polypose UK	112	59[a]	68:20[b]	11[b]
CCR précoce UK	42	41	24:18	n/a
Polypose Fr	75	51[c]	44:31	26[d]
CCR précoce Fr	56	40	24:32	n/a
Groupe contrôle	866	62	478:382[e]	n/a

Tableau 18. Caractéristiques cliniques et endoscopiques des patients et des contrôles.

Données manquantes pour : [a]33, [b]24, [c]3, [d]14, [e]6 individus. n/a=non applicable

La fréquence des variants au sein du groupe contrôle validait les conditions de l'équilibre d'Hardy-Weinberg. La liste des variants identifiés par dHPLC et sélectionnés dans la database dbSNO est détaillée dans le tableau 19. Les fréquences de chaque variants au sein du groupe témoin et patients sont données dans le tableau 20.

id	dbSNP	Position chromosome	Avant trans-cription	Avant translation	Allèles	fréquence	méthode	génotypage	variation	Analyse
CCND1-3	rs36225071	69163123	-1931	-2140	G/C	n/a	dbSNP	UK&French	Cas & contrôles	inclus
CCND1-5	rs36225067	69163166	-1888	-2097	A/C	n/a	dbSNP	UK cases	Cas & n/a	exclus
CCND1-6	rs36225068	69163169	-1885	-2094	G/T	n/a	dbSNP	UK cases	invariant cas & n/a	exclus
CCND1-7	rs36225069	69163317	-1737	-1946	A/T	n/a	dbSNP	UK&French	Cas & contrôles	inclus
CCND1-10	rs36225073	69163457	-1597	-1806	G/A	n/a	dbSNP	UK	Cas & contrôles	inclus
CCND1-11	rs36225394	69163500	-1554	-1763	A/G	n/a	dbSNP	UK	invariant	exclus
CCND1-13	rs3212859	69163888	-1166	-1375	G/T	0.033	dbSNP	UK	Cas & contrôles	inclus
CCND1-14	rs954618	69163891	-1163	-1372	A/G	0	dbSNP	UK	invariant	exclus
CCND1-15	rs954619	69163915	-1139	-1348	C/T	0	dbSNP	UK	Contrôles	inclus
CCND1-17	n/a	69164259	-795	-1004	T/C	n/a	dHPLC	UK	Cas & contrôles	inclus
CCND1-18	n/a	69164806	-248	-457	T/C	n/a	dHPLC	UK	Cas & contrôles	inclus
CCND1-16	rs3212862	69165001	-53	-262	C/A	0.006	dbSNP	UK	invariant	exclus
CCND1-19	rs55911137	69165181	128	-82	G/C	n/a	dHPLC	UK&French	cases	inclus
CCND1-1	rs1050971	69167169	597	Ex2-27 N130D	A/G	n/a	dbSNP	UK	invariant	exclus
CCND1-2	rs1131439	69167170	598	Ex2-26 N130S	A/G	n/a	dbSNP	UK	invariant	exclus
CCND1-22	n/a	69171856	697	IVS3-87	G/C	n/a	dHPLC	UK	cases	inclus
CCND1-20	rs3862792	69172037	878	Ex4-55 F223F	C/T	0.033	dHPLC	UK	Cas & contrôles	inclus
CCND1-21	rs9344	69172091	932	Ex4-1 P241P	G/A	0.517	dHPLC	UK&French	Cas & contrôles	inclus
CCND1-23 [a]	rs678653	69175918	1784	Ex5+852	C/G	0.319	dHPLC	UK cases	Cas & n/a	exclus
CCND1-24 [a]	n/a	69176073	1939	Ex5+1007	G/A	n/a	dHPLC	UK cases	Cas & n/a	exclus
CCND1-25[a]	n/a	69176371	2237	Ex5+1305	T/G	n/a	dHPLC	UK cases	Cas & n/a	exclus
CCND1-26[a]	n/a	69176880	2475	Ex5-1543	G/A	n/a	dHPLC	UK cases	Cas & n/a	exclus
CCND1-27	rs3212906	69177849	3715	Ex5-574	T/A	0.006	dHPLC	UK	Cas & contrôles	inclus
CCND1-28	rs55666306	69177952	3818	Ex5-471	G/A	n/a	dHPLC	UK	Cas	inclus
CCND1-29	rs3212907	69177957	3823	Ex5-466	C/T	0.050	dHPLC	UK	Cas & contrôles	inclus
CCND1-30	rs7178	69178211	4077	Ex5-213	A/G	0.100	dHPLC	UK&French	Cas & contrôles	exclus

Tableau 19. Détails des variants testés sur CCND1

[a]non validé par génotypage
n/a : non disponible
Méthode: dbSNP : données obtenues avec la database ; dHPLC : variant identifié par dHPLC.

id	dbSNP	allèle majeur/ mineur	Cas UK	Cas Fr	Contrôles
CCND1-1	rs1050971	A/G	0,000	n/t	0,000
CCND1-2	rs1131439	A/G	0,000	n/t	0,000
CCND1-3[a]	rs36225071	G/C	0,000	0,012	0,002
CCND1-7	rs36225069	A/T	0,016	0,008	0,005
CCND1-10	rs36225073	G/A	0,030	n/t	0,024
CCND1-11	rs36225394	A/G	0,000	n/t	0,000
CCND1-13	rs3212859	G/T	0,028	n/t	0,023
CCND1-14	rs954618	A/G	0,000	n/t	0,000
CCND1-15	rs954619	C/T	0,000	n/t	0,001
CCND1-16	rs3212862	C/A	0,000	n/t	0,000
CCND1-17	n/a	T/C	0,003	n/t	0,007
CCND1-18	n/a	T/C	0,013	n/t	0,008
CCND1-19	rs55911137	G/C	0,005	0,000	0,000
CCND1-22	n/a	G/C	0,003	n/t	0,000
CCND1-20	rs3862792	C/T	0,028	n/t	0,024
CCND1-21	rs9344	G/A	0,463	0,504	0,441
CCND1-27	rs3212906	T/A	0,007	n/t	0,006
CCND1-28	rs55666306	G/A	0,003	n/t	0,000
CCND1-29	rs3212907	C/T	0,034	n/t	0,049
CCND1-30[a]	rs7178	A/G	0,071	0,074	0,051

Tableau 20. Fréquence alléliques des variants *CCND1* chez les patients et les contrôles.

[a] testés dans un sous groupe de 222 contrôles, n/a=non disponible, n/t=non testés

La comparaison des fréquences alléliques selon les groupes (tableau 21) (cancer colorectal précoce ou polypose indéterminée) entre les patients anglais ou français ne m'était pas en évidence d'hétérogénéité (à l'exception du variant CCND1-3) (p=0,04). Ce résultat justifiait l'association des cas français et anglais pour une analyse globale dans les deux groupes (cancer colorectal précoce et polypose indéterminée).

id	dbSNP	polypes UK	CCR UK	Polypes Fr	CCR Fr	Polypes global	CCR global
CCND1-3	rs36225071	0/109	0/41	1/75	2/56	1/184	2/97
CCND1-7	rs36225069	3/109	2/42	1/75	1/55	4/184	3/97
CCND1-10	rs36225073	6/104	2/42				
CCND1-13	rs3212859	5/101	2/41				
CCND1-15	rs954619	0/109	0/42				
CCND1-17	n/a	1/108	0/42				
CCND1-18	n/a	4/107	0/42				
CCND1-19	rs55911137	1/77	0/30	0/75	0/55	1/152	0/85
CCND1-20	rs3862792	5/100	2/42				
CCND1-22	n/a	1/109	0/42				
CCND1-27	rs3212906	1/104	1/42				
CCND1-28	rs55666306	1/103	0/42				
CCND1-29	rs3212907	9/105	1/42				

Tableau 21. Nombre de patients porteurs d'un variant CCND1 rare ou intermédiaire.

polypes=polypose indéterminée ; CCR=CCR précoce ; n/a=not disponible

2. Variants rares, intermédiaires ou communs :

Vingt variants furent testés sur les cas et les contrôles anglais : 10 (50%) par identification sur les séquences analysées et 10 grâce à la database dbSNP. Quatre variants supplémentaires ont été identifiés par dHLPC (CCND1-23, 24, 25, 26) mais leur séquençage par le Sequenom n'a pas été possible. De même, deux autres variants (CCND1-5, 6) identifiés sur la database ne furent pas inclus dans l'analyse car impossible à séquencer.

Quatre variants étaient communs (fréquence dans la population contrôle > 10%) dont rs9344 : (CCND1-21, 23, 30 et rs7177). Les trois derniers furent exclus de l'analyse. De même les polymorphismes invariants dans les cohortes de cas et de contrôles (CCND1-1, 2, 11, 14, 16) furent également exclus. En revanche, les polymorphismes monomorphiques dans le population de témoins (CCND1-19, 22, 28) ou des cas (CCND1-15) furent inclus dans l'analyse. Quatorze variants (avec rs9344) furent donc étudiés : 13 rares ou intermédiaires et un variant commun. Six variants découverts par dHPLC n'étaient pas présents dans la database dbSNP (CCND1-17, 18, 22, 24, 25, 26). Trois variants (CCND1-13, 20 and 29)

avaient une fréquence reportée entre 3 et 5% et 2 (CCND1-16, 27) avaient une fréquence inférieure à 1%. (Tableau 21)

3. Etude d'association entre les variants :

Un variant (CCND1-7) présentait une fréquence significativement différente entre les cas (en regroupant les polyposes et les cancers précoces) et les témoins (p=0,03; OR 3,8, 95% CI 1,2-12,1) (Tableau 22).

id	dbSNP	allèle majeur/ mineur	Cas [b]	Contrôle[b]	p
CCND1-3[a]	rs36225071	G/C	0/150	1/222	0.63
CCND1-7	rs36225069	A/T	5/151	7/779	0.03
CCND1-10	rs36225073	G/A	8/146	35/743	0.67
CCND1-13	rs3212859	G/T	7/142	34/749	0.83
CCND1-15	rs954619	C/T	0/151	1/747	1.00
CCND1-17	n/a	T/C	1/150	11/736	0.70
CCND1-18	n/a	T/C	4/149	12/730	0.33
CCND1-19	rs55911137	G/C	1/107	0/732	0.24
CCND1-20	rs3862792	C/T	7/142	35/743	0.83
CCND1-22	n/a	G/C	1/151	0/750	0.17
CCND1-27	rs3212906	T/A	2/146	9/720	1.00
CCND1-28	rs55666306	G/A	1/145	0/749	0.16
CCND1-29	rs3212907	C/T	10/147	70/747	0.43

Tableau 22. Fréquence des porteurs de variants rares ou intermédiaires de CCND1 entre le groupe patient et le groupe contrôle.

[a] séquencés chez 222 contrôles, [b] nombre de cas ou contrôles portant l'allèle mineur / nombre total d'individus séquencés avec succès ; n/a=non disponible

Lorsque tous les variants rares étaient regroupés, une association significative était retrouvée (OR >2) entre l'ensemble des cas anglais et les contrôles ainsi que pour les cas de polyposes indéterminées (Tableau 23). Cette tendance n'était pas retrouvée pour les cas Français.

UK [a]	N	≥ 1 variant rare	Aucun variant rare	OR	95% CI	p
polypes	107	12	95	2,2	1,1-4,4	0,03
CCR	42	3	39	1,4	0,4-4,6	0,50
Tous les cas	149	15	134	2,0	1,1-3,7	0,04
contrôles 1	743	40	703	référence		
contrôles 2	213	8	205	référence		
Fr [b]						
polypes	75	2	73	0,7	0,1-3,4	1,00
CCR	55	3	52	1,5	0,4-5,8	0,70
Tous les cas	130	5	125	1,0	0,3-3,2	1,00

Tableau 23. Comparaison du nombre de porteurs de variants entre les cas anglais et français.

8 rares variants ont été testés sur les cas anglais (CCND1-7, 15, 17, 18, 19, 22, 27 et 28), 3 sur les échantillons français (CCND1-3, 7 et 19). Le nombre total d'individus était estimé à l'aide d'une moyenne harmonique afin de prendre en compte les variations du nombre de cas testés pour chaque variant.

[a] comparaison avec le groupe de contrôle 1

[b] comparaison avec le groupe de contrôle 2

Il y avait un déséquilibre de liaison important sur le locus comportant les polymorphismes CCND1-10, 13 and 20 ($r^2 > 0,8$, pour chaque paire de variants) avec les trois allèles mineurs sur le même haplotype (ATT) à l'exception d'un seul patient qui présentait l'haplotype ATC. Ces trois polymorphismes ont donc été associés et analysés comme un seul variant rare. L'analyse de cet agrégat ATT et des variants restants ne mettait plus en évidence de différence significative que ce soit en limitant l'analyse avec les variants rares ou intermédiaires (tableaux 24 et 25).

	N	≥ 1 variant rare	Aucun variant rare	OR	95% CI	p
Tous cas UK	143	31	112	1,2	0,8-1,8	0,49
Contrôles	744	142	602	référence		

Tableau 24. Patients anglais ayant des variants avec une fréquence allélique <0,05.

Neuf variants indépendants et 3 combinés en un haplotype (CCND1-10, 13 et 20) ont été inclus dans cette analyse. Le nombre total d'individus était estimé par une moyenne harmonique.

	N	haplotype ATT	autres haplotypes	OR	95% CI	p
Tous cas UK	147	6	141	0.9	0.4-2.1	1.00
Contrôles	751	35	716	référence		
		ATT ou CCND1-29*T	autres haplotypes et CCND1-29*C			
Tous cas UK	147	16	131	0.8	0.4-1.3	0.36
Contrôles	749	104	645	référence		

Tableau 25. Patients anglais ayant des variants intermédiaires avec une fréquence allélique entre 1 et 5%.

Un variants indépendant et un combiné en un haplotype (CCND1-10, 13 et 20) ont été inclus dans cette analyse. Le nombre total d'individus était estimé par une moyenne harmonique.

4. Analyse de la perte de fonction :

Les conséquences des polymorphismes sur la fonction protéique ont été estimées par l'approche « *in silico* ». Le logiciel Alibaba 2.1. a permis d'analyser les variants du promoteur : certains variants (CCND1-7, 10, 17) altéraient le site de fixation du facteur de transcription Sp1 appartenant à la famille Sp/KLF, CCND1-3 altérait une boite CACCC qui est également un site de fixation pour les facteurs de transcription Sp/KLF et CCND1-13 créait un site Cdx2. CCND1-18 ne semblait pas avoir d'influence sur les site de liaison des facteurs de transcription.

Pour les variants situés dans la zone codante du gène, l'analyse a porté sur la création ou l'élimination d'un site d'épissage alternatif. L'allèle mineur rare C du variant CCND1-22 est sensé, selon le logiciel HSF, modifier un site de liaison des protéines SR (en activant un ESE (*exonic splicing enhancer*) et en interrompant un ESS (*exonic splicing silencer)*. Le variant re9344 est également connu pour modifier un site d'épissage alternatif et prédisposer à la création d'un transcrit codant pour la cycline D1b. Sur ce transcrit, l'exon 5 n'est pas présent (Howe, D *et al.* 2001).

La majorité des variants situés aux extrémités 3' et 5' ne semblaient pas modifier l'épissage à l'exception de CCND1-19 : selon la database UTRdb il s'agit d'un

polymorphisme silencieux mais le logiciel HSF indique qu'il pourrait être responsable de la création d'un site ESE.

L'ensemble des modifications prévisibles est détaillé dans le tableau 26.

id	Allèle mineur	Position sur le gène	Effet	Conservé
CCND1-3	C	-2140	-CACCC-bi	Oui
CCND1-7	T	-1946	-Sp1/+Sp1	Oui
CCND1-10	A	-1806	+Sp1	Oui
CCND1-13	T	-1375	+Cdx2	Oui
CCND1-15	T	-1348	-CDP-CR1	Oui
CCND1-17	C	-1004	-Sp1	n/a
CCND1-18	C	-457	Pas de modification	n/a
CCND1-19	C	-82	Nouveau ESE	n/a
CCND1-22	C	intron 3	Site de fixation altéré/ nouveau ESE/ESS altéré	n/a
CCND1-20	T	exon 4	synonyme	Oui
CCND1-21	A	exon 4	Nouveau site de fixation/élimination du site donneur	Oui
CCND1-27	A	3'UTR		Oui
CCND1-28	A	3'UTR		n/a
CCND1-29	T	3'UTR		Oui
CCND1-30	G	3'UTR		Oui

Tableau 26. Effet présumés des variants CCND1 déterminé *in silico.*

CACCC, Sp1, Cdx2, CDP-CR1 = facteurs de transcription
ESE = exonic splicing enhancer, ESS = exonic splicing silencer
+/- = gain ou perte du site d'action correspondant au facteur de transcription/
n/a = information non disponible

5. Le variant rs9344 G/A

Le variant rs9344 était plus fréquemment retrouvé chez les cas Anglais et Français que dans la population des témoins. Ce résultat était cependant à la limite de la significativité (Tableau 20). Le groupe de cas français avec un cancer colorectal précoce était celui avec la plus grande fréquence (0,53). Cette incidence était néanmoins comparable avec les données publiées concernant la population européenne. En revanche, les cas Anglais présentaient une fréquence plus faible que le reste de la population européenne en relation avec les résultats de la British 1958 Birth Cohort (fréquence de l'allèle A = 0,44).

L'étude des déséquilibres de liaison à l'aide du logiciel Haploview entre rs9344 et les variants rares ou intermédiaires des populations françaises et anglaises indique que le déséquilibre était faible. Le coefficient de corrélation entre l'haplotype des 3 variants CCND1-10, 13 et 20 et rs9344 était de 0,02. L'étude du tableau de contingence des fréquences de l'haplotype et de rs9344 mettait cependant en évidence un déséquilibre de liaison significatif (Tableau 27).

CCND1-10, 13, 20\CCND1-21	GG	GA/AA	P
haplotype ATT	19	14	0,001
Autres haplotypes	199	489	

Tableau 27. Déséquilibre de liaison entre les haplotypes des loci CCND1-10,13 et 20 et l'allèle CCND1-21 (rs9344).

(rs9344), r^2=0,02, r=0,14, OR=3,3, 95% CI=1,6-6,8.

Le polymorphisme rs9344 n'est pas un facteur indépendant prédisposant au cancer. Des évènements additionnels sont nécessaires pour induire la production de la Cycline D1b. Nous avons donc étudié la relation entre la présence de variants rares sur le gène *CCND1* et l'existence du polymorphisme rs9344. Les patients Français ou Anglais ayant au moins un variant rare possédaient également plus fréquemment le polymorphisme rs9344 que les autres. La même observation était faite dans la population contrôle bien que l'association était plus faible (Tableau 28). Cette association est néanmoins possiblement liée au déséquilibre gamétique.

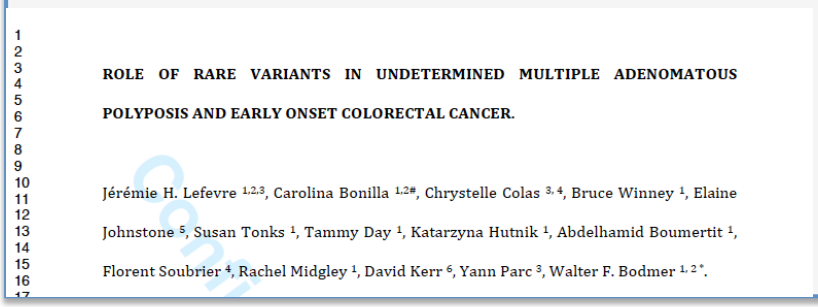
	GG	GA/AA	OR	95% CI	p
Cas UK ≥ 1 rv	6	8	2,2	0,7-6,8	0,21
Cas UK sans rv	34	100			
Cas français ≥ 1 rv	2	3	2,1	0,3-13,1	0,60
Cas français sans rv	30	94			
Tous les cas ≥ 1 rv	4	7	1,7	0,5-5,9	0,48
Tous les cas sans rv	68	198			
contrôles ≥ 1 rv	15	25	1,4	0,7-2,7	0,38
contrôles no rv	218	510			

Tableau 28. Répartition des patients selon la présence de variants rares et l'haplotype du variant

CCND1-21 (rs9344),

8 variants rares ont été testés dans le groupe UK, 3 dans le groupe français et dans le groupe global. rv = variant rare

ROLE DES VARIANTS RARES DANS LES POLYPOSES INDETERMINEES ET LES CANCERS COLORECTAUX PRECOCES.



ROLE OF RARE VARIANTS IN UNDETERMINED MULTIPLE ADENOMATOUS POLYPOSIS AND EARLY ONSET COLORECTAL CANCER.

Jérémie H. Lefevre [1,2,3], Carolina Bonilla [1,2#], Chrystelle Colas [3, 4], Bruce Winney [1], Elaine Johnstone [5], Susan Tonks [1], Tammy Day [1], Katarzyna Hutnik [1], Abdelhamid Boumertit [1], Florent Soubrier [4], Rachel Midgley [1], David Kerr [6], Yann Parc [3], Walter F. Bodmer [1, 2*].

Annexe F, page 187

OBJECTIF:

Etudier le rôle des variants rares (<1% de fréquence) ou intermédiaire (entre 1-5%) de différents gènes au sein d'une population de cancer colorectaux de survenue précoce (<50 ans) ou de polypose indéterminée (sans mutation identifiée sur *APC* ou *MUTYH*).

METHODES :

1. Sélection des patients et extraction de l'ADN:

Elles ont été décrites dans l'article précédent.

2. Sélection des variants:

Les variants rares ou intermédiaires étaient choisis selon les données de la littérature. Les variants associés à des cancers digestifs, colorectaux ou d'autres affections néoplasiques (principalement le sein et la prostate) ont donc été analysés. Les variants des gènes *BCRA* considérés comme non pathogéniques ou de signification indéterminée également. Quelques variants communs (*CDH1* rs16260, *MTHFR* rs1801133 et *TP53* rs1042522) ont été également séquencés car ils étaient considérés comme augmentant le risque de CCR. .

3. Génotypage:

77 variants ont été séquencés sur la population de patients et les contrôles. Ils appartenaient à 17 gènes. Parmi ces variants, 16 n'ont été séquencés que sur 227 contrôles. La liste complète des variants analysés est donnée dans le tableau 29.

Chromosome	variant	Génétique	dbSNP	Allèle maj	Allèle mineur	MAF cases[a]	MAF controls[a]	Type
1	MTHFR-1	A222V	rs1801133	C	T	0.338	0.336	Polymorphisme
1	EPHB2-1	R80H	n/a	G	A	0.003	0.001	Variant rare
1	EPHB2-3	I361V	rs56180036	A	G	0.007	0.001	Variant rare
1	EPHB2-4	R569W	n/a	C	T	0.003	0.001	Variant rare
1	*EPHB2-7*	M883V	n/a	A	G	0.003	0.000	Variant rare
1	EPHB2-8	K1019X	n/a	A	T	0.000	0.000	Variant rare
1	EXO1-2	E109K	n/a	G	A	0.000	0.001	Variant rare
1	EXO1-12	D249N	rs61750993	G	A	0.020	0.007	Variant rare
1	*EXO1-4*	L410R	n/a	T	G	0.004	0.000	Variant rare
1	*EXO1-10*	G759E	rs4150001	G	A	0.003	0.009	Variant rare
2	*MSH2-8*	E808X	rs34986638	G	T	0.000	0.000	Variant rare
3	MLH1-1	G22A	rs41295280	G	C	0.003	0.000	Variant rare
3	MLH1-3	R182G	n/a	A	G	0.000	0.000	Variant rare
3	CTNNB1-1	N287S	rs35288908	A	G	0.003	0.000	Variant rare
5	APC-1	R141S	n/a	G	T	0.000	0.000	Variant rare
5	APC-3	R653G	n/a	A	G	0.000	0.000	Variant rare
5	APC-6	N1118D	n/a	A	G	0.000	0.000	Variant rare
5	APC-7	L1129S	n/a	T	C	0.000	0.003	Variant rare
5	APC-8	T1292M	n/a	C	T	0.000	0.000	Variant rare
5	APC-10	I1307K	rs1801155	T	A	0.000	0.000	Variant rare
5	APC-11	E1317Q	rs1801166	G	C	0.011	0.007	Variant rare
5	APC-12	D1714N	n/a	G	A	0.000	0.000	Variant rare
5	APC-13	S1970R	n/a	A	C	0.000	0.000	Variant rare
5	APC-14	A2119V	n/a	C	T	0.000	0.000	Variant rare
5	APC-15	G2502S	rs2229995	G	A	0.010	0.021	Intermédiaire
5	APC-16	R2505Q	n/a	G	A	0.000	0.002	Variant rare
5	*APC-17*	S2621C	rs72541816	C	G	0.003	0.011	Intermédiaire

Chromosome	variant	Génétique	dbSNP	Allèle maj	Allèle mineur	MAF cas [a]	MAF contrôles[a]	Type
5	APC-18	A2690T	n/a	G	A	0.000	0.000	Variant rare
5	APC-19	8489 delA	n/a	A	del	0.000	0.000	Variant rare
5	*APC-20*	8636 C/A	n/a	C	A	0.024	0.019	Intermédiaire
7	PMS2-1	T511A	rs2228007	A	G	0.023	0.029	Intermédiaire
7	*PMS2-2*	T597S	rs1805318	A	T	0.034	0.020	Intermédiaire
7	PMS2-3	M622I	rs1805324	G	A	0.017	0.022	Intermédiaire
13	BRCA2-7	N372H	rs144848	T	G	0.320	0.286	Polymorphisme
13	BRCA2-8	S384F	rs41293475	C	T	0.000	0.003	Variant rare
13	BRCA2-27	R2034C	rs1799954	C	T	0.004	0.004	Variant rare
13	BRCA2-30	V2138F	rs11571659	G	T	0.000	0.000	Variant rare
13	BRCA2-33	I2490T	rs11571707	T	C	0.000	0.000	Variant rare
13	*BRCA2-35*	D2665G	rs28897745	A	G	0.003	0.000	Variant rare
13	BRCA2-37	V2728I	rs28897749	G	A	0.014	0.003	Variant rare
13	*BRCA2-51*	K2729N	n/a	G	T	0.000	0.000	Variant rare
13	*BRCA2-48*	P3194Q	rs28897760	C	A	0.000	0.000	Variant rare
14	*MLH3-1*	V741F	rs28756990	G	T	0.014	0.004	Variant rare
14	*MLH3-2*	M809V	rs61752722	A	G	0.003	0.009	Variant rare
14	MLH3-4	S817R	n/a	A	C	0.000	0.000	Variant rare
14	*MLH3-5*	S845G	rs28756992	A	G	0.014	0.033	Intermédiaire
16	AXIN1-4	D495E	n/a	C	G	0.004	0.009	Variant rare
16	AXIN1-6	R841Q	rs34015754	G	A	0.007	0.007	Variant rare
16	CDH1-1	-1128	rs13335980	A	T	0.000	0.001	Variant rare
16	*CDH1-2*	-284	rs16260	C	A	0.246	0.308	Polymorphisme
16	CDH1-3	IVS1+6	rs3743674	T	C	0.102	0.121	Polymorphisme
16	CDH1-4	Ex2 49-2	n/a	A	G	0.000	0.000	Variant rare
16	CDH1-5	IVS4+10	rs33963999	G	C	0.070	0.080	Polymorphisme
16	CDH1-6	T340A	rs116093741	A	G	0.000	0.000	Variant rare
16	CDH1-7	A592T	rs35187787	G	A	0.003	0.007	Variant rare

Chromosome	variant	Génétique	dbSNP	Allèle maj	Allèle mineur	MAF cases[a]	MAF controls[a]	Type
16	CDH1-8	T599S	n/a	C	G	0.000	0.001	Variant rare
16	CDH1-9	A617T	rs33935154	G	A	0.000	0.000	Variant rare
16	CDH1-10	A634V	n/a	C	T	0.000	0.001	Variant rare
17	TP53-1	R72P	rs1042522	G	C	0.255	0.248	Polymorphisme
17	BRCA1-6	Q356R	rs1799950	A	G	0.042	0.060	Polymorphisme
17	BRCA1-8	R496H	rs28897677	G	A	0.000	0.001	Variant rare
17	BRCA1-14	T790A	rs41286298	A	G	0.000	0.000	Variant rare
17	BRCA1-16	T826K	rs28897683	C	A	0.000	0.001	Variant rare
17	BRCA1-22	E1038G	rs16941	A	G	0.338	0.342	Polymorphisme
17	BRCA1-26	E1250K	rs28897686	G	A	0.000	0.000	Variant rare
17	*BRCA1-28*	S1512I	rs1800744	G	T	0.003	0.004	Variant rare
17	BRCA1-32	M1628T	rs4986854	T	C	0.000	0.000	Variant rare
17	AXIN2-1	N412S	rs115931022	A	G	0.000	0.007	Variant rare
18	SMAD4-1	A118A	n/a	G	A	0.003	0.005	Variant rare
22	*CHEK2-1*	1100 delC	n/a	C	del	0.008	0.003	Variant rare

Tableau 29. Variants génotypés dans les groupes de patients anglais et les contrôles.

Italique: variant séquencé dans le sous-groupe de 227 témoins.
Soulignés : variant analysés.
n/a: non disponible.
[a]MAF: fréquence allélique mineure

Quatre variants ont été analysés à l'aide d'enzyme de restriction (CDH1-2, CHEK2-1, CTNNB1-1, MSH2-8). Deux variants d'*APC* (APC-10 : I1307K et APC-11 : E1317Q) ont été étudiés à l'aide de deux PCR allèles spécifiques comme précédemment publié (Fearnhead, NS et al. 2004).

Les séquences des amorces et les détails des produits de PCR sont donnés dans le tableau 30. Les variants restants ont été séquencés selon la technique de spectrométrie de masse MALDI-TOF (iPLEX Gold assay (Sequenom Inc., San Diego, CA, USA)) (Sauer, S et al. 2002).

4. Analyse statistique :

L'équilibre d'Hardy-Weinberg était vérifié au sein des groupes contrôles en utilisant le logiciel PLINK v.1.07 (Purcell, S *et al.* 2007). Les études d'associations entre les cas et les témoins étaient réalisées avec le même logiciel. La comparaison des fréquences alléliques était faite à l'aide du test exact de Fisher. Les tests étaient bilatéraux avec un risque alpha de 0,05. Le calcul des odd-ratios combinés était fait en utilisant le test de Mantel-Haenszel.

5. Analyse fonctionnelles in silico :

L'étude de l'influence des variants sur la fonction sur la protéine des différents variants était réalisée grâce à des logiciels disponibles sur internet : PolyPhen-2 et SNPs&GO. FastSNP et F-SNP étaient également utilisés pour les variants situés sur des portions non-codantes. (Yuan, HY *et al.* 2006; (Lee, PH *et al.* 2008).

variant id	variant	dbSNP	major/ minor allele	Amorce sens (5'-3')	Amorce Anti-sens (5'-3')	Enzyme	fragment (bp)	Fragments coupé (bp)	Allèle coupant
CDH1-2	-284	rs16260	C/A	CCAGGTCTTAGTGAGCCACC	CCACCCGGCCTCGCATAGAC	HincII	133	25+108	A
CHEK2-1	1100delC	n/a	C/del	CCCTTTTGTACTGAATTTTAGA*G*TA	ACAAGAAACTTCAGGCGCCAAG TAG	ScaI	116	24+92	C
CTNNB1-1	N287S	rs35288908	A/G	AAAATGGCAGTGCGTTTAGC	TGTCCAATGCTCCATGAAAA	DdeI	173	49+124	G
MSH2-8	E808X	rs34986638	G/T	TGGTGCTTTTTGCATGTTTG	TCCTACCCCCAGACTGTGAA	AflII	291	99+192	T
APC-10	I1307K	rs1801155	T/A	AGAAAAGATTGGAACTAGGTCAG CTG	TGAGTGGGGTCTCCTGAACATA		251		
APC-11	E1317Q	rs1801166	G/C	AGAAAAGATTGGAACTAGGTCAG CTC	TGAGTGGGGTCTCCTGAACATA		223		

Tableau 30. Séquence des amorces et enzymes utilisés pour le génotypage des variants par RFLP.

RESULTATS:

1. Populations :

La population était similaire à celle de l'étude précédente (tableau 18).

2. Sélection des variants :

Parmi les 70 variants sélectionnés, 24 (34%) étaient monomorphiques dans les cas et les contrôles anglais et n'ont donc pas été utilisés dans l'analyse. Sur les 46 variants restants, 10 étaient monomorphiques dans la population de patients et 5 uniquement dans le groupe contrôle.

Trente et un variants avaient une fréquence allélique inférieure à 1% dans la population contrôle et ont donc été considérés comme des variants rares. Sept étaient considérés comme intermédiaires et 8 étaient des variants communs (fréquence allélique > 5%). Si la fréquence était déterminée en utilisant la population combinée témoins + patients, seul un variant (APC-17) passait du groupe intermédiaire à rare.

Tous les variants dans le groupe contrôle satisfaisaient l'équilibre d'Hardy-Weinberg et ce en utilisant une correction de Bonferroni (p=0,001 (0,05/46). Sans appliquer la correction de Bonferroni, 2 variants étaient en déséquilibre dans le groupe contrôle (TP53-1 et BRCA1-6, $p<0,05$) et 3 dans le groupe patients (EPHB2-3, EXO1-12 et BRCA1-22, $p<0,05$).

3. Analyse d'association des variants :

3.1. Comparaison cas vs. contrôles dans la population anglaise.

L'étude des fréquences alléliques entre ces deux groupes a mis en évidence quatre différences significatives pour les variants (EXO1-12, MLH1-1, CTNNB1-1 et BRCA2-37, $p<0,05$; Tableau 31). Le variant EXO2-12 était plus fréquent chez les patients avec un CCR de survenue précoce alors que les 3 autres variants étaient observés dans la population des polyposes indéterminées.

id	variant	dbSNP	allèle majeur/ mineur	MAF cas	MAF contrôles	p [a]	MAF polypes	MAF CCR	p [b]
MTHFR-1	A222V	rs1801133	C/T	0.338	0.336	0.95	0.337	0.342	0.94
EPHB2-1	R80H	n/a	G/A	0.003	0.001	0.19	0.005	0.000	0.53
EPHB2-3	I361V	rs56180036	A/G	0.007	0.001	0.07	0.010	0.000	0.37
EPHB2-4	R568W	n/a	C/T	0.003	0.001	0.19	0.000	0.012	0.11
EPHB2-7	M883V	n/a	A/G	0.003	0.000	0.22	0.005	0.000	0.53
EXO1-2	E109K	n/a	G/A	0.000	0.001	0.66	0.000	0.000	n/a
EXO1-12	D249N	rs61750993	G/A	0.020	0.007	0.03	0.018	0.024	0.74
EXO1-4	L410R	n/a	T/G	0.004	0.000	0.18	0.005	0.000	0.54
EXO1-10	G759E	rs4150001	G/A	0.003	0.009	0.38	0.000	0.013	0.11
MLH1-1	G22A	rs41295280	G/C	0.003	0.000	0.03	0.005	0.000	0.53
CTNNB1-1	N287S	rs35288908	A/G	0.003	0.000	0.05	0.005	0.000	0.42
APC-7	L1129S	n/a	T/C	0.000	0.003	0.37	0.000	0.000	n/a
APC-11	E1317Q	rs1801166	G/C	0.011	0.007	0.46	0.014	0.007	0.59
APC-15	G2502S	rs2229995	G/A	0.010	0.021	0.19	0.009	0.012	0.83
APC-16	R2505Q	n/a	G/A	0.000	0.002	0.43	0.000	0.000	n/a
APC-17	S2621C	rs72541816	C/G	0.003	0.011	0.25	0.000	0.012	0.12
APC-20	8636 C/A	n/a	C/A	0.024	0.019	0.66	0.024	0.025	0.96
PMS2-1	T511A	rs2228007	A/G	0.023	0.029	0.60	0.023	0.024	0.55
PMS2-2	T597S	rs1805318	A/T	0.034	0.020	0.23	0.043	0.012	0.19
PMS2-3	M622I	rs1805324	G/A	0.017	0.022	0.59	0.014	0.024	0.98
BRCA2-7	N372H	rs144848	T/G	0.320	0.286	0.23	0.329	0.298	0.60
BRCA2-8	S384F	rs41293475	C/T	0.000	0.003	0.38	0.000	0.000	n/a
BRCA2-27	R2034C	rs1799954	C/T	0.004	0.004	0.90	0.000	0.012	0.12
BRCA2-35	D2665G	rs28897745	A/G	0.003	0.000	0.22	0.005	0.000	0.53
BRCA2-37	V2728I	rs28897749	G/A	0.014	0.003	0.02	0.019	0.000	0.21
MLH3-1	V741F	rs28756990	G/T	0.014	0.004	0.16	0.005	0.036	0.04

id	variant	dbSNP	allèle majeur/ mineur	MAF cas	MAF contrôles	p [a]	MAF polypes	MAF CCR	p [b]
MLH3-2	M809V	rs61752722	A/G	0.003	0.009	0.37	0.005	0.000	0.53
MLH3-5	S845G	rs28756992	A/G	0.014	0.033	0.09	0.005	0.036	0.04
AXIN1-4	D495E	n/a	C/G	0.004	0.009	0.44	0.007	0.000	0.37
AXIN1-6	R841Q	rs34015754	G/A	0.007	0.007	0.92	0.010	0.000	0.45
CDH1-1	-1128	rs13335980	A/T	0.000	0.001	0.53	0.000	0.000	n/a
CDH1-2	-284	rs16260	C/A	0.246	0.308	0.07	0.218	0.291	0.12
CDH1-3	IVS1+6	rs3743674	T/C	0.102	0.121	0.37	0.116	0.071	0.27
CDH1-5	IVS4+10	rs33963999	G/C	0.070	0.080	0.57	0.051	0.119	0.04
CDH1-7	A592T	rs35187787	G/A	0.003	0.007	0.45	0.005	0.000	0.53
CDH1-8	T599S	n/a	C/G	0.000	0.001	0.65	0.000	0.000	n/a
CDH1-10	A634V	n/a	C/T	0.000	0.001	0.52	0.000	0.000	n/a
TP53-1	R72P	rs1042522	G/C	0.255	0.248	0.80	0.270	0.220	0.38
BRCA1-6	Q356R	rs1799950	A/G	0.042	0.060	0.22	0.043	0.038	0.53
BRCA1-8	R496H	rs28897677	G/A	0.000	0.001	0.65	0.000	0.000	n/a
BRCA1-16	T826K	rs28897683	C/A	0.000	0.001	0.65	0.000	0.000	n/a
BRCA1-22	E1038G	rs16941	A/G	0.338	0.342	0.90	0.346	0.317	0.64
BRCA1-28	S1512I	rs1800744	G/T	0.003	0.004	0.82	0.005	0.000	0.86
AXIN2-1	N412S	rs115931022	A/G	0.000	0.007	0.19	0.000	0.000	n/a
SMAD4-1	A118A	n/a	G/A	0.003	0.005	0.75	0.005	0.000	0.53
CHEK2-1	1100 delC	n/a	C/del	0.008	0.003	0.31	0.009	0.007	0.88

Tableau 31. Liste des variants analysés dans la population de patients anglais et les contrôles.

n/a : not disponible. MAF: fréquence allélique mineure ; Nombre en gras : p≤0.05.
[a]p pour la comparaison cas vs. contrôle ; [b] p pour la comparaison entre les polyposes et les CCR précoces.

Quatre autres variants avaient des résultats à la limite de la significativité : un variant rare : EPHB2-3 (p=0,07) et un variant communs CDH1-2 (p=0,07). L'allèle A de CDH1-2 semblait être un facteur protecteur du CCR. En passant d'une comparaison de fréquence allélique à une comparaison du nombre de personnes portant le variant, seul BRCA2-37 restait avec une différence significative et un odd-ratio de 4,1 (1,2-14,3, p=0,05). Les données sont détaillées dans le tableau 32.

Variant rare	Cas [a]	Contrôles [b]	p
EPHB2-1	1/149	1/778	0.30
EPHB2-3	1/145	2/746	0.41
EPHB2-4	1/149	1/775	0.30
EPHB2-7	1/149	0/224	0.40
EXO1-2	0/150	1/751	1.00
EXO1-4	1/125	0/226	0.36
EXO1-10	1/145	4/225	0.65
EXO1-12	5/150	10/745	0.15
MLH1-1	1/147	0/740	0.17
CTNNB1-1	1/174	0/702	0.20
APC-7	0/149	4/743	1.00
APC-11	4/176	11/745	0.50
APC-16	0/150	3/729	1.00
BRCA2-8	0/147	4/748	1.00
BRCA2-27	1/142	6/740	1.00
BRCA2-35	1/149	0/224	0.40
BRCA2-37	4/148	5/744	0.05
MLH3-1	4/144	2/225	0.21
MLH3-2	1/148	4/227	0.65
AXIN1-4	1/117	14/749	0.71
AXIN1-6	2/146	11/746	1.00
CDH1-1	0/150	2/748	1.00
CDH1-7	1/147	11/747	0.70
CDH1-8	0/151	1/742	1.00
CDH1-10	0/150	2/731	1.00
BRCA1-8	0/151	1/735	1.00
BRCA1-16	0/149	1/744	1.00
BRCA1-28	1/149	2/225	1.00
AXIN2-1	0/128	10/737	0.37
SMAD4-1	1/148	7/739	1.00
CHEK2-1	3/178	1/179	0.37

Tableau 32. Nombre de porteurs de variants rares.

En combinant ensemble tous les variants rares, la proportion de patients porteurs de ces polymorphismes était significativement plus importante que dans le groupe contrôle. Cette observation persistait quel que soit le groupe contrôle utilisé (227 ou 866) (Tableau 33).

L'odd-ratio combiné, estimé à l'aide du test de Mantel-Haenszel, en utilisant celui calculé avec l'effectif total des contrôles (OR1) et celui calculé avec le groupe réduit de contrôles était de 1,2 (95% CI, 0,8-1,8, p=0,42). L'effet des variants rares devenait beaucoup plus important en se limitant aux variants ayant une fréquence allélique inférieure à 0,5% : OR 1,8, 95% CI, 1,0-3,1, p=0,05; Tableau 33).

rare variant	Cas	Contrôles	p
MAF < 1% (n=31)			
porteurs/non porteurs 1	24/146	108/739	
OR 1 (95% CI)	1,13 (0,70-1,81)		0,63
porteurs/non porteurs 2	13/146	13/217	
OR 2 (95% CI)	1,49 (0,68-3,24)		0,33
OR combiné (95% CI)	1,21 (0,80-1,82)		0,42
MAF < 0,5% (n=23)			
porteurs/non porteurs 1	11/149	41/740	
OR 1 (95% CI)	1,33 (0,67-2,65)		0,41
porteurs/non porteurs 2	11/146	5/215	
OR 2 (95% CI)	3,24 (1,10-9,52)		0,04
OR combiné (95% CI)	1,77 (0,97-3,08)		0,05

Tableau 33. Odd-ratio combiné du risque de cancer-colorectal chez les porteurs d'un variant rare.

Les variants de fréquence intermédiaire étaient associés à un effet protecteur non significatif (odd-ratio combiné : 0,8, 95% CI, 0,5-1,1, p=0,18) (Tableau 34).

Variants intermédiaires	Cas	Contrôles	p
APC-15	3/151	31/750	0,25
APC-17	1/146	5/224	0,41
APC-20	7/144	6/157	0,78
PMS2-1	7/149	43/742	0,70
PMS2-2	9/147	9/227	0,46
PMS2-3	5/149	30/714	0,82
MLH3-5	4/148	15/226	0,10
porteurs/non porteurs 1	15/145	104/700	
OR 1 (95% CI)	0,70 (0,39-1,23)		0,21
porteurs/non porteurs 2	21/141	35/195	
OR 2 (95% CI)	0,83 (0,46-1,49)		0,53
OR combiné (95% CI)	0,76 (0,51-1,14)		0,18

Tableau 34. Odd-ratio combiné du risque de cancer-colorectal chez les porteurs d'un variant intermédaire.

L'OR 1 était calculé avec la population la plus importante de contrôles et l'OR 2 en utilisant le sous-groupe de 227 contrôles.

En modifiant la méthode de définition de fréquence allélique (en regroupant les cas et les témoins), APC-17 passait dans la catégorie des variants rares (<1%) et deux variants n'avaient plus une fréquence allélique < 0,5%. L'odd-ratio combiné des variants rares (<1%) était alors de 1,1 (95% CI, 0,8-1,7, p=0,54) et celui des variants rares (<0,5%) de 1,6 (95% CI, 0,8-2,9, p=0,17).

3.2. Comparaison entre polyposes et cancer colorectal précoce.

La comparaison par type d'affection des patients anglais mettait en évidence un sur-risque provoqué par les variants ayant une fréquence < 0,5% pour les patients avec une polypose (OR combiné : 1,9; 95% IC, 1,0-3,5; p=0,05; Tableaux 35 et 36).

Variant rare	Polypes multiples	CCR	Contrôles	P [a]	P [b]
EPHB2-1	1/107	0/42	1/778	0,23	1,00
EPHB2-3	1/103	0/42	2/746	0,32	1,00
EPHB2-4	0/107	1/42	1/775	1,00	0,10
EPHB2-7	1/107	0/42	0/224	0,32	1,00
EXO1-2	0/108	0/42	1/751	1,00	1,00
EXO1-4	1/91	0/34	0/226	0,29	1,00
MLH1-1	1/105	0/42	0/740	0,12	1,00
CTNNB1-1	1/106	0/68	0/702	0,13	1,00
APC-7	0/107	0/42	4/743	1,00	1,00
APC-16	0/108	0/42	3/729	1,00	1,00
BRCA2-8	0/105	0/42	4/748	1,00	1,00
BRCA2-27	0/101	1/41	6/740	1,00	0,05
BRCA2-35	1/107	0/42	0/224	0,32	1,00
BRCA2-37	4/106	0/42	5/744	0,02	1,00
MLH3-1	1/102	3/42	2/225	1,00	0,03
CDH1-1	0/108	0/42	2/748	1,00	1,00
CDH1-8	0/109	0/42	1/742	1,00	1,00
CDH1-10	0/109	0/41	2/731	1,00	1,00
BRCA1-8	0/109	0/42	1/744	1,00	1,00
BRCA1-16	0/107	0/42	1/734	1,00	1,00
BRCA1-28	1/107	0/42	2/225	1,00	1,00
SMAD4-1	1/106	0/42	7/739	1,00	1,00
CHEK2-1	2/111	1/67	1/179	0,56	0,47

Tableau 35. Comparaison de nombre de porteurs de variants rares entre les groupes polypose et CCR précoce.

[b] comparaison polyposes vs. contrôles, [c] comparaison CCR précoces vs. contrôles.

Variant rare	Polypes multiples	CCR	Contrôles	P [a]
polypes vs contrôles				
porteurs/non porteurs 1	9/106		41/740	
OR 1 (95% CI)	1,53 (0,72-3,24)			0,26
porteurs/non porteurs 2	7/103		5/215	
OR 2 (95% CI)	2,92 (0,91-9,43)			0,06
OR combiné (95% CI)	1,87 (0,98-3,48)			0,05
CCR précoces vs contrôles				
porteurs/non porteurs 1		2/43	41/740	
OR 1 (95% CI)		0,84 (0,20-3,59)		0,81
porteurs/non porteurs 2		4/42	5/215	
OR 2 (95% CI)		4,10 (1,06-5,89)		0,03
OR combiné (95% CI)		1,72 (0,73-5,27)		0,25

Tableau 36. Odd-ratio combiné du risque de cancer-colorectal chez les patients avec une polypose indéterminée ou un CCR de survenue précoce et porteur d'un variant rare (fréquence < 0,5%)

Le variant BRCA2-37 était plus fréquemment observé dans les cas de polyposes que dans les contrôles et n'était pas retrouvé chez les patients ayant un CCR de survenue précoce. En revanche, MLH3-1, était plus fréquemment observé dans le groupe de CCR précoce que dans le groupe contrôle et BRCA2-27 n'était observé que dans le groupe des CCR précoces.

Globalement, sur les 31 variants avec une fréquence < 1%, 14 étaient présents dans le groupe des polyposes indéterminées et seulement 4 dans le groupe des CCR précoces. Cette différence était significative ($p<0,05$). Le faible nombre de patients dans le groupe des CCR précoces pouvant cependant induire un biais.

La fréquence allélique était différente entre les deux groupes de patients pour deux variants rares (MLH3-1 et MLH3-5) et pour un variant commun (CDH1-5) (tableau 35). Pour ces trois polymorphismes, la fréquence était plus importante dans le groupe des CCR précoces.

3.3. Comparaison entre les groupes de patients anglais et français.

Vingt et un variants ont été séquencés dans les deux groupes (16 rares, 4 intermédiaires et 2 communs) (tableau 37).

id	variant	Allèle majeur/mineur	Cas français (N=131)	Cas UK (N=182)	Type
EPHB2-1	R80H	G/A	0,000	0,003	Variant rare
EPHB2-4	R569W	C/T	0,000	0,003	Variant rare
EPHB2-7	M883V	A/G	0,000	0,003	Variant rare
EXO1-4	L410R	T/G	0,000	0,004	Variant rare
EXO1-10	G759E	G/A	0,015	0,003	Variant rare
MSH2-8	E808X	G/T	0,000	0,000	Variant rare
CTNNB1-1	N287S	A/G	0,000	0,003	Variant rare
APC-10	I1307K	T/A	0,000	0,000	Variant rare
APC-11	E1317Q	G/C	0,004	0,011	Variant rare
APC-17	S2621C	C/G	0,008	0,003	Variant inter.
APC-20	8636 C/A	C/A	0,051	0,024	Variant inter.
PMS2-2	T597S	A/T	0,008	0,034	Variant inter.
BRCA2-7	N372H	T/G	0,273	0,320	Polymorphisme
BRCA2-35	D2665G	A/G	0,000	0,003	Variant rare
BRCA2-37	V2728I	G/A	0,000	0,014	Variant rare
BRCA2-48	P3194Q	C/A	0,000	0,000	Variant rare
MLH3-1	V741F	G/T	0,016	0,014	Variant rare
MLH3-2	M809V	A/G	0,012	0,003	Variant rare
MLH3-5	S845G	A/G	0,042	0,014	Variant inter.
CDH1-2	-284	C/A	0,315	0,246	Polymorphisme
BRCA1-28	S1512I	G/A	0,004	0,003	Variant rare
CHEK2-1	1100 delC	C/del	0,000	0,008	Variant rare

Tableau 37. Liste des variants analysés sur la population française.

Trois variants n'étaient pas retrouvés dans les deux populations (MSH2-8, APC-10 et BRCA2-48). Huit variants ne furent identifiés que dans la population anglaise (EPHB2-1, EPHB2-4, EPHB2-7, EXO1-4, CTNNB1-1, BRCA2-35, BRCA2-37, CHEK2-1) alors que tous les variants présents dans la population française étaient également observés dans la population anglaise. Il n'y avait pas de différence entre les nombres de variants rares (fréquence < 1%) entre les deux populations. En revanche les anglais présentaient plus de variants avec une fréquence <0,5% (p=0,02) (Tableau 38)

rare variant	Cas UK	Cas Fr
EPHB2-1	1/149	0/130
EPHB2-4	1/149	0/124
EPHB2-7	1/149	0/129
EXO1-4	1/125	0/130
EXO1-10	1/145	4/130
CTNNB1-1	1/174	0/130
APC-11	4/176	1/131
BRCA2-35	1/149	0/130
BRCA2-37	4/148	0/131
MLH3-2	1/148	3/129
MLH3-1	4/144	4/127
BRCA1-28	1/149	1/129
CHEK2-1	3/178	0/131
MAF < 1% (n=13)		
porteurs/non porteurs	24/149	13/128
OR 1 (95% CI)	1,59 (0,71-3,02)	0.20
MAF < 0.5% (n=10)		
porteurs/non porteurs	18/148	5/128
OR 1 (95% CI)	3,14 (1,13-8,69)	0,02

Tableau 38. Comparaison du nombre de porteurs de variants rares entres les anglais et les français.

3.4. Analyse *in silico* de l'impact de la mutation sur la fonction.

Nous avons étudié l'effet de chaque variant à l'aide des logiciels PolyPhen-2 et SNPs&GO. Selon les résultats donnés par Polyphen-2, 11 variants étaient classés comme probablement délétères, 6 comme délétères et 22 comme bénins (figures 28 et 29).

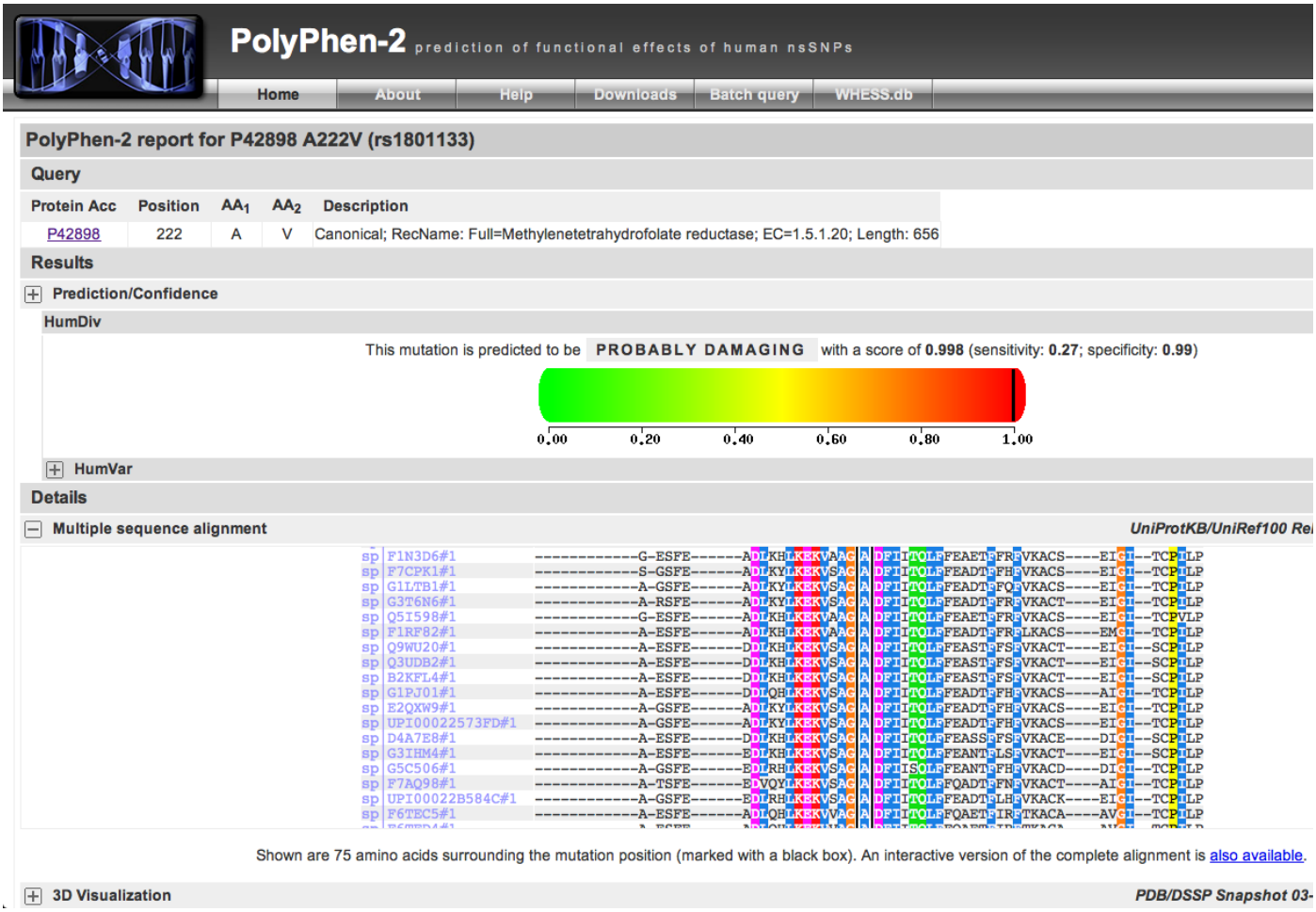


Figure 28. Utilisation du logiciel PolyPhen2 sur le variant rs1801133.

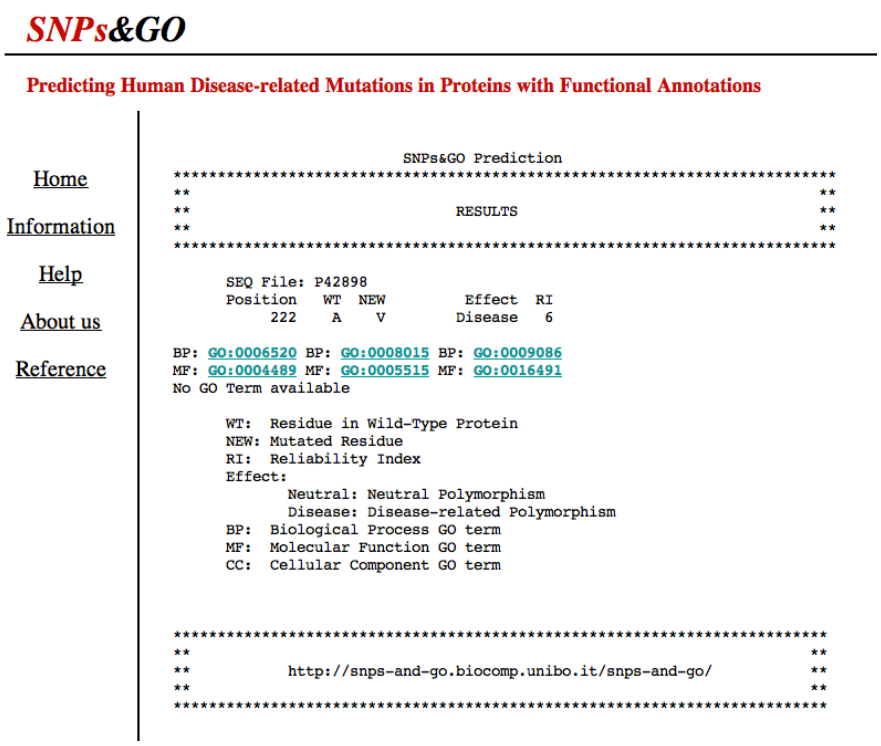


Figure 29. Utilisation du logiciel SNPs&GO sur le variant rs1801133.

Selon le logiciel SNP&GO, 18 variants étaient pathologiques et 21 neutres.

Bien que la comparaison des deux études *in silico* semble montrer des résultats similaires en valeurs absolus, il existait néanmoins de nombreuses différences (tableau 39).

id	variant	dbSNP	Allèle majeur/mineur	PolyPhen-2/FastSNP	SNPs&GO
MTHFR-1	A222V	rs1801133	C/T	Probablement délétère	Délétère
EPHB2-1	R80H	n/a	G/A	Probablement délétère	Délétère
EPHB2-3	I361V	rs56180036	A/G	Bénin	Neutre
EPHB2-4	R568W	n/a	C/T	Probablement délétère	Délétère
EPHB2-7	M883V	n/a	A/G	Probablement délétère	Neutre
EXO1-2	E109K	n/a	G/A	Probablement délétère	Neutre
EXO1-12	D249N	rs61750993	G/A	Probablement délétère	Neutre
EXO1-4	L410R	n/a	T/G	Probablement délétère	Neutre
EXO1-10	G759E	rs4150001	G/A	Bénin	Neutre
MLH1-1	G22A	rs41295280	G/C	Probablement délétère	Neutre
CTNNB1-1	N287S	rs35288908	A/G	Bénin	Délétère
APC-7	L1129S	n/a	T/C	Probablement délétère	Délétère
APC-11	E1317Q	rs1801166	G/C	Bénin	Neutre
APC-15	G2502S	rs2229995	G/A	Bénin	Neutre
APC-16	R2505Q	n/a	G/A	Probablement délétère	Neutre
APC-17	S2621C	rs72541816	C/G	Bénin	Neutre
APC-20	8636 C/A	n/a	C/A	n/a	n/a
PMS2-1	T511A	rs2228007	A/G	Bénin	Neutre
PMS2-2	T597S	rs1805318	A/T	Bénin	Neutre
PMS2-3	M622I	rs1805324	G/A	Bénin	Neutre
BRCA2-7	N372H	rs144848	T/G	Bénin	Délétère
BRCA2-8	S384F	rs41293475	C/T	Probablement délétère	Délétère
BRCA2-27	R2034C	rs1799954	C/T	Bénin	Délétère
BRCA2-35	D2665G	rs28897745	A/G	Probablement délétère	Délétère
BRCA2-37	V2728I	rs28897749	G/A	Bénin	Délétère
MLH3-1	V741F	rs28756990	G/T	Bénin	Neutre
MLH3-2	M809V	rs61752722	A/G	Bénin	Neutre
MLH3-5	S845G	rs28756992	A/G	Bénin	Neutre
AXIN1-4	D495E	n/a	C/G	Bénin	Neutre
AXIN1-6	R841Q	rs34015754	G/A	Probablement délétère	Délétère
CDH1-1	-1128	rs13335980	A/T	-S8 site fixation	n/a
CDH1-2	-284	rs16260	C/A	touche TF site fixation	n/a
CDH1-3	IVS1+6	rs3743674	T/C	Site de fixation	n/a
CDH1-5	IVS4+10	rs33963999	G/C	intronic enhancer	n/a
CDH1-7	A592T	rs35187787	G/A	Bénin	Neutre
CDH1-8	T599S	n/a	C/G	Bénin	Neutre
CDH1-10	A634V	n/a	C/T	Probablement délétère	Délétère
TP53-1	R72P	rs1042522	G/C	Bénin	Délétère
BRCA1-6	Q356R	rs1799950	A/G	Probablement délétère	Délétère

BRCA1-8	R496H	rs28897677	G/A	Bénin	Délétère
BRCA1-16	T826K	rs28897683	C/A	Probablement délétère	Délétère
BRCA1-22	E1038G	rs16941	A/G	Probablement délétère	Délétère
BRCA1-28	S1512I	rs1800744	G/T	Bénin	Délétère
AXIN2-1	N412S	rs115931022	A/G	Bénin	Neutre
SMAD4-1	A118A	n/a	G/A	Synonyme	n/a
CHEK2-1	1100 delC	n/a	C/del	p.T367fsX15	n/a

Tableau 39. Analyse *in silico* des variants testés dans la population anglaise.

Pour les variants rares (fréquence <1%), 9 étaient probablement délétères, 5 possiblement et 14 bénins selon PolyPhen2 et 13 pathologiques/15 neutres pour SNPs&GO. En se limitant aux variants avec une fréquence <0,5%, le ratio de variants pathologiques augmentait de 50% à 60%. Tous les variants intermédiaires étaient au contraire estimés comme bénins ou neutres. A coté des substitutions, un variant était silencieux (SMAD4-1), un aboutissait à un codon stop (CHEK2-1) et 5 étaient situés dans des régions non codantes.

Les variants CDH1-1 sur le promoteur de CDH1 étaient sensés provoquer la disparation d'une séquence codant pour un site de fixation d'un facteur de transcription et le variant CDH1-2 diminuait l'efficacité transcriptionelle de 68% par une altération du site de fixation.

DISCUSSION - PERSPECTIVES

Les travaux rapportés dans cette thèse de sciences portaient sur deux formes de CCR familiaux, la polypose liée à *MUTYH* et les formes non expliquées de polyposes ou de cancer colorectaux précoces. Même si ces cas restent rares au vu de la prévalence du CCR sporadique, elles constituent le sujet idéal pour explorer la cancérogenèse colorectale. C'est en effet en étudiant des familles présentant une polypose adénomateuse diffuse que le gène *APC* a été identifié. La découverte en 2002 de l'implication du gène *MUTYH* dans certains cas de polyposes adénomateuses a fourni une explication à de nombreux patients porteurs de polypes mais sans mutation sur *APC*. Jusqu'alors ils devaient avoir la même surveillance endoscopique et clinique que les patients porteurs d'une mutation germinale sur *APC*.

Notre première étude a confirmé le rôle important de ce gène suppresseur de tumeur puisqu'il expliquait un cas sur 5 de polyposes indéterminées. Ce taux montait à 27% en se limitant aux patients ayant au moins 20 polypes. Ces résultats ont depuis été confirmés par l'expertise publiée par l'INCA (Buecher, B *et al.* 2012).

Le deuxième point fondamental est le mode de transmission récessif de la MAP. Le conseil génétique est donc complètement différent avec un risque pour la descendance passant de 50% à moins de 1%. Le risque des porteurs monoallélique d'une mutation sur *MUTYH* est encore sujet à discussion pour certains mais la majorité des auteurs reconnaît un risque similaire à celui d'un patient ayant un antécédent de CCR au premier degré (Jones, N *et al.* 2009; Buecher, B *et al.* 2012) .

Enfin, la découverte de la MAP avec son phénotype proche de la PAF, permet de battre en brèche le concept même de polypose atténuée. Avant sa découverte le sous-groupe de patients avec une AFAP (moins de 100 polypes avec une survenue plus tardive) avait plus rarement une mutation identifiée sur *APC*. De surcroit, la corrélation génotype-phénotype qui poussait certains auteurs à recommander une conservation rectale avec confection d'une

anastomose iléo-rectale en cas de mutation située avant le codon 1250 n'a pas été retrouvée dans notre étude. Ces données confirment les recommandations actuelles de l'INCA sur la prise en charge des polyposes adénomateuses (Collectif. 2009) : elles doivent être basées sur les données de l'endoscopie et non pas sur la mutation. Il est donc très probable, qu'à coté des variations d'expression phénotypique de la PAF, une majorité de polyposes adénomateuses atténuées étaient en fait des MAP.

La confirmation du diagnostic de MAP repose sur l'identification des deux mutations sur *MUTYH*. Soixante treize variants ont été identifiés sur *MUTYH* dont 45% sont délétères (Buecher, B *et al.* 2012). Y165C (c.494G>A; p.Tyr165Cys) et G382D (c.1145G>A; p.Gly382Asp) sont les mutations les plus fréquemment observées dans les populations caucasiennes avec une fréquence allélique estimée dans la population générale à 0,2 et 0,6%, respectivement. Les recommandations pour la stratégie d'analyse du gène sont soit un séquençage complet des 16 exons ou une analyse séquentielle s'arrêtant dès que deux mutations ont été identifiées. Dans ce dernier cas, on débute par le séquençage des 6 exons porteurs de plus de 95% des mutations : 7, 9, 10, 12, 13 et 14. Notre étude sur le type de mutation portée sur *MUTYH* par les patients originaires d'Afrique du Nord permet de faciliter la séquence d'analyse en débutant toujours par le séquençage de l'exon 13. Nous avons observé que plus de 75% des patients provenant d'Afrique du Nord était homozygotes ou hétérozygotes pour cette mutation alors qu'elle n'était observée que dans moins de 5% des cas chez les patients caucasiens. La consanguinité fréquemment observée dans notre étude et qui a déjà fait l'objet de nombreuse publication explique le taux d'homozygotie pour cette mutation (Bittles, A 2001; Hussain, R *et al.* 2001).

L'inactivation du système BER entraine l'accumulation de transversions sur différents gènes comme *APC* ou *k-ras*. L'étude de la famille H. nous a permis de décrire la première inactivation de *MLH1* par *MUTYH*. Ces données ont été intégrées par les généticiens et il est maintenant admis que la présence d'un phénotype MSI ne doit pas éliminer le diagnostic de

MAP même si dans la majorité des cas, les patients présentent un phénotype MSS (Buecher, B *et al.* 2012). L'apparente similitude avec le syndrome de Lynch va même plus loin avec l'association des polypes avec des adénomes sébacés évoquant un syndrome de Muir-Torre (Guillen-Ponce, C *et al.* 2010). L'interaction entre *MYH* et certains gènes du système MMR dont le complexe hMUTSa (constitué de MSH2 et MSH6) est probablement une des explications.

Cependant, aucune explication n'a encore été donnée quant aux mécanismes aboutissant au cancer en cas de mutation biallélique sur *MUTYH*. En effet, contrairement au syndrome de Lynch, les deux gènes sont inactivés dès la naissance et il doit exister des mécanismes qui compensent la perte d'efficacité du système BER. De plus, une étude récente a trouvé un risque de CCR indépendant du nombre de polypes (Nieuwenhuis, MH *et al.* 2012) ce qui est une différence majeure avec la PAF. Parmi 254 patients porteurs bialléliques de mutation sur *MUTYH*, 147 (58%) ont développé un CCR. Alors qu'environ la moitié de cet effectif avait moins de 50 polypes, le risque pour les patients ayant moins de 50 polypes était 43% et celui des patients avec plus de 50 polypes de 46% (p=0,647). Il existait également 4 patients sans aucun polype au moment de la découverte de la mutation biallélique dont deux avaient déjà un CCR.

La polypose liée à MUTYH est donc une affection beaucoup plus large que ne laissait penser la première description d'Al-Tassan. En fonction des gènes, des voies de la cancérogenèse touchés par les transversions et des variabilités d'expression, un patient porteur de mutation biallélique sur MUTYH peut se présenter comme :

- une polypose adénomateuse familiale classique (plus de 100 polypes, âge de diagnostic jeune, affection extra-colique) (Sampson, JR *et al.* 2003; (de Ferro, SM *et al.* 2009) avec même un pseudo-dominance en cas de consanguinité (Etudes A et B)

- une polypose adénomateuse atténuée (cas le plus fréquent) (Jones, S *et al.* 2002) et (Etude A)
- une polypose mixte associant aux adénomes des polypes festonnés et hyperplasiques (Boparai, KS *et al.* 2008; (Castells, A 2008). Certains patients présentant tous les critères diagnostiques d'une polypose hyperplasique. On voit ici le rôle des transversions sur *k-ras* conduisant aux polypes festonnés et celles sur *APC* conduisant aux adénomes.
- Un syndrome de Lynch (Etude B)
- Une forme familiale de CCR type X (famille avec validation des critères d'Amsterdam sans mutation identifiée sur un gène du système MMR et phénotype MSS) (Jo, WS *et al.* 2005)
- Un cancer colorectal sporadique sans polypose. (Nieuwenhuis, MH *et al.* 2012)

Toutes ces observations posent plusieurs questions : faut-il renommer le syndrome MAP comme le propose James Church en *MYH-Associated Neoplasia* (Church, J *et al.* 2012) et élargir les situations cliniques devant lesquelles une recherche de mutation sur *MUTYH* devra être proposée ? Y a t-il un test de dépistage simple, à l'instar de l'IHC ou de la PCR pour le syndrome de Lynch, pour dépister les porteurs de mutation sur *MUTYH* ?

Cette dernière question fera l'objet du prochain Master M2 au sein du laboratoire d'oncogénétique de la Pitié-Salpétrière. La spécificité de la transversion sur le codon 12 de *k-ras* et sa simplicité en font un excellent test de dépistage de patients porteurs de mutation sur *MUTYH*. En étudiant la prévalence des mutations sur *k-ras* au sein d'une population de polypose *APC* négative et de CCR sporadique ainsi que le taux de patients porteurs d'une ou deux mutations sur *MUTYH* en cas de transversion trouvée sur *k-ras*, nous allons déterminer les caractéristiques de ce test. En cas de VPP>10%, il serait licite de le proposer à l'ensemble de la population opérée d'un CCR comme c'est désormais le cas à Saint-Antoine pour la

recherche de perte d'expression de protéines MMR et du phénotype MSI (Canard, G *et al.* 2011).

Malgré ces avancées dans la génétique des CCR et des polypes, il reste environ 20% de patients avec une polypose adénomateuse endoscopique sans mutation prouvée sur *APC* ou *MUTYH*. Le mosaïcisme germinal pourrait expliquer un cinquième des cas de PAF de novo (Aretz, S *et al.* 2007; Hes, FJ *et al.* 2008). Le séquençage d'autres gènes comme *SMAD4* ou *BMPR1A* pourrait permettre de reclasser certaines polyposes. Enfin l'étude d'autres gènes impliqués dans la voie Wnt comme *AXIN2, PPP2R1B, WIF1, SFRPA* qui ont été impliqués dans la cancérogenèse colorectale pourrait fournir des explications pour ces cas (Cheah, PY *et al.* 2009; Tuohy, TM *et al.* 2010). Nous avons déjà réalisé ces explorations sur 25 patients et identifiés 4 explications germinales sur *APC* : un cas de mosaïcisme germinal, une large délétion de l'exon 14, une mutation germinale intronique et une mutation non identifiées par les anciennes amorces. L'étude de 13 polyposes mixtes permis de redresser le diagnostic de polypose juvénile après identification d'une mutation germinale sur *SMAD4* et nouvel examen histologique (Mongin, C *et al.* 2011) (Annexe 4).

Pour les patients n'ayant toujours pas d'explication génétique à leur polypose, l'exploration d'autres gènes de la voie Wnt ou du TGFB est nécessaire. Par exemple *MCC* a été décrit comme muté dans certains CCR (Fukuyama, R *et al.* 2008). D'autres mécanismes pourraient également être impliqués : inactivation épigénétique germinale, micro-ARN ou les variants rares (Suter, CM *et al.* 2004; Wang, X *et al.* 2009; Wong, HL *et al.* 2010).

L'hypothèse des variants rares propose qu'une proportion importante de la susceptibilité héritée à une affection relativement fréquente soit due à l'addition d'effets provenant de variants appartenant à différent gènes, ayant une fréquence allélique faible, un effet dominant et une action indépendante. Chaque variant conférant un sur-risque modéré mais délectable de provoquer l'apparition de l'affection (Bodmer, W *et al.* 2010).

Nous avons exploré une centaine de variants sur 18 gènes différents et comparé leur fréquence entre les cas (CCR précoce ou polypose indéterminée) et un groupe 800 témoins sains (Etudes E et F). Comme déjà observé, nous avons confirmé que l'association de variants rares était associéz à un sur-risque de CR précoce ou de polyposes avec un odd-ratio aux alentours de 2. Dans l'étude spécifique sur la cycline D1, l'effet était plus important pour expliquer l'apparition de polypes multiples. Cet effet disparaissait en associant les variants intermédiaires avec une fréquence allélique <5% aux variants rares. Cette observation confirme que plus la fréquence allélique st faible plus l'OR est important et est appuyée par la théorie de l'évolution qui stipule qu'un gène conférant une affection mortelle doit être soumis à une pression de sélection négative (Bulmer, MG 1976).

Les différences de prévalence de certains variants entre les deux populations de patients (française et anglaise) confirme également que la recherche sur les variants rares doit prendre en compte un probable effet fondateur (Bodmer, W *et al.* 2008). L'implication fonctionnelle d'un variant sur la fonction est également une étape indispensable pour valider la théorie des variants rares. Nous avons observé quelques différences de l'analyse fonctionnelle in silico entre les logiciels utilisés dans l'étude F. Des analyses *in vitro* seront donc nécessaires pour valider les différents variants associés à un sur-risque d'affection colorectale (Bodmer, W *et al.* 2010).

Les études à venir vont donc devoir se concentrer sur les variants les plus significatifs notamment BRCA2-37, avec des effectifs de patients plus importants.

Un autre aspect de la théorie des variants rares est l'association synthétique entre variants commun et rares. Elle correspond à une situation où l'effet observé d'un variant commun est en fait lié à un déséquilibre de liaison entre ce variant et plusieurs variants rares qui sont agrégés en un haplotype. Par exemple, un variant commun présent dans 20% des cas et qui semble mathématiquement expliquer 1% d'une susceptibilité à une affection pourrait

simplement signaler l'effet de 2 ou 3 variants qui individuellement augmentent le risque de quelque pourcents. Cette théorie de l'association synthétique pourrait expliquer une partie des résultats de GWAS sur les variants communs identifiés (Dickson, SP *et al.* 2010). En effet, nous avons observé dans l'étude E une association significative entre le variant commun rs9344 et l'haplotype constitué des 3 variants rares CCND1-10/13/20.

L'étude de populations de malades sélectionnés (comme les polyposes indéterminées ou les CCR de survenue précoce) constitue donc la première étape dans la recherche des variants rares impliqués dans les affections néoplasiques colorectales. Les résultats apportés devront donc être confirmés au niveau fonctionnel, sur de plus grandes cohortes et au sein d'autres bassins de populations. C'est par exemple le but du consortium EPICOLON qui a déjà collecté dans les hôpitaux espagnols plus de 4000 échantillons d'ADN de malades et de contrôles appareillés sur l'âge et le sexe (Castellvi-Bel, S *et al.* 2012). Le variant rs11466445 du gène TGFBR1*6A/9A avait par exemple était associé à un sur-risque de CCR (Daley, D *et al.* 2007). Les méta-analyses publiées ultérieurement étaient contradictoires sur le véritable effet observé et une étude récente publiée par l'équipe d'Oxford a infirmé toute implication de ce variant après une analyse sur plus de 3000 cas de CCR et 3000 témoins (Carvajal-Carmona, LG *et al.* 2010) (Annexe 5).

Cette théorie est encore sujette à un certain nombre de réserves pour une fraction des généticiens (Gibson, G 2011). En effet, le risque de voir un jumeau présenter la même affection que son frère est plus important que le prédit le modèle des variants rares (Bodmer, W *et al.* 2008). L'influence d'autres facteurs génétiques est donc à prendre en compte. De plus, l'effet additif apporté par chaque variant au risque de maladie est contradictoire avec sa dominance et dans tous les cas les mécanismes sous tendus sont extrêmement difficiles à explorer. L'évolution de la prévalence de certaines affections (dont le CCR) en quelques générations va également contre la théorie des variants rares comme seule explication et confirme, entre autre, l'importance des facteurs environnementaux. Il existe enfin, des

problèmes statistiques liés à la multiplication des tests et par le calcul des fréquences alléliques au sein de la population contrôle (au lieu de la calculer en sommant la population malade et témoin) qui expose à un risque d'erreur de type I (Lemire, M 2011; Pearson, RD 2011). Dans l'étude E nous n'avons pas, par exemple, utilisé de correction de Bonferroni car nous avons estimé que l'analyse serait trop stringente alors que les variants testés sont théoriquement responsables d'une altération fonctionnelle.

Toutes ces réserves rendent indispensables la poursuite des recherches et la confirmation des résultats obtenus.

CONCLUSION

Alors qu'initialement la cancérogenèse colorectale semblait liée à une seule voie de signalisation cellulaire, codée par quelques gènes avec un schéma simple de l'adénome tubuleux à l'adénocarcinome, l'étude de formes familiales et les progrès de la biologie moléculaire ont permis d'ouvrir une boite de Pandore… Il existe maintenant plusieurs voies de la cancérogenèse avec différents types histologiques de polypes conduisant au cancer. Le nombre de gènes impliqués a augmenté exponentiellement en même temps que les modes de transmission se complexifiaient. Il existe maintenant plusieurs syndromes de transmission dominante, la polypose liée à *MUTYH* de transmission récessive et les variants rares qui confèrent un sur risque modéré mais réel au patient.

L'implication clinique de ces découvertes est évidente pour les formes syndromiques notamment pour le conseil génétique et le dépistage des apparentés ; elle reste pour le moment très limitée pour les formes familiales ou les polyposes indéterminées. Les différents travaux rapportés dans cette thèse représentent bien les différentes voies actuelles de recherche sur la cancérogenèse colorectale. Il est probable que la polypose liée à *MUTYH* va

permettre d'avancer dans la compréhension globale du CCR. Le syndrome HNPCC, par sa fréquence (environ 5% des CCR) reste le plus prometteur en terme santé publique. Une revue récente vient de schématiser l'état des connaissances sur la cancérogenèse du cancer colorectal… (Hindorff, LA *et al.* 2011))

La route est encore très longue (figure 30)….

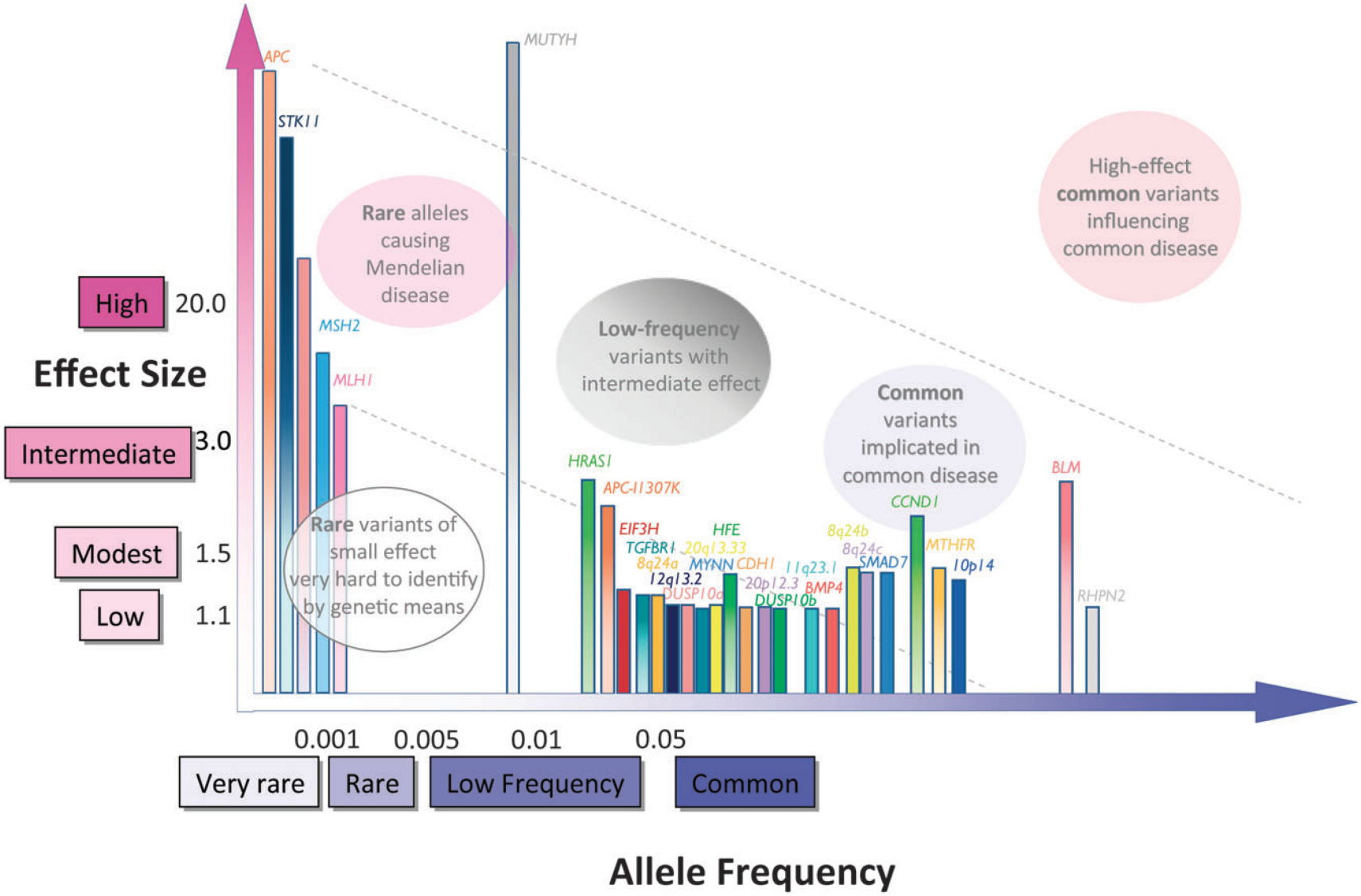


Figure 30. Représentation schématique des gènes impliqués dans la cancérogenèse colorectale.

REFERENCES

Ahn, SJ, Costa, J and Emanuel, JR (1996). "PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR." *Nucleic Acids Res* **24**(13): 2623-5.

Al-Tassan, N, Chmiel, NH, Maynard, J, Fleming, N, Livingston, AL, Williams, GT*, et al.* (2002). "Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors." *Nat Genet* **30**(2): 227-32.

Al-Tassan, N, Eisen, T, Maynard, J, Bridle, H, Shah, B, Fleischmann, C*, et al.* (2004). "Inherited variants in MYH are unlikely to contribute to the risk of lung carcinoma." *Hum Genet* **114**(2): 207-10.

Allan, JM, Shorto, J, Adlard, J, Bury, J, Coggins, R, George, R*, et al.* (2008). "MLH1 -93G>A promoter polymorphism and risk of mismatch repair deficient colorectal cancer." *Int J Cancer* **123**(10): 2456-9.

Allegra, CJ, Jessup, JM, Somerfield, MR, Hamilton, SR, Hammond, EH, Hayes, DF*, et al.* (2009). "American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy." *J Clin Oncol* **27**(12): 2091-6.

Aretz, S, Stienen, D, Friedrichs, N, Stemmler, S, Uhlhaas, S, Rahner, N*, et al.* (2007). "Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP)." *Hum Mutat* **28**(10): 985-92.

Aretz, S, Stienen, D, Uhlhaas, S, Stolte, M, Entius, MM, Loff, S*, et al.* (2007). "High proportion of large genomic deletions and a genotype phenotype update in 80 unrelated families with juvenile polyposis syndrome." *J Med Genet* **44**(11): 702-9.

Balmana, J, Balaguer, F, Castellvi-Bel, S, Steyerberg, EW, Andreu, M, Llor, X*, et al.* (2008). "Comparison of predictive models, clinical criteria and molecular tumour screening for the identification of patients with Lynch syndrome in a population-based cohort of colorectal cancer patients." *J Med Genet* **45**(9): 557-63.

Barrett, JC, Fry, B, Maller, J and Daly, MJ (2005). "Haploview: analysis and visualization of LD and haplotype maps." *Bioinformatics* **21**(2): 263-5.

Beggs, AD, Latchford, AR, Vasen, HF, Moslein, G, Alonso, A, Aretz, S*, et al.* (2010). "Peutz-Jeghers syndrome: a systematic review and recommendations for management." *Gut* **59**(7): 975-86.

Bittles, A (2001). "Consanguinity and its relevance to clinical genetics." *Clin Genet* **60**(2): 89-98.

Bodmer, W and Bonilla, C (2008). "Common and rare variants in multifactorial susceptibility to common diseases." *Nat Genet* **40**(6): 695-701.

Bodmer, W and Tomlinson, I (2010). "Rare genetic variants and the risk of cancer." *Curr Opin Genet Dev* **20**(3): 262-7.

Boland, CR and Goel, A (2010). "Microsatellite instability in colorectal cancer." *Gastroenterology* **138**(6): 2073-2087 e3.

Boparai, KS, Dekker, E, Van Eeden, S, Polak, MM, Bartelsman, JF, Mathus-Vliegen, EM*, et al.* (2008). "Hyperplastic polyps and sessile serrated adenomas as a phenotypic expression of MYH-associated polyposis." *Gastroenterology* **135**(6): 2014-8.

Brosens, LA, Langeveld, D, van Hattem, WA, Giardiello, FM and Offerhaus, GJ (2011). "Juvenile polyposis syndrome." *World J Gastroenterol* **17**(44): 4839-44.

Brosens, LA, van Hattem, A, Hylind, LM, Iacobuzio-Donahue, C, Romans, KE, Axilbund, J*, et al.* (2007). "Risk of colorectal cancer in juvenile polyposis." *Gut* **56**(7): 965-7.

Buecher, B, Bonaiti, C, Buisine, MP, Colas, C and Saurin, JC (2012). "French experts report on MUTYH-associated polyposis (MAP)." *Fam Cancer*.

Bulmer, MG (1976). "The effect of selection on genetic variability: a simulation study." *Genet Res* **28**(2): 101-17.

Bulow, C, Vasen, H, Jarvinen, H, Bjork, J, Bisgaard, ML and Bulow, S (2000). "Ileorectal anastomosis is appropriate for a subset of patients with familial adenomatous polyposis." *Gastroenterology* **119**(6): 1454-60.

Butterworth, AS, Higgins, JP and Pharoah, P (2006). "Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis." *Eur J Cancer* **42**(2): 216-27.

Caillie, F, Paye, F, Desaint, B, Bennis, M, Lefevre, JH, Parc, Y*, et al.* (2012). "Severe Duodenal Involvement in Familial Adenomatous Polyposis Treated by Pylorus-Preserving Pancreaticoduodenectomy." *Ann Surg Oncol*.

Canard, G, Lefevre, JH, Colas, C, Coulet, F, Svrcek, M, Lascols, O*, et al.* (2011). "Screening for Lynch Syndrome in Colorectal Cancer: Are We Doing Enough?" *Ann Surg Oncol*.

Carvajal-Carmona, LG, Churchman, M, Bonilla, C, Walther, A, Lefevre, JH, Kerr, D*, et al.* (2010). "Comprehensive assessment of variation at the transforming growth factor beta type 1 receptor locus and colorectal cancer predisposition." *Proc Natl Acad Sci U S A* **107**(17): 7858-62.

Caspari, R, Olschwang, S, Friedl, W, Mandl, M, Boisson, C, Boker, T*, et al.* (1995). "Familial adenomatous polyposis: desmoid tumours and lack of ophthalmic lesions

(CHRPE) associated with APC mutations beyond codon 1444." *Hum Mol Genet* **4**(3): 337-40.

Castells, A (2008). "MYH-Associated Polyposis: Adenomas and Hyperplastic Polyps, Partners in Crime?" *Gastroenterology*.

Castellvi-Bel, S, Ruiz-Ponte, C, Fernandez-Rozadilla, C, Abuli, A, Munoz, J, Bessa, X*, et al.* (2012). "Seeking genetic susceptibility variants for colorectal cancer: the EPICOLON consortium experience." *Mutagenesis* **27**(2): 153-9.

Cheadle, JP, Dolwani, S and Sampson, JR (2003). "Inherited defects in the DNA glycosylase MYH cause multiple colorectal adenoma and carcinoma." *Carcinogenesis* **24**(7): 1281-2; author reply 1283.

Cheah, PY, Wong, YH, Chau, YP, Loi, C, Lim, KH, Lim, JF*, et al.* (2009). "Germline bone morphogenesis protein receptor 1A mutation causes colorectal tumorigenesis in hereditary mixed polyposis syndrome." *Am J Gastroenterol* **104**(12): 3027-33.

Church, J, Heald, B, Burke, C and Kalady, M (2012). "Understanding MYH-associated neoplasia." *Dis Colon Rectum* **55**(3): 359-62.

Cleary, SP, Cotterchio, M, Jenkins, MA, Kim, H, Bristow, R, Green, R*, et al.* (2009). "Germline MutY human homologue mutations and colorectal cancer: a multisite case-control study." *Gastroenterology* **136**(4): 1251-60.

Coassin, S, Brandstatter, A and Kronenberg, F (2010). "Lost in the space of bioinformatic tools: a constantly updated survival guide for genetic epidemiology. The GenEpi Toolbox." *Atherosclerosis* **209**(2): 321-35.

Colebatch, A, Hitchins, M, Williams, R, Meagher, A, Hawkins, NJ and Ward, RL (2006). "The role of MYH and microsatellite instability in the development of sporadic colorectal cancer." *Br J Cancer* **95**(9): 1239-43.

Collectif. (2009). *Chirurgie prohylactique: syndrome HNPCC*.

Collectif. (2009). *Chirurgie prophylactique: Polypose adénomateurs familiale et polypose liée à MYH.*

Collectif. (2011). *La polypose assoicée aux mutations bi-alléliques du gène MUTYH.*

Daley, D, Morgan, W, Lewis, S, Willis, J, Elston, RC, Markowitz, SD*, et al.* (2007). "Is TGFBR1*6A a susceptibility allele for nonsyndromic familial colorectal neoplasia?" *Cancer Epidemiol Biomarkers Prev* **16**(5): 892-4.

de Ferro, SM, Suspiro, A, Fidalgo, P, Lage, P, Rodrigues, P, Fragoso, S*, et al.* (2009). "Aggressive phenotype of MYH-associated polyposis with jejunal cancer and intra-abdominal desmoid tumor: report of a case." *Dis Colon Rectum* **52**(4): 742-5.

de Leng, WW, Jansen, M, Carvalho, R, Polak, M, Musler, AR, Milne, AN*, et al.* (2007). "Genetic defects underlying Peutz-Jeghers syndrome (PJS) and exclusion of the polarity-associated MARK/Par1 gene family as potential PJS candidates." *Clin Genet* **72**(6): 568-73.

De Roock, W, Claes, B, Bernasconi, D, De Schutter, J, Biesmans, B, Fountzilas, G*, et al.* (2010). "Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis." *Lancet Oncol* **11**(8): 753-62.

Deng, G, Bell, I, Crawley, S, Gum, J, Terdiman, JP, Allen, BA*, et al.* (2004). "BRAF mutation is frequently present in sporadic colorectal cancer with methylated hMLH1, but not in hereditary nonpolyposis colorectal cancer." *Clin Cancer Res* **10**(1 Pt 1): 191-5.

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

Dickson, SP, Wang, K, Krantz, I, Hakonarson, H and Goldstein, DB (2010). "Rare variants create synthetic genome-wide associations." *PLoS Biol* **8**(1): e1000294.

Djureinovic, T, Skoglund, J, Vandrovcova, J, Zhou, XL, Kalushkova, A, Iselius, L*, et al.* (2006). "A genome wide linkage analysis in Swedish families with hereditary non-familial adenomatous polyposis/non-hereditary non-polyposis colorectal cancer." *Gut* **55**(3): 362-6.

Duval, A and Hamelin, R (2002). "Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability." *Cancer Res* **62**(9): 2447-54.

Fearnhead, NS, Wilding, JL, Winney, B, Tonks, S, Bartlett, S, Bicknell, DC*, et al.* (2004). "Multiple rare variants in different genes account for multifactorial inherited susceptibility to colorectal adenomas." *Proc Natl Acad Sci U S A* **101**(45): 15992-7.

Flossmann, E and Rothwell, PM (2007). "Effect of aspirin on long-term risk of colorectal cancer: consistent evidence from randomised and observational studies." *Lancet* **369**(9573): 1603-13.

Fodde, R, Smits, R and Clevers, H (2001). "APC, signal transduction and genetic instability in colorectal cancer." *Nat Rev Cancer* **1**(1): 55-67.

Frayling, IM, Beck, NE, Ilyas, M, Dove-Edwin, I, Goodman, P, Pack, K*, et al.* (1998). "The APC variants I1307K and E1317Q are associated with colorectal tumors, but not always with a family history." *Proc Natl Acad Sci U S A* **95**(18): 10722-7.

Fukuyama, R, Niculaita, R, Ng, KP, Obusez, E, Sanchez, J, Kalady, M*, et al.* (2008). "Mutated in colorectal cancer, a putative tumor suppressor for serrated colorectal cancer, selectively represses beta-catenin-dependent transcription." *Oncogene* **27**(46): 6044-55.

Gallione, CJ, Repetto, GM, Legius, E, Rustgi, AK, Schelley, SL, Tejpar, S*, et al.* (2004). "A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4)." *Lancet* **363**(9412): 852-9.

Gebert, JF, Dupon, C, Kadmon, M, Hahn, M, Herfarth, C, von Knebel Doeberitz, M*, et al.* (1999). "Combined molecular and clinical approaches for the identification of families with familial adenomatous polyposis coli." *Ann Surg* **229**(3): 350-61.

Gellad, ZF and Provenzale, D (2010). "Colorectal cancer: national and international perspective on the burden of disease and public health impact." *Gastroenterology* **138**(6): 2177-90.

Giardiello, FM, Krush, AJ, Petersen, GM, Booker, SV, Kerr, M, Tong, LL*, et al.* (1994). "Phenotypic variability of familial adenomatous polyposis in 11 unrelated families with identical APC gene mutation." *Gastroenterology* **106**(6): 1542-7.

Gibson, G (2011). "Rare and common variants: twenty arguments." *Nat Rev Genet* **13**(2): 135-45.

Grady, WM (2003). "Genetic testing for high-risk colon cancer patients." *Gastroenterology* **124**(6): 1574-94.

Guillen-Ponce, C, Castillejo, A, Barbera, VM, Pascual-Ramirez, JC, Andrada, E, Castillejo, MI*, et al.* (2010). "Biallelic MYH germline mutations as cause of Muir-Torre syndrome." *Fam Cancer* **9**(2): 151-4.

Half, E, Bercovich, D and Rozen, P (2009). "Familial adenomatous polyposis." *Orphanet J Rare Dis* **4**: 22.

Hampel, H, Frankel, WL, Martin, E, Arnold, M, Khanduja, K, Kuebler, P*, et al.* (2008). "Feasibility of screening for Lynch syndrome among patients with colorectal cancer." *J Clin Oncol* **26**(35): 5783-8.

Hes, FJ, Nielsen, M, Bik, EC, Konvalinka, D, Wijnen, JT, Bakker, E*, et al.* (2008). "Somatic APC mosaicism: an underestimated cause of polyposis coli." *Gut* **57**(1): 71-6.

Hindorff, LA, Gillanders, EM and Manolio, TA (2011). "Genetic architecture of cancer and other complex diseases: lessons learned and future directions." *Carcinogenesis* **32**(7): 945-54.

Hor, T, Zalinski, S, Lefevre, JH, Shields, C, Attal, E, Tiret, E*, et al.* (2012). "Feasibility of laparoscopic restorative proctocolectomy without diverting stoma." *Dig Liver Dis* **44**(2): 118-22.

Houlston, RS, Webb, E, Broderick, P, Pittman, AM, Di Bernardo, MC, Lubbe, S*, et al.* (2008). "Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer." *Nat Genet* **40**(12): 1426-35.

Howe, D and Lynas, C (2001). "The cyclin D1 alternative transcripts [a] and [b] are expressed in normal and malignant lymphocytes and their relative levels are influenced by the polymorphism at codon 241." *Haematologica* **86**(6): 563-9.

Hussain, R, Bittles, AH and Sullivan, S (2001). "Consanguinity and early mortality in the Muslim populations of India and Pakistan." *Am J Hum Biol* **13**(6): 777-87.

Ikushima, H and Miyazono, K (2010). "TGFbeta signalling: a complex web in cancer progression." *Nat Rev Cancer* **10**(6): 415-24.

Isidro, G, Laranjeira, F, Pires, A, Leite, J, Regateiro, F, Castro e Sousa, F*, et al.* (2004). "Germline MUTYH (MYH) mutations in Portuguese individuals with multiple colorectal adenomas." *Hum Mutat* **24**(4): 353-4.

Issa, JP (2004). "CpG island methylator phenotype in cancer." *Nat Rev Cancer* **4**(12): 988-93.

Issa, JP (2008). "Colon cancer: it's CIN or CIMP." *Clin Cancer Res* **14**(19): 5939-40.

Jasperson, KW, Tuohy, TM, Neklason, DW and Burt, RW (2010). "Hereditary and familial colon cancer." *Gastroenterology* **138**(6): 2044-58.

Jasperson, KW, Vu, TM, Schwab, AL, Neklason, DW, Rodriguez-Bigas, MA, Burt, RW*, et al.* (2010). "Evaluating Lynch syndrome in very early onset colorectal cancer probands without apparent polyposis." *Fam Cancer* **9**(2): 99-107.

Jass, JR, Williams, CB, Bussey, HJ and Morson, BC (1988). "Juvenile polyposis--a precancerous condition." *Histopathology* **13**(6): 619-30.

Jo, WS, Bandipalliam, P, Shannon, KM, Niendorf, KB, Chan-Smutko, G, Hur, C*, et al.* (2005). "Correlation of polyp number and family history of colon cancer with germline MYH mutations." *Clin Gastroenterol Hepatol* **3**(10): 1022-8.

Johns, LE and Houlston, RS (2001). "A systematic review and meta-analysis of familial colorectal cancer risk." *Am J Gastroenterol* **96**(10): 2992-3003.

Jones, N, Vogt, S, Nielsen, M, Christian, D, Wark, PA, Eccles, D*, et al.* (2009). "Increased colorectal cancer incidence in obligate carriers of heterozygous mutations in MUTYH." *Gastroenterology* **137**(2): 489-94, 494 e1; quiz 725-6.

Jones, S, Emmerson, P, Maynard, J, Best, JM, Jordan, S, Williams, GT*, et al.* (2002). "Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C-->T:A mutations." *Hum Mol Genet* **11**(23): 2961-7.

Kambara, T, Simms, LA, Whitehall, VL, Spring, KJ, Wynter, CV, Walsh, MD*, et al.* (2004). "BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum." *Gut* **53**(8): 1137-44.

Kang, D, Nishida, J, Iyama, A, Nakabeppu, Y, Furuichi, M, Fujiwara, T*, et al.* (1995). "Intracellular localization of 8-oxo-dGTPase in human cells, with special reference to the role of the enzyme in mitochondria." *J Biol Chem* **270**(24): 14659-65.

Kemp, Z, Carvajal-Carmona, L, Spain, S, Barclay, E, Gorman, M, Martin, L*, et al.* (2006). "Evidence for a colorectal cancer susceptibility locus on chromosome 3q21-q24 from a high-density SNP genome-wide linkage scan." *Hum Mol Genet* **15**(19): 2903-10.

Kerber, RA, Neklason, DW, Samowitz, WS and Burt, RW (2005). "Frequency of familial colon cancer and hereditary nonpolyposis colorectal cancer (Lynch syndrome) in a large population database." *Fam Cancer* **4**(3): 239-44.

Kinzler, KW, Nilbert, MC, Su, LK, Vogelstein, B, Bryan, TM, Levy, DB*, et al.* (1991). "Identification of FAP locus genes from chromosome 5q21." *Science* **253**(5020): 661-5.

Klaus, A and Birchmeier, W (2008). "Wnt signalling and its impact on development and cancer." *Nat Rev Cancer* **8**(5): 387-98.

Knudsen, AL, Bisgaard, ML and Bulow, S (2003). "Attenuated familial adenomatous polyposis (AFAP). A review of the literature." *Fam Cancer* **2**(1): 43-55.

Knudson, AG, Jr. (1971). "Mutation and cancer: statistical study of retinoblastoma." *Proc Natl Acad Sci U S A* **68**(4): 820-3.

Kong, S, Amos, CI, Luthra, R, Lynch, PM, Levin, B and Frazier, ML (2000). "Effects of cyclin D1 polymorphism on age of onset of hereditary nonpolyposis colorectal cancer." *Cancer Res* **60**(2): 249-52.

Kong, S, Wei, Q, Amos, CI, Lynch, PM, Levin, B, Zong, J*, et al.* (2001). "Cyclin D1 polymorphism and increased risk of colorectal cancer at young age." *J Natl Cancer Inst* **93**(14): 1106-8.

Korinek, V, Barker, N, Morin, PJ, van Wichen, D, de Weger, R, Kinzler, KW*, et al.* (1997). "Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma." *Science* **275**(5307): 1784-7.

Laken, SJ, Petersen, GM, Gruber, SB, Oddoux, C, Ostrer, H, Giardiello, FM*, et al.* (1997). "Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC." *Nat Genet* **17**(1): 79-83.

Lamlum, H, Al Tassan, N, Jaeger, E, Frayling, I, Sieber, O, Reza, FB*, et al.* (2000). "Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q." *Hum Mol Genet* **9**(15): 2215-21.

Laurent-Puig, P, Agostini, J and Maley, K (2010). "[Colorectal oncogenesis]." *Bull Cancer* **97**(11): 1311-21.

Lee, PH and Shatkay, H (2008). "F-SNP: computationally predicted functional SNPs for disease association studies." *Nucleic Acids Res* **36**(Database issue): D820-4.

Lefevre, JH, Parc, Y, Kerneis, S, Goasguen, N, Benis, M, Parc, R*, et al.* (2008). "Risk factors for development of desmoid tumours in familial adenomatous polyposis." *Br J Surg* **95**(9): 1136-9.

Lefevre, JH, Rodrigue, CM, Mourra, N, Bennis, M, Flejou, JF, Parc, R*, et al.* (2006). "Implication of MYH in colorectal polyposis." *Ann Surg* **244**(6): 874-9; discussion 879-80.

Leggett, B and Whitehall, VL (2010). "Role of the serrated pathway in colorectal cancer pathogenesis." *Gastroenterol* **138**(6): 2088-2100.

Lemire, M (2011). "Defining rare variants by their frequencies in controls may increase type I error." *Nat Genet* **43**(5): 391-2; author reply 394-5.

Lichtenstein, P, Holm, NV, Verkasalo, PK, Iliadou, A, Kaprio, J, Koskenvuo, M*, et al.* (2000). "Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland." *N Engl J Med* **343**(2): 78-85.

Ligtenberg, MJ, Kuiper, RP, Chan, TL, Goossens, M, Hebeda, KM, Voorendt, M*, et al.* (2009). "Heritable somatic methylation and inactivation of MSH2 in families with

Lynch syndrome due to deletion of the 3' exons of TACSTD1." *Nat Genet* **41**(1): 112-7.

Lindor, NM (2009). "Familial colorectal cancer type X: the other half of hereditary nonpolyposis colon cancer syndrome." *Surg Oncol Clin N Am* **18**(4): 637-45.

Lindor, NM, Rabe, K, Petersen, GM, Haile, R, Casey, G, Baron, J*, et al.* (2005). "Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X." *JAMA* **293**(16): 1979-85.

Lipton, L, Halford, SE, Johnson, V, Novelli, MR, Jones, A, Cummings, C*, et al.* (2003). "Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway." *Cancer Res* **63**(22): 7595-9.

Lipton, L, Sieber, OM, Thomas, HJW, Hodgson, SV, Tomlinson, IPM and Woodford-Richens, K (2003). "Germline mutations in the TGF-beta and Wnt signalling pathways are a rare cause of the "multiple" adenoma phenotype." *Journal of Medical Genetics* **40**(4): e35.

Liu, K and Muse, SV (2005). "PowerMarker: an integrated analysis environment for genetic marker analysis." *Bioinformatics* **21**(9): 2128-9.

Lynch, HT, Smyrk, T, McGinn, T, Lanspa, S, Cavalieri, J, Lynch, J*, et al.* (1995). "Attenuated familial adenomatous polyposis (AFAP). A phenotypically and genotypically distinctive variant of FAP." *Cancer* **76**(12): 2427-33.

Macrae, F, du Sart, D and Nasioulas, S (2009). "Familial adenomatous polyposis." *Best Pract Res Clin Gastroenterol* **23**(2): 197-207.

Manolio, TA (2010). "Genomewide association studies and assessment of the risk of disease." *N Engl J Med* **363**(2): 166-76.

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

Matsuo, S, Eguchi, S, Azuma, T, Hidaka, M, Yamaguchi, S, Hayashi, T*, et al.* (2001). "Attenuated familial adenomatous polyposis associated with advanced rectal cancer in a 16-year-old boy: report of a case." *Surg Today* **31**(11): 1020-3.

Mitry, E, Rachet, B, Quinn, MJ, Cooper, N and Coleman, MP (2008). "Survival from cancer of the colon in England and Wales up to 2001." *Br J Cancer* **99 Suppl 1**: S26-9.

Miyaki, M, Iijima, T, Yamaguchi, T, Hishima, T, Tamura, K, Utsunomiya, J*, et al.* (2005). "Germline mutations of the MYH gene in Japanese patients with multiple colorectal adenomas." *Mutat Res* **578**(1-2): 430-3.

Miyoshi, Y, Ando, H, Nagase, H, Nishisho, I, Horii, A, Miki, Y*, et al.* (1992). "Germ-line mutations of the APC gene in 53 familial adenomatous polyposis patients." *Proc Natl Acad Sci U S A* **89**(10): 4452-6.

Mongin, C, Coulet, F, Lefevre, J, Colas, C, Svrcek, M, Eyries, M*, et al.* (2011). "Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists." *Clin Genet*.

Nagase, H, Miyoshi, Y, Horii, A, Aoki, T, Ogawa, M, Utsunomiya, J*, et al.* (1992). "Correlation between the location of germ-line mutations in the APC gene and the number of colorectal polyps in familial adenomatous polyposis patients." *Cancer Res* **52**(14): 4055-7.

Narayan, S and Roy, D (2003). "Role of APC and DNA mismatch repair genes in the development of colorectal cancers." *Mol Cancer* **2**: 41.

Neklason, DW, Kerber, RA, Nilson, DB, Anton-Culver, H, Schwartz, AG, Griffin, CA*, et al.* (2008). "Common familial colorectal cancer linked to chromosome 7q31: a genome-wide analysis." *Cancer Research* **68**(21): 8993-7.

Nielsen, M, Morreau, H, Vasen, HF and Hes, FJ (2011). "MUTYH-associated polyposis (MAP)." *Crit Rev Oncol Hematol* **79**(1): 1-16.

Niessen, RC, Hofstra, RM, Westers, H, Ligtenberg, MJ, Kooi, K, Jager, PO*, et al.* (2009). "Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome." *Genes Chromosomes Cancer* **48**(8): 737-44.

Nieuwenhuis, MH, Lefevre, JH, Bulow, S, Jarvinen, H, Bertario, L, Kerneis, S*, et al.* (2011). "Family history, surgery, and APC mutation are risk factors for desmoid tumors in familial adenomatous polyposis: an international cohort study." *Dis Colon Rectum* **54**(10): 1229-34.

Nieuwenhuis, MH and Vasen, HF (2007). "Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): a review of the literature." *Crit Rev Oncol Hematol* **61**(2): 153-61.

Nieuwenhuis, MH, Vogt, S, Jones, N, Nielsen, M, Hes, FJ, Sampson, JR*, et al.* (2012). "Evidence for accelerated colorectal adenoma--carcinoma progression in MUTYH-associated polyposis?" *Gut* **61**(5): 734-8.

Olschwang, S, Hamelin, R, Laurent-Puig, P, Thuille, B, De Rycke, Y, Li, YJ*, et al.* (1997). "Alternative genetic pathways in colorectal carcinogenesis." *Proc Natl Acad Sci U S A* **94**(22): 12122-7.

Pabalan, N, Bapat, B, Sung, L, Jarjanazi, H, Francisco-Pabalan, O and Ozcelik, H (2008). "Cyclin D1 Pro241Pro (CCND1-G870A) polymorphism is associated with increased cancer risk in human populations: a meta-analysis." *Cancer Epidemiol Biomarkers Prev* **17**(10): 2773-81.

Parc, Y, Mabrut, JY and Shields, C (2011). "Surgical management of the duodenal manifestations of familial adenomatous polyposis." *Br J Surg* **98**(4): 480-4.

Parry, S, Win, AK, Parry, B, Macrae, FA, Gurrin, LC, Church, JM*, et al.* (2011). "Metachronous colorectal cancer risk for mismatch repair gene mutation carriers: the advantage of more extensive colon surgery." *Gut* **60**(7): 950-7.

Pearson, RD (2011). "Bias due to selection of rare variants using frequency in controls." *Nat Genet* **43**(5): 392-3; author reply 394-5.

Peltomaki, P and Vasen, H (2004). "Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database." *Dis Markers* **20**(4-5): 269-76.

Pendlebury, S, Duchesne, F, Reed, KA, Smith, JL and Kerr, DJ (2003). "A trial of adjuvant therapy in colorectal cancer: the VICTOR trial." *Clin Colorectal Cancer* **3**(1): 58-60.

Penna, CP, Tiret, E, Kartheuser, AH, Olschwang, S and Parc, R (1992). "[Value of screening of familial adenomatous polyposis for the prevention of colorectal cancer]." *Gastroenterol Clin Biol* **16**(3): 210-4.

Peters, U, Hutter, CM, Hsu, L, Schumacher, FR, Conti, DV, Carlson, CS*, et al.* (2011). "Meta-analysis of new genome-wide association studies of colorectal cancer risk." *Hum Genet*.

Pino, MS and Chung, DC (2010). "The chromosomal instability pathway in colon cancer." *Gastroenterology* **138**(6): 2059-72.

Purcell, S, Neale, B, Todd-Brown, K, Thomas, L, Ferreira, MA, Bender, D*, et al.* (2007). "PLINK: a tool set for whole-genome association and population-based linkage analyses." *Am J Hum Genet* **81**(3): 559-75.

Rampino, N, Yamamoto, H, Ionov, Y, Li, Y, Sawai, H, Reed, JC*, et al.* (1997). "Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype." *Science* **275**(5302): 967-9.

Robbins, DH and Itzkowitz, SH (2002). "The molecular and genetic basis of colon cancer." *Med Clin North Am* **86**(6): 1467-95.

Sampson, JR, Dolwani, S, Jones, S, Eccles, D, Ellis, A, Evans, DG*, et al.* (2003). "Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH." *Lancet* **362**(9377): 39-41.

Sampson, JR and Jones, N (2009). "MUTYH-associated polyposis." *Best Pract Res Clin Gastroenterol* **23**(2): 209-18.

Sauer, S and Gut, IG (2002). "Genotyping single-nucleotide polymorphisms by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry." *J Chromatogr B Analyt Technol Biomed Life Sci* **782**(1-2): 73-87.

Schmeler, KM, Lynch, HT, Chen, LM, Munsell, MF, Soliman, PT, Clark, MB*, et al.* (2006). "Prophylactic surgery to reduce the risk of gynecologic cancers in the Lynch syndrome." *N Engl J Med* **354**(3): 261-9.

Sieber, OM, Lipton, L, Crabtree, M, Heinimann, K, Fidalgo, P, Phillips, RK*, et al.* (2003). "Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH." *N Engl J Med* **348**(9): 791-9.

Slupska, MM, Baikalov, C, Luther, WM, Chiang, JH, Wei, YF and Miller, JH (1996). "Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutY gene whose function is required for the repair of oxidative DNA damage." *J Bacteriol* **178**(13): 3885-92.

Solomon, DA, Wang, Y, Fox, SR, Lambeck, TC, Giesting, S, Lan, Z*, et al.* (2003). "Cyclin D1 splice variants. Differential effects on localization, RB phosphorylation, and cellular transformation." *J Biol Chem* **278**(32): 30339-47.

Souza, RF, Appel, R, Yin, J, Wang, S, Smolinski, KN, Abraham, JM*, et al.* (1996). "Microsatellite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumours." *Nat Genet* **14**(3): 255-7.

Spirio, L, Olschwang, S, Groden, J, Robertson, M, Samowitz, W, Joslyn, G*, et al.* (1993). "Alleles of the APC gene: an attenuated form of familial polyposis." *Cell* **75**(5): 951-7.

Stoffel, E, Mukherjee, B, Raymond, VM, Tayob, N, Kastrinos, F, Sparr, J*, et al.* (2009). "Calculation of risk of colorectal and endometrial cancer among patients with Lynch syndrome." *Gastroenterology* **137**(5): 1621-7.

Stupart, DA, Goldberg, PA, Algar, U and Ramesar, R (2009). "Surveillance colonoscopy improves survival in a cohort of subjects with a single mismatch repair gene mutation." *Colorectal Dis* **11**(2): 126-30.

Suter, CM, Martin, DI and Ward, RL (2004). "Germline epimutation of MLH1 in individuals with multiple cancers." *Nat Genet* **36**(5): 497-501.

Sweet, K, Willis, J, Zhou, XP, Gallione, C, Sawada, T, Alhopuro, P*, et al.* (2005). "Molecular classification of patients with unexplained hamartomatous and hyperplastic polyposis." *JAMA* **294**(19): 2465-73.

Takao, M, Zhang, QM, Yonei, S and Yasui, A (1999). "Differential subcellular localization of human MutY homolog (hMYH) and the functional activity of adenine:8-oxoguanine DNA glycosylase." *Nucleic Acids Res* **27**(18): 3638-44.

Trano, G, Sjursen, W, Wasmuth, HH, Hofsli, E and Vatten, LJ (2010). "Performance of clinical guidelines compared with molecular tumour screening methods in

identifying possible Lynch syndrome among colorectal cancer patients: a Norwegian population-based study." *Br J Cancer* **102**(3): 482-8.

Tulupova, E, Kumar, R, Hanova, M, Slyskova, J, Pardini, B, Polakova, V*, et al.* (2008). "Do polymorphisms and haplotypes of mismatch repair genes modulate risk of sporadic colorectal cancer?" *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* **648**(1-2): 40-5.

Tuohy, TM, Done, MW, Lewandowski, MS, Shires, PM, Saraiya, DS, Huang, SC*, et al.* (2010). "Large intron 14 rearrangement in APC results in splice defect and attenuated FAP." *Hum Genet* **127**(3): 359-69.

van der Luijt, RB, Vasen, HF, Tops, CM, Breukel, C, Fodde, R and Meera Khan, P (1995). "APC mutation in the alternatively spliced region of exon 9 associated with late onset familial adenomatous polyposis." *Hum Genet* **96**(6): 705-10.

van Roon, EH, van Puijenbroek, M, Middeldorp, A, van Eijk, R, de Meijer, EJ, Erasmus, D*, et al.* (2010). "Early onset MSI-H colon cancer with MLH1 promoter methylation, is there a genetic predisposition?" *BMC Cancer* **10**: 180.

Vasen, HF, Moslein, G, Alonso, A, Aretz, S, Bernstein, I, Bertario, L*, et al.* (2008). "Guidelines for the clinical management of familial adenomatous polyposis (FAP)." *Gut* **57**(5): 704-13.

Vasen, HF, van der Luijt, RB, Slors, JF, Buskens, E, de Ruiter, P, Baeten, CG*, et al.* (1996). "Molecular genetic tests as a guide to surgical management of familial adenomatous polyposis." *Lancet* **348**(9025): 433-5.

Volikos, E, Robinson, J, Aittomaki, K, Mecklin, JP, Jarvinen, H, Westerman, AM*, et al.* (2006). "LKB1 exonic and whole gene deletions are a common cause of Peutz-Jeghers syndrome." *J Med Genet* **43**(5): e18.

Walther, A, Houlston, R and Tomlinson, I (2008). "Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis." *Gut* **57**(7): 941-50.

Walther, A, Johnstone, E, Swanton, C, Midgley, R, Tomlinson, I and Kerr, D (2009). "Genetic prognostic and predictive markers in colorectal cancer." *Nature Publishing Group* **9**(7): 489-499.

Wang, X, Lam, EK, Zhang, J, Jin, H and Sung, JJ (2009). "MicroRNA-122a functions as a novel tumor suppressor downstream of adenomatous polyposis coli in gastrointestinal cancers." *Biochem Biophys Res Commun* **387**(2): 376-80.

Wasielewski, M, Vasen, H, Wijnen, J, Hooning, M, Dooijes, D, Tops, C*, et al.* (2008). "CHEK2 1100delC is a susceptibility allele for HNPCC-related colorectal cancer." *Clin Cancer Res* **14**(15): 4989-94.

Wiesner, GL, Daley, D, Lewis, S, Ticknor, C, Platzer, P, Lutterbaugh, J*, et al.* (2003). "A subset of familial colorectal neoplasia kindreds linked to chromosome 9q22.2-31.2." *Proc Natl Acad Sci U S A* **100**(22): 12961-5.

Winney, B, Boumertit, A, Day, T, Davison, D, Echeta, C, Evseeva, I*, et al.* (2012). "People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population." *Eur J Hum Genet* **20**(2): 203-10.

Wong, HL, Peters, U, Hayes, RB, Huang, WY, Schatzkin, A, Bresalier, RS*, et al.* (2010). "Polymorphisms in the adenomatous polyposis coli (APC) gene and advanced colorectal adenoma risk." *Eur J Cancer* **46**(13): 2457-66.

Woodage, T, King, SM, Wacholder, S, Hartge, P, Struewing, JP, McAdams, M*, et al.* (1998). "The APCI1307K allele and cancer risk in a community-based study of Ashkenazi Jews." *Nat Genet* **20**(1): 62-5.

Wu, JS, Paul, P, McGannon, EA and Church, JM (1998). "APC genotype, polyp number, and surgical options in familial adenomatous polyposis." *Ann Surg* **227**(1): 57-62.

Yuan, HY, Chiou, JJ, Tseng, WH, Liu, CH, Liu, CK, Lin, YJ*, et al.* (2006). "FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization." *Nucleic Acids Res* **34**(Web Server issue): W635-41.

ANNEXES

ARTICLES RAPPORTES DANS LA THESE

A. Implication of MYH in colorectal polyposis (Ann Surg)
B. MYH biallelic mutation can inactivate the two genetic pathways of colorectal cancer by APC or MLH1 transversions. (Fam Cancer)
C. APC, MYH and the correlation gentoype-phenotype in colorectal polyposis. (Ann Surg Oncol)
D. Frequent mutation in North-African patients with MUTYH-Associated polyposis. (Clin Genet).
E. Cyclin D1 rare variants in UK multiple adenoma and early-onset colorectal cancer patients. (J Hum Genet)
F. Variants rares

Article A. Implication of MYH in colorectal polyposis (Ann Surg)

Implication of MYH in Colorectal Polyposis

Jérémie H. Lefevre, MD,† Christelle M. Rodrigue, MD,† Najat Mourra, MD,‡ Malika Bennis, MD,* Jean-François Flejou, MD,‡ Rolland Parc, MD, FACS,* Emmanuel Tiret, MD,* Christian Gespach, PhD,† and Yann R. Parc, MD, PhD†*

Objective: The aim of this study was to determine the frequency of MYH mutations in one large population of polyposis patients without APC mutation identified.

Summary Background Data: Familial adenomatous polyposis (FAP) is the most known inherited colorectal cancer syndrome. In 70% to 80% of polyposis patients, an APC mutation is found. Patients with polyposis but no APC mutation are considered as APC-muted patients and followed as their relatives accordingly. Biallelic mutation of MYH has been found to responsible of colorectal polyposis and cancer in an autosomal recessive pattern of inheritance.

Methods: Between 1978 and 2004, 433 patients were operated for polyposis. A mutation on APC was identified in 322 patients. Among the remaining patients, 44 were identified as possible MYH-muted patients and contacted, and 31 signed informed consent. Clinical data were obtained from the patients' medical notes. Germline mutation of MYH was searched by sequencing the whole gene. To confirm the deleterious effects of biallelic MYH mutation, transversions on K-*ras* and APC were searched.

Results: There were 9 women and 22 men with a mean age of 53.9 years (range, 22–68 years) at the time of diagnosis. The mean number of polyps was 62.8 (range, 11–266). Eighteen patients (58.1%) had a colorectal cancer. We found biallelic MYH mutation in 6 patients (19.3%; 95% confidence interval, 5.2%–33.5%) and 5 (83.3%) had transversions in K-*ras* and/or APC.

Conclusion: MYH is a new gene responsible for about 1.4% of all adenomatous polyposis and about 20% of adenomatous polyposis without APC mutation identified. Search for MYH biallelic mutation in these patients should be systematic as it changes their and relatives'surveillance.

(*Ann Surg* 2006;244: 874–880)

From the *Department of Digestive Surgery, the †Unit INSERM U673, and the ‡Department of Pathology, Hôpital Saint-Antoine (AP/HP), University Pierre et Marie Curie, Paris, France.

Reprints: Yann Parc, MD, PhD, Department of Digestive Surgery, Hôpital Saint-Antoine, Assistance Publique-Hôpitaux de Paris, Université Pierre et Marie Curie, Paris VI., 184 rue du Faubourg Saint-Antoine, 75012, Paris, France. E-mail: yann.parc@sat.ap-hop-paris.fr.

Copyright © 2006 by Lippincott Williams & Wilkins

ISSN: 0003-4932/06/24406-0874
DOI: 10.1097/01.sla.0000246937.54435.50

Colorectal cancer (CRC) is the third most frequent cancer in Western countries.[1] Inherited factors are thought to play a major role in sporadic colorectal carcinogenesis as more than 25% of all CRC are associated with some family history.[2] However, well-established CRC predisposal genes as APC,[3] MLH1, or MHS2[4] account for only a minority of cases. Nevertheless, they are implicated in 2 well-known genetic syndromes representing 3% to 5% of all CRC: familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC). FAP is the commonest adenomatous polyposis syndrome with an autosomal dominant pattern of inheritance. Classic FAP is characterized by the development of multiple adenomas (at least 100) during the second decade with a high risk of developing CRC;[5] when FAP is diagnosed after the age of 40 years, 73% of patients have already developed CRC.[6] Extracolonic manifestations of FAP include congenital hypertrophy of the retinal pigmented epithelium (CHRPE), desmoid tumors, and duodenal, periampullary or ampullary adenoma. Some patients develop attenuated FAP (AFAP), which is characterized by the presence of fewer polyps (<100), proximal colonic predominance of the polyps and the later age of onset of polyp and CRC development.[7]

Direct sequencing of the APC gene is considered to be the most accurate diagnostic test for FAP or AFAP.[8] APC mutations are responsible for 70% to 80% of the classic forms of FAP but are found in less than 30% in AFAP.[9–11] Possible explanations are germinal mosaïsm and large genetic deletions, which are not identified by classic techniques. Until recently, patients without mutation were considered as APC carriers and surveyed accordingly. Theirs relatives (siblings and offspring) had the same surveillance, which is associated with considerable cost as well as anxiety. The endoscopy protocol includes a full colonoscopy every year after the teenage years to search for adenomatous polyps.

In 2002, Al-Tassan et al[12] discovered that the biallelic mutation of MYH increased the risk of CRC. MYH[13] is a base excision repair (BER) gene like MTH1 or OGG1. It's a DNA glycosylase responsible for the removal of adenines from DNA that have been mispaired with 8-hydroxiguanine (8oxoG). 8oxoG is a nucleotide product of oxidative reaction and can readily mismatch with adenine.[14,15] If the BER system is deficient, as in the case of biallelic mutation of MYH, the mispaired adenines will lead to an accumulation of somatic transversions G:C→T:A in specific growth-regulatory genes such as APC or K-*ras*. Previous studies have

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

shown that patients with mutations in MYH have numerous polyps (but never thousands) and some extracolonic features as osteomas, duodenal polyps, and CHRPE.[16] The clinical differentiation between patients with de novo APC mutation (20% of all FAP patients[17]) and MYH mutated patients can be difficult as there is no polyposis history in either of the cases.

The aim of our study was to determine the frequency of MYH biallelic mutations in one large population of polyposis patients without APC mutation.

PATIENTS AND METHODS

Patients

Between January 1978 and September 2004, 433 patients were operated for colorectal polyposis at our institution. Among these patients, 351 had genetic testing for APC mutation. Patients with proven APC mutation were not included in our study. For patients not tested for APC mutation, personal and familial history was reviewed. Patients with a family history evoking an APC mutation (≥3 first-degree relatives developing colorectal polyposis) were also excluded from the study. The remaining patients were contacted and invited to participate in the study. Approval from the Commission National de l'Informatique et des Libertés (French National Data Processing Agency) was obtained (Bulletin Officiel de la Ville de Paris, May, 15, 2004).

Age, gender, cancer family history, and details of surgical procedures were obtained by data collection from patients' medical notes (family history of cancer being recorded by surgeons for each patient and routinely detailed in the patient's notes). The family history of all patients included in the present study was also checked by direct interview.

DNA Extraction

For each patient, tumor DNA and normal control DNA were extracted from frozen tissue sections using the Qiamp-Kit (Qiagen Inc, Santa Clarita, CA). For tumor DNA, only those areas containing >70% tumor cells were used. The corresponding normal control DNA for each patient was extracted from normal colonic tissue, which was checked by a histopathologist to ensure absence of tumor cells in the sample.

PCR Amplification and Sequencing

Sixteen primer pairs were used to amplified the coding region and exon/intron junctions of MYH (GenBank ID: 4595) as previously described by Al-Tassan et al.[12] PCR was performed on normal DNA to find germline mutation. To confirm the deleterious effects of biallelic MYH mutation, we looked for transversions on K-*ras* for every patient with available tumor DNA and on APC for every patient with a biallelic mutation on MYH. We amplified the first exon of K-*ras* (GenBank ID: 3845) and the mutation cluster region (codon 1250 to 1550) of APC (GenBank ID: 324) with 2 primers pairs (APC1: codon 1182 to 1415, APC2: codon 1390 to 1590). PCR reactions were performed in 50-μL reaction mixtures containing 60 ng of template DNA, 50 μmol/L of each oligonucleotide primer pair, 1.25 unit of AmpliTaq Gold DNA polymerase (Applied Biosystems, Applera, Courtaboeuf, France), and 0.2 mmol/L of dNTP (Invitrogen, Life Technologies, Cergy-Pontoise, France). Amplifications were realized in a Thermocycler GeneAmp PCR System 2700 (Applied Biosystems). The sequences of the primers used, the length of the PCR products, the $MgCL_2$ concentration, and the annealing temperature are shown in Table 1. PCR products were run on a 2% standard agarose gel

TABLE 1. Primers of MYH, K-*ras* and APC, Annealing Temperature

Exon	Sequence	Length (BP)	T°C
MYH 1	5′-GAAGCTGCGGGAGCTGAAA-3′/5′-ATCCCCGACTGCCTGAACC-3′	133	58
MYH 2	5′-CTGCATTTGGCTGGGTCTTT-3′/5′-CGCACCTGGCCCTTAGTAAG-3′	263	56
MYH 3	5′-AGCCTGTGCAGGGATGATTG-3′/5′-CAACCCCAGATGAGGAGTTAGG-3′	272	57
MYH 4	5′-CTCATCTGGGGTTGCATTGA-3′/5′-GGGTTGGCATGAGGACACTG-3′	167	52
MYH 5	5′-GGGCAGGTCAGCAGTGTC-3′/5′-TACACCCACCCCAAAGTAGA-3′	189	52
MYH 6	5′-TACTTTGGGGTGGGTGTAGA-3′/5′-AAGAGATCACCCGTCAGTCC-3′	185	54
MYH 7	5′-GGGACTGACGGGTGATCTCT-3′/5′-TTGGAGTGCAAGACTCAAGATT-3′	186	54
MYH 8	5′-CCAGGAGTCTTGGGTGTCTT-3′/5′-AGAGGGGCCAAAGAGTTAGC-3′	240	58
MYH 9	5′-AACTCTTTGGCCCCTCTGTG-3′/5′-GAAGGGAACACTGCTGTGAAG-3′	196	54
MYH 10	5′-GTGCTTCAGGGGTGTCTGC-3′/5′-TGTCATAGGGCAGAGTCACTCC-3′	262	60
MYH 11	5′-TAAGGAGTGACTCTGCCCTATG-3′/5′-GCCAAGAGGGGCTTTAGG-3′	248	54
MYH 12	5′-AGCCCCTCTTGGCTTGAGTA-3′/5′-TGCCGATTCCCTCCATTCT-3′	298	60
MYH 13	5′-AGGGCAGTGGCATGAGTAAC-3′/5′-GGCTATTCCGCTGCTCACTT-3′	242	54
MYH 14	5′-TTGGCTTTTGAGGCTATATCC-3′/5′-CATGTAGGAAACACAAGGAAGTA-3′	256	58
MYH 15	5′-TGAAGTTAAGGGCAGAACACC-3′/5′-GTTCACCCAGACATTCGTTAGT-3′	205	58
MYH 16	5′-AGGACAAGGAGAGGATTCTCTG-3′/5′-GGAATGGGGGCTTTCAGA-3′	224	60
K-*ras*	5′-GTACTGGTGGAGTATTTGAT -3′/5′-ACTCATGAAAATGGTCAG-3′	290	57
APC 1	5′-TGCCACAGATATTCCTTCATC-3′/5′-GCCACTTACCATTCCACTGC-3′	675	57
APC 2	5′-CTTCTGTCAGTTCACTTGATAG-3′/5′-GCTTTACGTGATGACTTTGTTG-3′	707	57

© 2006 Lippincott Williams & Wilkins

(InVitrogen), then eluted and purified using the QIAquick kit (Qiagen) according to the manufacture protocol. DNA fragments were then sequenced in both directions by MillGen Biotechnologies (Prologue Biotech, Labege, France).

Immunohistochemical Staining

Immunohistochemical staining for MSH2 and MLH1 was performed as previously described.[18]

RESULTS

Patients

In the 433 colonic polyposis patients operated at our institution between 1978 and 2004, including 199 relatives from 82 FAP families, APC sequencing was performed in 351 patients and a mutation was identified in 322 (92%). In 29 patients, no mutation was identified. Patients without APC mutation were statistically older than patients with a mutation (35.7 years vs. 28.0 years, $P = 0.0367$). Eighty-two patients had not been tested.

Among the 111 patients without an APC mutation, 47 had a family history of FAP and were excluded from the study. Of 64 patients, 11 had died and 9 were lost to follow-up. The remaining 44 patients were approached to participate in the study and 31 (70.4%) signed informed consent.

Clinical data of the patients are detailed in Table 2. There were 9 women and 22 men, and the mean age at the time of diagnosis was 53.9 years (±7.3; median, 55; range, 22–68 years). The mean number of polyps was 62.8 (±44.5; median, 40; range, 11–266). Eighteen patients (58.1%) were operated on as they had a cancer. Cancers were in the left colon or rectum in 58.1% (18 of 31 cancers). Two patients had extracolonic manifestations: patient 14 had hypertrophy of the pigmented retinal epithelium and 2 patients (patients 3 and 14) had duodenal polyps.

MYH Analysis

In 6 patients, a biallelic MYH mutation was found (19.3%; 95% CI, 5.2%–33.5%) (Fig. 1, Table 3). The most frequently reported missense changes Y165C and G382D were found in 3 patients. One or 2 MYH polymorphisms were identified in 10 patients (32.2%): Q324H (n = 9; 29.0%), V22M (n = 1; 3.2%) and S501F (n = 2; 6.4%) (Table 3).

TABLE 2. Clinical Data of the Patients

No.	Gender	Surgery	Familial History	Personal History	Age at the Time of Diagnosis	Polyps (n)	Cancer	Cancer Side	TNM
1	M	AIA	Grandfather: CRC		42	156	3	Right colon (n = 3)	pT3NOM0
2	M	AIR	Father & sister: polyps	Hyperplasic polyps	55	60	0		
3	M	AIA			67	70	1	Left colon	T2N0M0
4	M	AIR	Sister and niece: polyposis		60	100	0		
5	M	AIR			53	15	1	Right colon	T3N0
6	M	AIR		Polyps resected	51	56	0		
7	F	AIA	Mother: polyps		59	127	1	Right colon	T3N0M0
8	M	ACR->AIR			56	11	2	Left colon	T3N1M0
9	M	AIR			68	11	2	Right & left colon	T4N1M0
10	M	AIR			64	17	4	Right colon	T3N1M0
11	F	AIA	Grandfather & cousin: Polyps		34	200	0		
12	M	AIA	Polyposis in 2 brothers (#31)		55	266	3	Rectum	T3N0M0
13	M	AIA	Father: stomach cancer	Polyps resected	60	15	1	Rectum	T1N0M0
14	M	AIA		CHPER	60	30	2	Left colon & rectum	T2N1M1
15	F	AIA			64	32	1	Rectum	T2N0M0
16	M	AIR		Polyps resected	59	20	0		
17	M	ACR	Uncle: CRC		49	40	0		
18	F	AIR	Father: polyps		22	13	2	Left colon (2 times)	T2N0M0
19	F	AIR			57	100	0		
20	M	AIA			52	50–100	1	Left colon	Tis
21	F	AIR	Grandmother & aunt: CRC	Uterus cancer	61	12	2	Right & left colon	T2N0M0
22	M	ACA->AIA			55	20	2	Right & left colon	T2N0M0
23	F	AIR			50	32	1	Right colon	T1N0M0
24	F	AIR			64	24	1	Left colon	T1N0M0
25	M	AIR		Polyps resected	59	25	0		
26	M	AIA			43	40	0		
27	M	AIR	Father and sister: polyps		48	100	0		
28	F	AIR			41	30	0		
29	M	AIR	Brother: polyposis		53	70	0		
30	M	AIR			60	100	0		
31	M	AIA	Polyposis in 2 brothers (#12)		49	80	1	Left colon	T3N0M0

AIA indicates ileoanal anastomosis; AIR, ileorectal anastomosis; ACR, colorectal anastomosis, ACA, coloanal anastomosis.

© 2006 Lippincott Williams & Wilkins

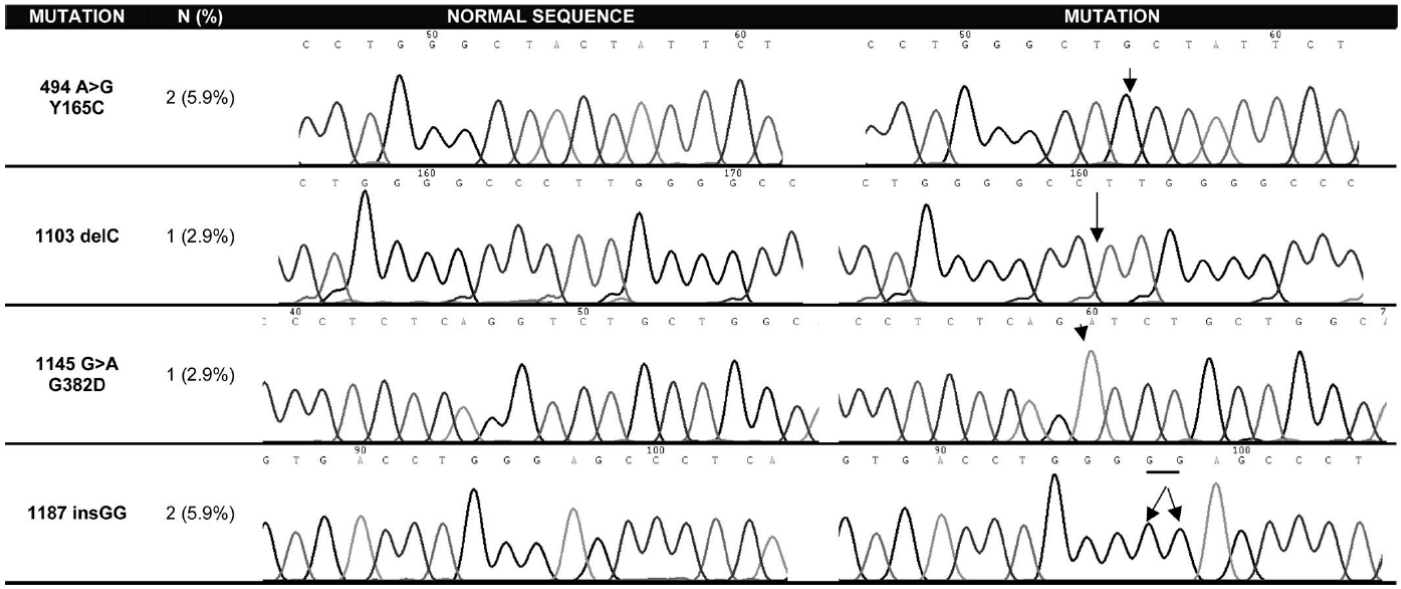


FIGURE 1. MYH biallelic mutations.

TABLE 3. MYH, K-*ras* and APC Mutations

Patient	MYH	K-*ras*	APC
1	**Y165C/Y165C**	**34 G>T, G12C***	
4	Q324H	†	
5	Q324H		
8		35 G>A, G12D	
9	Q324H	35 G>C, G12A	
10	Q324H/S501F		
12	**1187 ins GG/1187 ins GG**	**34 G>T, G12C***	**1554 T>A, 1867 G>T***
14	**1103 del C/1103 del C**		**1858 G>T***
15	**Y165C/Y165C**	**34 G>T, G12C***	**1579 G>T***
16	Q324H		
19	Q324H/S501F	†	
20	V22M	35 G>C, G12A	
25	Q324H		
26	**G382D/G382D**		**1609 G>T***
28	Q324H		
30	Q324H		
31	**1187 ins GG/1187 ins GG**		**1554 T>A**

Patient in bold font had pathogenic biallelic mutation on MYH.
*Transversion G:C→T:A.
†No tumoral DNA available.

K-ras and APC Analysis

Amplification of the first exon of K-*ras* was performed in 24 patients. Six patients were found to have a mutation (25%). Three patients who had a biallelic mutation of MYH also had a mutation of K-*ras*, and in all cases it was a transversion G:C→T:A. The 3 others mutations on K-*ras* happened in patients without a pathogenic variant of MYH, and none of them was a transversion (Table 3). The study of APC in the 6 patients muted on MYH showed 4 transversions and 2 other mutations (two T>A in the 2 brothers, patients 12 and 31) (Table 3). All transversions observed in K-*ras* or APC occurred in *G*AA or *G*AAAA sequences. With these 2 sequences, we found in 5 patients (83.3%) with a biallelic mutation on MYH, 1 or 2 transversions in K-*ras* and/or APC.

Immunohistochemical Staining

Among the 31 patients who were included in this study, only 1 patient (patient 31) lost expression of MLH1.

Clinical Characteristics of Biallelic MYH Mutation Patients

There was 5 males and 1 female patients that demonstrated a biallelic MYH mutation. The mean number of polyps was 100 (median, 60; range, 30–266). One third of the

© 2006 Lippincott Williams & Wilkins

patients with MYH mutation had more than 100 polyps. Among the 22 patients with more than 20 polyps, the frequency of biallelic mutation on MYH was 27.2% (95% CI, 8.2%–35.5%), and none of the patients with less than 20 polyps was found to have a MYH biallelic mutation. The mean age of patients with MYH mutation at the time of the operation was 52.2 years (median, 52 years; range, 42–64 years). Two extracolonic features were seen in 1 patient: duodenal polyposis and CHRPE. Ten cancers occurred in 5 of the 6 patients (85.6%) with biallelic MYH mutation. The cancers were mainly localized on the left colon (n = 2 in the colon, n = 5 in the rectum).

DISCUSSION

Beside classic FAP with a germline mutation of APC, some patients have numerous adenomatous polyps, a high risk of CRC, and 2 wild-type copies of APC. Since 2002, the MYH gene, which belongs to the BER system involved in the DNA repair of oxidative lesion, is known to predispose to recessive inheritance of numerous adenomatous polyps.[12] The aim of our work was to study the frequency of MYH mutations in one large single-center population of polyposis patients without APC mutation and in a further step to adapt the surveillance offered to the relatives of newly diagnosed polyposis patients without APC mutation.

Among our 31 patients without an APC mutation, 6 (19.3%) had a biallelic germline mutation of MYH. Two of them had more than 100 polyps. Given this observation, the number of polyps cannot be viewed as a pathognomonic feature of FAP any more. Five of the 6 patients had CRC; they were not known to have colorectal polyposis and only diagnosed when they became symptomatic. Our group[6] has shown that in 141 patients with FAP and mutation of APC, 71 were symptomatic at the time of the operation (as were the MYH patients in this study). The mean age of those patients was 40 years and 32 (45%) had a cancer at the time of the operation. It therefore appears that polyps and cancers develop later in patients with biallelic mutation of MYH than in those with APC mutation. One patient had extracolonic features: CHRPE (n = 1) and duodenal polyps (n = 1). Thus, confronted with a patient with adenomatous polyposis, the 2 main clinical features that should evoke MYH biallelic mutation are the age at the time of diagnosis and the absence of a family history of polyposis, especially in siblings and offspring. Concerning sex ratio, if in our study 5 of the 6 patients with biallelic MYH mutation were males, it seems that this observation is purely incidental. Thus, gender should not be considered for MYH mutation search. If no mutation is found at APC testing, MYH genetic testing should be undertaken.

Most investigators explore only the exons 7 and 13[19,20] where the 2 most frequent mutations in MYH Y165C and G382D occur. We would have diagnosed only 3 of our 6 patients with MYH biallelic mutation had we followed this strategy. We therefore chose to explore the entire coding sequence of MYH. Moreover, recently novel mutations of other exons have been described[21] as well as specific mutations of MYH in different ethic populations such as E466X in Indian cases,[22] which demonstrates the importance of screening the whole coding sequence of MYH.

Among our 6 patients muted on MYH, 5 had typical transversions on APC and/or K-*ras* demonstrating the pathogenicity of biallelic MYH mutation. Patient 31 was the only one in whom no transversion could be identified. However, in this patient, a loss of expression was identified by immunohistochemical staining, demonstrating a tumor with microsatellite instability phenotype. A search for mutation or methylation of hMLH1 was performed and showed no abnormality. Study of the DNA sequence of hMLH1 (GenBank 29729888) shows the presence of several GAA sequences that could be the site of transversion and then explain the inactivation of hMLH1 by the MYH pathway. Moreover, patient 31 is the brother of patient 12 (both patients having the same MYH mutation: 1187insGG/1187insGG) who was found to have a microsatellite stable tumor with normal expression of hMLH1 and hMSH2 but transversions on K-*ras* and APC. Thus, biallelic MYH mutation could induce tumorogenesis by the 2 pathways known in CRC, mainly by inactivation of APC and K-*ras* and in some case also by inactivating hMLH1 and thereby the mismatch repair system.

Despite the search of germline mutation in MYH, there is still no explanation for the polyposis in 25 of the 31 patients (80.6%) included in this study. These patients with no genetic etiology are still considered as APC mutant and they and their families are submitted to intensive clinical and endoscopy follow-up. Alternate splicing modifications of APC or MYH and unknown genes implicated in the colonic cancers may be the explanation for these polyposis cases. Lipton et al[23] have recently proven the involvement of germline mutation in BMPR1A in adenomatous polyposis. This gene belongs to the TGF-β family known to cause the juvenile polyposis syndrome.

Depending on number and location of the polyps, patients with MYH biallelic mutations should undergo total colectomy or restorative proctocolectomy, just like APC-muted patients. As previously described,[16] we found 1 patient with duodenal polyposis; surveillance of the upper gastrointestinal tract for duodenal polyps should therefore be continued. The main difference with FAP concerns the siblings and offspring. When an APC mutation is identified, the risk for each child to have FAP is 50%, whereas it is less than 1% (heterozygote frequency[24] divided by 2) for a child of a MYH biallelic muted parent. For the siblings of an APC muted patient, the risk is 50%, compared with 25% with a MYH biallelic muted parent. Moreover, identification of biallelic mutation on MYH in a patient with FAP or AFAP phenotype but without APC mutation allows specific investigation of the relatives. Surveillance of siblings may then be modified accordingly. Nevertheless, the real risk for monoallelic carrier of pathogenic mutation of MYH is still unknown. As in every disease with a recessive pattern of transmission, heterozygous patients should be healthy, but Kambara et al[25] showed a higher level of loss of heterozygosity in 1p where the locus of MYH is localized and a higher level of transversion in APC and K-*ras* in patients with a single pathogenic variant of MYH. There was, however, no significant differ-

© 2006 Lippincott Williams & Wilkins

ence of CRC rates between the general population and the monoallelic carriers of MYH.

Reducing the number of unnecessary colonoscopies would save money and avoid the discomfort of bowel preparation, the risks of anesthesia, and of the examination itself as well as any associated anxiety. For those polyposis patients with neither mutation of APC nor MYH, other gene mutations should be searched for, but until another explanation is found they have to be surveyed as if they were APC mutants.

CONCLUSION

MYH is a gene responsible for 1.4% of all adenomatous polyposis and 20% of adenomatous polyposis without mutation of the APC gene. The risk of transmission to offspring is less than 1% as it is a recessive inherited disease. Colonoscopic surveillance can thus be dramatically reduced in the offspring of the index patient. For these reasons, we propose searching for MYH biallelic mutations to all patients with colorectal polyposis without APC mutation.

REFERENCES

1. Landis SH, Murray T, Bolden S, et al. Cancer statistics, 1999. *CA Cancer J Clin*. 1999;49:8–31.
2. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol*. 2001;96:2292–2303.
3. Kinzler VW, Nilbert MC, Su LK, et al. Identification of FAP locus genes from chromosome 5q21. *Science*. 1991;253:661–665.
4. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell*. 1996;87:159–170.
5. Lynch HT, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet*. 1999;36:801–818.
6. Penna C, Tiret E, Kartheuser AH, et al. Value of screening of familial adenomatous polyposis for prevention of colorectal cancer. *Gastroenterol Clin Biol*. 1992;16:210–214.
7. Lynch HT, Smyrk T, McGinn T, et al. Attenuated familial adenomatous polyposis (AFAP): a phenotypically and genotypically distinctive variant of FAP. *Cancer*. 1995;76:2889–2892.
8. Giardiello FM, Brensinger JD, Petersen GM. AGA technical review on hereditary colorectal cancer and genetic testing. *Gastroenterology*. 2001; 121:198–213.
9. Lamlum H, Al Tassan N, Jaeger E, et al. Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q. *Hum Mol Genet*. 2000;9:2215–2221.
10. Armstrong JG, Davies DR, Guy SP, et al. APC mutations in familial adenomatous polyposis families in the Northwest of England. *Hum Mutat*. 1997;10:376–380.
11. Heinimann K, Thompson A, Locher A, et al. Nontruncating APC germ-line mutations and mismatch repair deficiency play a minor role in APC mutation-negative polyposis. *Cancer Res*. 2001;61:7616–7622.
12. Al-Tassan N, Chmiel NH, Maynard J, et al. Inherited variants of MYH associated with somatic G:C→T:A mutations in colo-rectal tumors. *Nat Genet*. 2002;30:227–232.
13. Slupska MM, Baikalov C, Luther WM, et al. Cloning and sequencing a human homolog (MYH) of the Escherichia coli MutY gene whose function is required for the repair of oxidative damage. *J Bacteriol*. 1996;178:3885–3892.
14. Slupphaug G, Kavli B, Krokan HE. The interacting pathways for prevention and repair of oxidative DNA damage. *Mutat Res*. 2003;531: 231–251.
15. Ames BN, Gold LS. Endogenous mutagens and the causes of aging and cancer. *Mutat Res*. 1991;250:3–16.
16. Sieber OM, Lipton L, Crabtree M, et al. Multiple colorectal adenomas, classic adenomatous polyposis ans germ-Line mutations in MYH. *N Engl J Med*. 2003;348:791–799.
17. Grady WM. Genetic testing for high-risk colon cancer patient. *Gastroenterology*. 2003;124:1574–1594.
18. Rigau V, Sebbagh N, Olschwang S, et al. Microsatellite instability in colorectal carcinoma: the comparison of immunohistochemistry and molecular biology suggests the interest of hMSH6 immunostaining. *Arch Pathol Lab Med*. 2003;127:694–700.
19. Gismondi V, Meta M, Bonelli L, et al. Prevalence of the Y165C, G382D and 1395delgga germline mutations of the *myh* gene in Italian patients with adenomatous polyposis coli and colorectal adenomas *Int J Cancer*. 2004;109:680–684.
20. Enholm S, Hienonen T, Suomalainen A, et al. Proportion and phenotype oh MYH-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients. *Am J Pathol*. 2003;163:827–832.
21. Alhopuro P, Parker AR, Lehtonen R, et al. A novel functionally deficient MYH variant in individuals with colorectal adenomatous polyposis. *Hum Mutat*. 2005;26:393.
22. Sampson JR, Dolwani S, Jones S, et al. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet*. 2003;362:39–41.
23. Lipton L, Sieber OM, Thomas HJW, et al. Germline mutations in the TGF-β and Wnt pathways are a rare cause of the multiple adenoma phenotype. *J Med Genet*. 2003;40:e35.
24. Wang L, Baudhuin LM, Boardman LA, et al. MYH mutations in patients with attenuated and classic polyposis and with young-onset colorectal cancer without polyps. *Gastroenterology*. 2004;127:9–16.
25. Kambara T, Whitehall VLJ, Spring KJ, et al. Role of inherited defects of MYH in the development of sporadic colorectal cancer. *Genes Chromosomes Cancer*. 2004;40:1–9.

Discussions

DR. NEIL MORTENSEN: Thank you for a beautifully presented and explained paper. Life used to be so simple. FAP was all about abnormalities of the APC gene, and now we find that is not the case. While APC gene abnormalities are the cause in about 80% of patients with FAP, you've looked at that other 20% and you've only found a small proportion having MYH abnormalities. I wonder if you could please tell us the answers to 2 questions?

Are there patients within the APC group who also have MYH gene abnormalities, which would, if you like, give a more severe phenotype, perhaps?

Do you think that there is, within that residual, let's say 15% to 18% of patients in whom you haven't found any abnormality, the place for some kind of international collaboration using GNY searches on huge numbers of patients to try and find out what's going on in those last 15% to 18%?

DR. YANN PARC: No, we did not find any MYH mutations in patients already with an APC mutation; however, we tested only a few patients with an APC mutation. In the literature, such association of mutations has not been reported and, yes, we are ready to collaborate in a study to find other genes. A family case has been reported in which the BRMA gene was considered to be responsible for the polyposis. However, much larger series are required to confirm this finding and allow the identification of other genes that could be responsible for colorectal polyposis.

DR. MARIO MORINO: This is a very interesting study but, in the abstract, you have proposed a different follow-up strategy for these patients and you did not speak of this point in your presentation. Do you think there are clinical implica-

© 2006 Lippincott Williams & Wilkins

Article B. MYH biallelic mutation can inactivate the two genetic pathways of colorectal cancer by APC or MLH1 transversions. (Fam Cancer)

Author's personal copy

Familial Cancer (2010) 9:589–594
DOI 10.1007/s10689-010-9367-0

MYH biallelic mutation can inactivate the two genetic pathways of colorectal cancer by APC or MLH1 transversions

Jérémie H. Lefevre · Chrystelle Colas · Florence Coulet · Carolina Bonilla · Najat Mourra · Jean-Francois Flejou · Emmanuel Tiret · Walter Bodmer · Florent Soubrier · Yann Parc

Published online: 18 July 2010
© Springer Science+Business Media B.V. 2010

Abstract MYH associated polyposis is a hereditary syndrome responsible for early colorectal cancer with a distinct genetic pathway from the Familial Adenomatous Polyposis or the Hereditary Non Polyposis Colorectal Cancer syndrome. We have studied a family with three members bearing a biallelic mutation in MYH at c.1185_1186dup. One patient who developed colon cancer had loss of expression of MLH1 on tumoral tissue and microsatellite instability (MSI) phenotype. Analysis of MLH1 based on his blood sample revealed no germline mutation or large genomic deletion. No methylation of the promoter was identified in tumoral DNA. No transversion mutations were identified in *APC* or *KRAS* in tumor DNA of this patient. Loss of expression of MLH1 was due to a transversion in intron 7 at position +5 (c.588 + 5G > T) leading to a complete deletion of exon 7 at the RNA level. This observation demonstrates that MLH1 can be a target of MYH transversions leading to MSI phenotype.

Keywords MYH gene · APC gene · Adenomatous familial polyposis · MYH Associated polyposis · Colorectal cancer · MLH1 gene

J. H. Lefevre (✉) · E. Tiret · Y. Parc
Department of Digestive Surgery, Hôpital Saint-Antoine AP-HP, University Pierre et Marie Curie Paris VI, 184 rue du Faubourg Saint-Antoine, 75571 Paris, France
e-mail: jeremie.lefevre@sat.aphp.fr

J. H. Lefevre · C. Bonilla · W. Bodmer
Department of Clinical Pharmacology, Cancer and Immunogenetics Laboratory, University of Oxford, Old Road Campus Research Building, Old Road Campus, Oxford OX3 9DQ, UK

C. Colas · F. Coulet · F. Soubrier
Laboratory of Oncogenetics and Angiogenetic, Groupe hospitalier Pitié-Salpêtrière, University Pierre et Marie Curie Paris VI, Paris, France

N. Mourra · J.-F. Flejou
Department of Pathology, Hopital Saint-Antoine, University Pierre et Marie Curie Paris VI, Paris, France

J. H. Lefevre · C. Colas · N. Mourra · J.-F. Flejou · E. Tiret · Y. Parc
Hopital Saint-Antoine AP-HP, 184 rue du Faubourg Saint-Antoine, 75012 Paris, France

C. Colas · F. Coulet · F. Soubrier
Groupe hospitalier Pitié-Salpêtrière, 47-83 bld de l'Hopital, 75013 Paris, France

Introduction

Several hereditary syndromes have been implicated in colorectal cancer (CRC) and the tree most frequent syndromes are: familial adenomatous polyposis (FAP) due to *APC* gene mutation (GeneID: 324), Hereditary Non Polyposis Colorectal Cancer (HNPCC) due to MisMatch Repair (MMR) gene mutation (mostly: *MLH1* (GeneID 4292) and *MHS2* (GeneID 4436)) and more recently, MYH Associated Polyposis (MAP) syndrome due to *MUTYH*, a Base Excision Repair (BER) gene mutation (GeneID 4595). Thus far, there is no description of the microsatellite instability (MSI) phenotype resulting from constitutional MYH biallelic mutations. We have studied a family with three members bearing a biallelic mutation in MYH at c.1185_1186dup. One patient who developed colon cancer had loss of expression of MLH1 on tumoral tissue and microsatellite instability (MSI) phenotype.

Materials and methods

DNA isolation

Genomic DNA was purified from peripheral blood leukocytes using the QIAamp DNA blood mini kit and tumoral DNA was extracted from frozen tissue sections using the QIAamp DNA miniKit® following the manufacturer's instructions (QIAGEN, Courtaboeuf, France).

PCR amplification and sequencing of the coding regions of hMLH1 and MYH

All *hMLH1 and MYH* exons were amplified by polymerase chain reaction (PCR) using primer designed with primer3 software (Frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). The resulting sequence data were analyzed with the SeqScape software, version 2.5 (Applied Biosystems, Foster City, CA) and aligned with hMLH1 and MYH reference sequences (Accession Number respectively: NM_000249.1 and NM_012222).

Somatic mutations on BRAF

Analysis of BRAF was performed by automatic sequencing. The fragment encompassing exon 15 was amplified by PCR with primers forward: 5′-TTGACTTCTAAGAGGAA AGATGAAGT-3′ and reverse 5′-AGCATCTCAGGGCCA AAAAT-3′.

Somatic mutations on APC and/or K-ras

To confirm the deleterious effects of biallelic MYH mutation, we looked for transversions on K-ras for every patient where tumour DNA was available and on APC for every patient with a biallelic mutation on MYH. We amplified the first exon of K-ras (GenBank ID: 3845) and the mutation cluster region (codon 1250–1550) of APC (GenBank ID: 324) with two primer pairs (APC1: codon 1182–1415, APC2: codon 1390–1590).

Multiplex ligation-dependent probe amplification (MLPA)

The SALSA MLPA P003 MLH1/MSH2 probemix kit (MRC-Holland®, Netherlands. www.mrc-holland.com) was used according to the manufacturer's instructions to screen rearrangements of one or more exons of the *hMLH1, hMSH2* genes. The multiplex probe mix contained 16 exon probe pairs for MSH2 and 19 exon probe pairs for MLH1 and 7 control probe pairs specific for DNA sequences outside the MSH2 ad MLH1 genes. Normalization of samples was done against the mean control ratios. DNA samples from unaffected individuals were used as controls in each experiment. Patient samples were compared to unaffected individuals samples. Peak heights of each fragment were compared to those of control samples and deletions or duplications were suspected when peak height differed by more than 30%.

Methylation status of MLH1 and MLH2

The methylation status of the promoter regions of both MLH1 and MSH2 was assessed as described previously [1], using an *Hpa*II based PCR assay and a Methylation Specific PCR (MSP). All DNA samples were chemically treated with sodium bisulfite according to the CpGenome™ DNA Modification Kit and subsequent methylation-specific PCRs were carried out as described by the manufacturer (Serologicals Corporation). Two PCRs were performed for each chemically modified DNA sample, one with primers annealing specifically to the methylated gene and the other to the originally unmethylated DNA. PCR products were separated on 6% non-denaturating polyacrylamide gels (Invitrogen™) and bands were visualized by staining with ethidium bromide. All PCRs were performed with positive controls for both methylated and unmethylated DNA.

Microsatellite status

A pentaplex PCR using the mononucleotide markers, BAT-25, BAT-26, NR-21, NR-22 and NR-24, was performed according to a previously reported technique [2].

Extraction of RNA and cDNA synthesis

For preparation of RNA, total RNA was extracted from cells using Trizol® reagent according to the manufacturer's instructions. After isolation, the RNA was used for cDNA synthesis using 0.5 μg of total RNA from each sample, $pd(N)_6$ random hexamer and dNTPs (1 mM), 200U M-MLV reverse transcriptase, and 20 U RNase inhibitor. Reverse transcription was carried out for 50 min at 37°C and stopped by heating (to) 65°C for 5 min.

Results

Family H and MYH mutations

Three members of a Morocco family who had surgery for colorectal cancer associated with adenomatous polyps were included in the study (Fig. 1). Patient II_1, a 55-year old male, was transferred to our unit because of rectal bleeding that lead to the diagnosis of diffuse colorectal polyposis. A restorative proctocolectomy with an ileal pouch anal

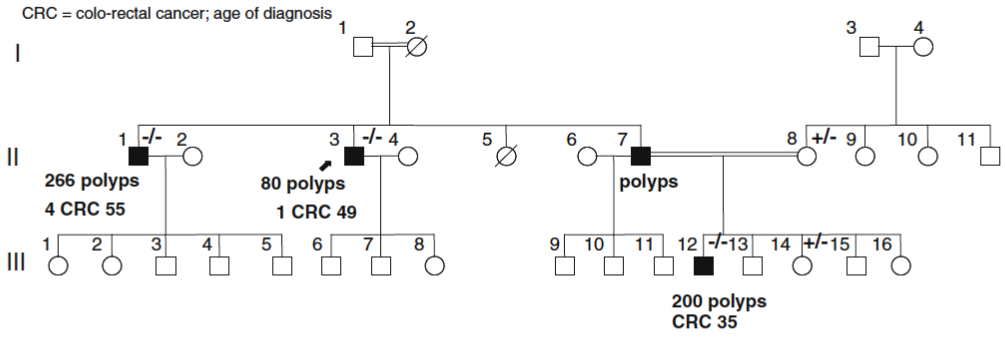


Fig. 1 Pedigree of familly H. *CRC* colo-rectal cancer; age of diagnosis

anastomosis was performed. Histopathological examination revealed 4 colorectal cancers and 266 adenomatous polyps. Systematic immunohistochemical staining of *MLH1* and *MSH2* proteins on tumoral and normal tissues showed a normal expression of both proteins. In this young patient without a family history of CRC with multiple polyps, a search for a de novo *APC* mutation was performed. Sequencing of the entire *APC* gene did not reveal any germline mutation. The 16 exons of *MYH* were analysed and biallelic mutations c.1185_1186dup, p.Glu396GlyfsX43 were identified (Fig. 2). During the time period required for identification of the mutation, the patient's relatives were offered colonoscopy. Two relatives, a brother (II_3) and a nephew (III_{12}) were found to have multiple polyps and the same biallelic mutations in *MYH*.

MLH1 inactivation

Immunohistochemical staining of *MLH1* revealed a loss of expression in the cancer tissue of patient II_3 and was normal for patient III_{12}. Instability was confirmed in patient II_3's tumour by PCR on mononucleotide markers BAT-25, BAT-26, NR-21, NR-22 and NR-24. No MSI phenotype was found for patient II_1 and for patient III_{12}.

Methylation analysis

We studied the methylation status of the promoter regions of *MLH1* by using an *Hpa*II based PCR assay and a methylation specific PCR. The promoter was unmethylated. We screened for the V600E mutation of the *BRAF* gene in tumoral tissue of patient II_3 by sequencing. Both alleles of *BRAF* were wild type.

MLH1 sequencing

No germline point-mutation was identified from DNA sequencing. To identify large deletions, we used Multiplex ligation-dependent probe amplification (MLPA). No germline deletion or duplication was identified in either the *MLH1* or the *MSH2* genes. As there was no germline mutation or methylation that could explain the loss of expression of *MLH1* and the MSI phenotype of patient II_3's tumour, we looked for the presence of a somatic mutation in *MLH1*.

Somatic mutation of MLH1

Sequence analysis of *MLH1* revealed that 54 GAA sequences were present within the 756 triplets of the coding sequence. Of these, 34 GAA sequences coded for

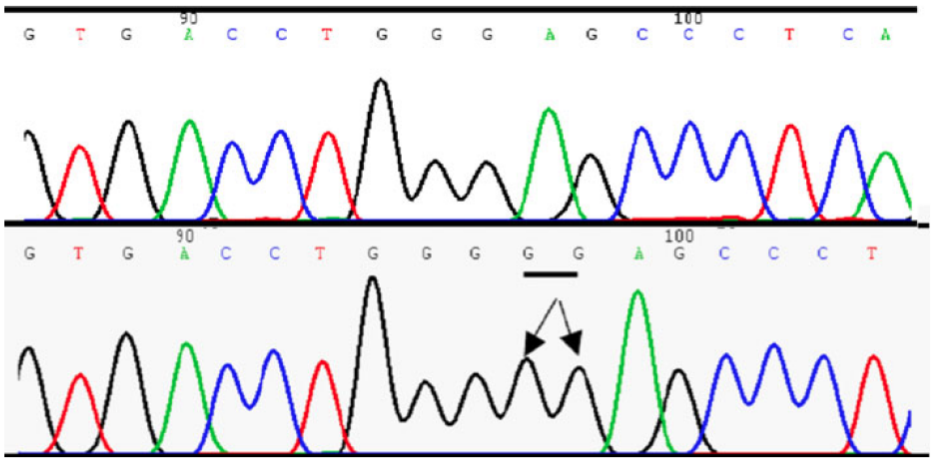


Fig. 2 Normal sequence (*upper sequence*) and patient II_2 lymphocyte DNA (*lower sequence*) of MYH gene showing 1186_1187insGG germline mutation

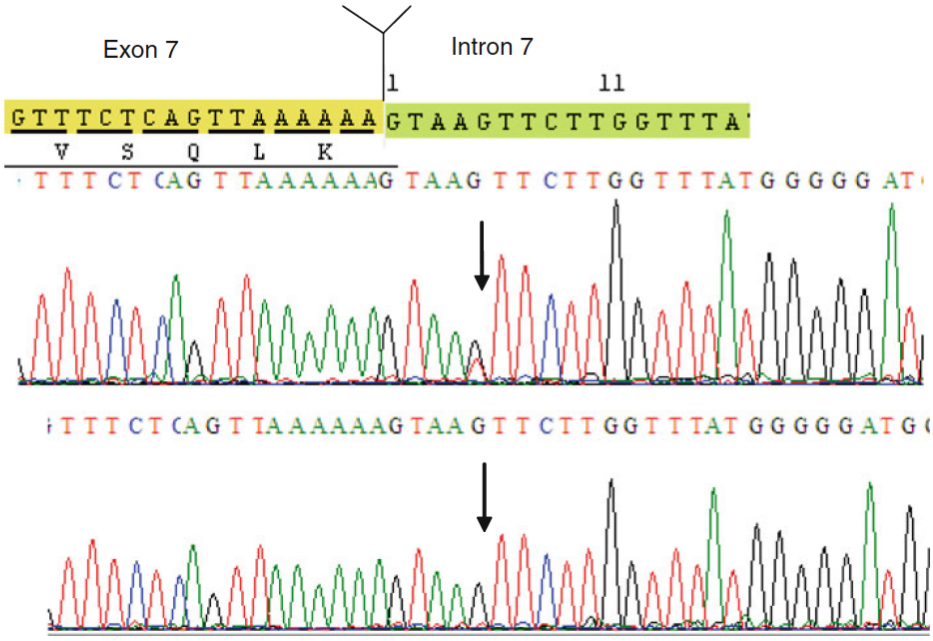


Fig. 3 Sequences of MLH1 in tumoral tissue (*upper sequence*) and in lymphocyte DNA (*lower sequence*) showing a transversion G > T in the intron 7 at position +5

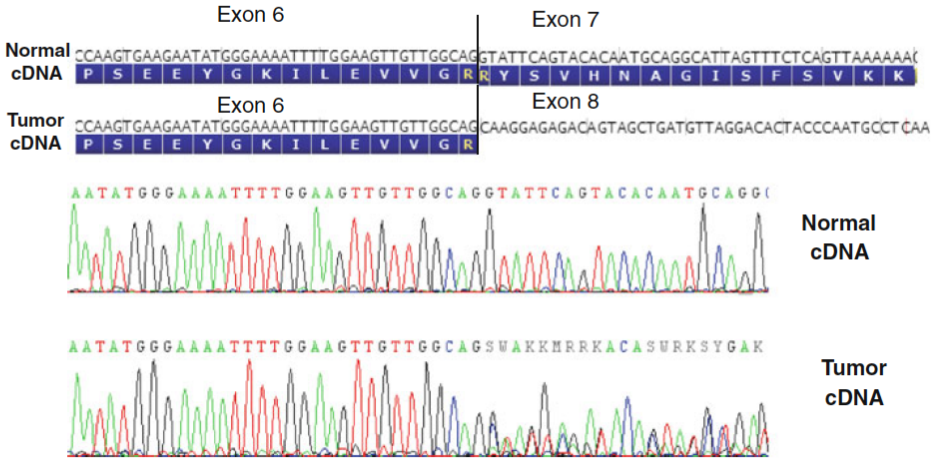


Fig. 4 Impact of c.588 + 5G > T mutation on RNA replication of MLH1

glutamic acid, which can result in a stop codon (TAA) after a G > T transversion. As transversion may also occur on other triplets, we performed an analysis of the entire coding sequence of *MLH1* in tumoral DNA compared with the normal sequence found in the blood sample. This analysis revealed a transversion in intron 7 at position +5 (c.588 + 5G > T) (Fig. 3). This transversion occurred in a AAGTT sequence. RNA extraction and RT–PCR, analysis of the cDNA of MLH1 showed that exon 7 was completely deleted leading to a nonfunctional RNA (Fig. 4).

Discussion

Since the discovery of *MYH* by Al-Tassan et al. [3], all patients with colonic polyposis and either a family history of CRC or young age at the time of diagnosis and no germline mutation on *APC* should be screened for *MYH* mutation. Unlike the usual recessive autosomal transmission of polyposis related to *MYH*, the pseudo-dominant pattern of inheritance in family H was explained by the consanguinity between patients at two generations (Fig. 1). Despite this atypical presentation, the endoscopic findings were similar to the MAP syndrome already described.

MLH1 inactivation was proven by the loss of expression and the microsatellite instability. Numerous publications have shown a high correlation between microsatellite instability and immunohistochemical staining showing loss of the MLH1 protein [4].

The loss of expression of MLH1 in a biallelic MYH mutated patient has already been described [5] but it was due to a methylation of the promotor. In most of the sporadic tumours with MSI phenotype, *MLH1* is also inactivated by epigenetic modification of its promoter [1, 6]. The

MLH1 promoter undergoes transcriptional silencing by hypermethylation at CpG islands (i.e., clusters of cytosine residues followed by a guanosine) and is responsible for the lack of expression of this protein, and accounts for the MSI phenotype [7]. In the present study, methylation of the promotor was ruled out by *Hpa*II based PCR assay and by BRAF amplification. The somatic mutation V600E of the *BRAF* gene (GeneID: 673) has been described to be associated with promoter methylation of *MLH1* [8] in more than 80% of cases but not with the presence of germline mutations in *MLH1* or *MSH2* [9]. These two observations provided evidence against methylation being the basis of the MSI phenotype of patient II_3's tumour.

Even though the presence of multiple adenomas made the HNPCC syndrome unlikely, a MLH1 sequencing was performed and revealed no germinal mutation. Large deletions in DNA MMR genes, in which the breakpoints occur outside the exons are not detectable by routine DNA sequencing but can be responsible of more than 22% of mutation negative MSI patients [10]. No germline deletion or duplication was identified in either the *MLH1* or the *MSH2* genes by MLPA.

MYH is a base excision repair (BER) gene like *MTH1* or *OGG1*. It is a DNA glycosylase responsible for the removal of adenines from the DNA, that have been mispaired with 8-hydroxiguanine (8oxoG). 8oxoG is a nucleotide product of oxidative reaction and can readily pair inappropriately with adenine [11]. If the BER system is deficient, as in the case of biallelic germline mutations of *MYH*, the mispaired adenines lead to an accumulation of somatic transversions G:C → T:A in specific growth-regulatory genes for which mutations are selected for in cancers, such as *APC* or *KRAS* [12, 13]. Transversions leading to a stop codon often occur in GAA sequences in *APC* and in codons 12 or 13 in *KRAS*. In order to confirm the involvement of *MYH* in the clinical syndrome of siblings II_1 et II_3, we extracted tumoral DNA from the histological specimens in order to identify somatic transversions G:C → T:A. We have already published the status of patients from other families with biallelic *MYH* mutations treated in our institution [12]. These patients had normal immunohistochemical staining for *MLH1* and *MSH2*. We sequenced the first exon of *KRAS* and the mutation cluster region of *APC* between codons 1182 and 1590 in tumor DNA. In all patients except patient II_3 one or more transversions were found in *KRAS* and/or *APC* [12].

As there was no germline mutation or methylation that could explain the loss of expression of *MLH1* and the MSI phenotype of patient II_3's tumour, we looked for the presence of a somatic mutation in *MLH1*. The transversion found in intron 7 at position +5 (c.588 + 5G > T) occurred in a AAGTT sequence. Surprisingly, none of the GAA sequences was found to undergo transversion. The GAA sequence specificity of *MYH* targeting seems limited to *APC* as *K-ras* transversion occurred in a GGT sequence and the *MLH1* transversion occurred in a GTT sequence. This mutation has never been described before in the international databases (InSIGHT, Mismatch Repair Genes Variants Database, www.insight-group.org) but another base change at the same position (c.588 + 5G > A) has been found as a germline mutation in at least 3 different HNPCC families (one in France (French database), one in Austria [14] and one in Germany [15]). Aberrant splicing with partial loss of exon 7 was demonstrated by different teams for this mutation leading to its classification as pathogenic. Each patient with a c.588 + 5G > A germline mutation described in the literature or in a database had a MSI phenotype in tumours and belonged to a family fulfilling the Amsterdam or Bethesda Criteria. Considering these data we assumed that the somatic mutation c.588 + 5G > T could also have a detrimental influence on splicing as confirmed by the complete deletion leading to a nonfunctional RNA.

Loss of expression of *MLH1* must involve two mutations or one mutation and loss of the wild type allele by LOH or one mutation and loss of expression of the wild type allele by methylation. In HNPCC tumours, somatic inactivation of the remaining wild-type allele may occur through different mechanisms: somatic mutation, promoter hypermethylation or loss of heterozygosity (LOH). The first two events have been eliminated by our previous studies on tumoral *MLH1* sequence. Loss of the wild-type allele in HNPCC tumours was shown by Hemminki et al. [16] and has been found to be the major mechanism for somatic second hits in most of the studies [17], even if up to 40% of the LOH events targeted the mutant allele [18]. We found no LOH on tumoral DNA of patient II_3 and RT_PCR on tumoral DNA showed the expression of both alleles. Moreover, the model of two hits is based on a first germline mutation followed by a somatic mutation, while in our patient both hits are somatic. The second hit may reside deep within an intron or at considerable distance upstream or downstream as described by others authors [19]. However, the results of the immunohistochemistry and the MSI phenotype indicates that both alleles of MLH1 were non functional.

Our findings implicate for the first time *MLH1* as a target of transversion by inactivation of *MYH* and the new mutation c.588 + 5G > T on MLH1. This mutation happened to arise as a result of the BER deficiency and was capable of providing a selective advantage to the tumor. The BER and MMR systems are already known to be related: MYH interacts with the MMR system and particularly with the hMUTSα complex (hMSH2/hMSH6) in order to localize the DNA daughter strand [20]. The MMR proteins are also implicated in the removal of 8-oxo-G like MYH [21]. Further studies are needed to explore the interaction between MYH and MLH1.

References

1. Cunningham JM, Christensen ER, Tester DJ, Kim CY, Roche PC, Burgart LJ et al (1998) Hypermethylation of the *hMLH1* promoter in colon cancer with microsatellite instability. Cancer Res 58:3455–3460
2. Parc Y, Gueroult S, Mourra N, Serfaty L, Flejou JF, Tiret E et al (2004) Prognostic significance of microsatellite instability determined by immunohistochemical staining of MSH2 and MLH1 in sporadic T3N0M0 colon cancer. Gut 53:371–375
3. Al-Tassan N, Chmiel NH, Maynard J et al (2002) Inherited variants of MYH associated with somatic G:C → T:A mutations in colorectal tumors. Nat Genet 30(2):227–232
4. Stone JG, Robertson D, Houlston RS (2001) Immunohistochemistry for MSH2, MHL1: a method for identifying mismatch repair deficient colorectal cancer. J Clin Pathol 54(6):484–487
5. Colebatch A, Hitchins M, Williams R, Meagher A, Hawkins NJ, Ward RL (2006) The role of MYH and microsatellite instability in the development of sporadic colorectal cancer. Br J Cancer 95(9):1239–1243
6. Kane MF, Loda M, Gaida GM et al (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 57(5):808–811
7. Herman JG, Umar A, Polyak K et al (1998) Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci USA 95(12): 6870–6875
8. Deng G, Bell I, Crawley S et al (2004) BRAF mutation is frequently present in sporadic colorectal cancer with methylated hMLH1, but not in hereditary nonpolyposis colorectal cancer. Clin Cancer Res 10(1 Pt 1):191–195
9. Domingo E, Laiho P, Ollikainen M et al (2004) BRAF screening as a low-cost effective strategy for simplifying HNPCC genetic testing. J Med Genet 41(9):664–668
10. Grabowski M, Mueller-Koch Y, Grasbon-Frodl E et al (2005) Deletions account for 17% of pathogenic germline alterations in MLH1, MSH2 in hereditary nonpolyposis colorectal cancer (HNPCC) families. Genet Test 9(2):138–146
11. Slupphaug G, Kavli B, Krokan HE (2003) The interacting pathways for prevention and repair of oxidative DNA damage. Mutat Res 531(1–2):231–251
12. Lefevre JH, Rodrigue CM, Mourra N et al (2006) Implication of MYH in colorectal polyposis. Ann Surg 244(6):874–880
13. Lipton L, Halford SE, Johnson V et al (2003) Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway. Cancer Res 63(22):7595–7599
14. Wolf B, Gruber S, Henglmueller S et al (2006) Efficiency of the revised Bethesda guidelines (2003) for the detection of mutations in mismatch repair genes in Austrian HNPCC patients. Int J Cancer 118(6):1465–1470
15. Pagenstecher C, Wehner M, Friedl W et al (2006) Aberrant splicing in MLH1 and MSH2 due to exonic and intronic variants. Hum Genet 119(1–2):9–22
16. Hemminki A, Peltomaki P, Mecklin JP et al (1994) Loss of the wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. Nat Genet 8(4):405–410
17. Tuupanen S, Karhu A, Jarvinen H, Mecklin JP, Launonen V, Aaltonen LA (2007) No evidence for dual role of loss of heterozygosity in hereditary non-polyposis colorectal cancer. Oncogene 26(17):2513–2517
18. Zhang J, Lindroos A, Ollila S et al (2006) Gene conversion is a frequent mechanism of inactivation of the wild-type allele in cancers from MLH1/MSH2 deletion carriers. Cancer Res 66(2): 659–664
19. Borrego S, Wright FA, Fernandez RM et al (2003) A founding locus within the RET proto-oncogene may account for a large proportion of apparently sporadic Hirschsprung disease and a subset of cases of sporadic medullary thyroid carcinoma. Am J Hum Genet 72(1):88–100
20. Gu Y, Parker A, Wilson TM, Bai H, Chang DY, Lu AL (2002) Human MutY homolog, a DNA glycosylase involved in base excision repair, physically and functionally interacts with mismatch repair proteins human MutS homolog 2/human MutS homolog 6. J Biol Chem 277(13):11135–11142
21. Colussi C, Parlanti E, Degan P et al (2002) The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. Curr Biol 12(11):912–918

Article C. APC, MYH and the correlation gentoype-phenotype in colorectal polyposis. (Ann Surg Oncol)

Ann Surg Oncol
DOI 10.1245/s10434-008-0297-0

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – GASTROINTESTINAL ONCOLOGY

APC, MYH, and the Correlation Genotype-Phenotype in Colorectal Polyposis

Jérémie H. Lefevre, MD[1], Yann Parc, MD, PhD[1], Magali Svrcek, MD[2], Solen Kernéis, MD[1], Chrystelle Colas, MD[3], Conor Shields, MD[1], Jean-François Flejou, MD[2], Rolland Parc, MD, FACS[1], and Emmanuel Tiret, MD[1]

[1]Department of Digestive Surgery, Hopital Saint-Antoine (AP-HP), University Paris VII Pierre & Marie Curie, Paris 75012, France; [2]Department of Pathology, Hopital Saint-Antoine (AP-HP), University Paris VII Pierre & Marie Curie, Paris 75012, France; [3]Laboratory of Angiogenetics and Oncogenetics, Hopital Pitié-Salpétrière (AP-HP), University Paris VII Pierre & Marie Curie, Paris 75013, France

ABSTRACT

Background. Familial adenomatous polyposis (FAP) has been divided into two entities: classical (CFAP) and attenuated (AFAP). With the discovery of MYH associated polyposis (MAP) syndrome, the clinical differences have become unclear. The aim of our study was to investigate patients with polyposis treated in our institution for a correlation between genotype and phenotype.

Methods. Between 1978 and 2007, 515 patients were followed. Four groups were identified: AFAP, CFAP, MAP, and no-mutation patients. Clinical, surgical, histological, and genetic data of patients were collected and compared. Two ranges of mutations responsible for AFAP were used.

Results. Patient breakdown was CFAP patients ($n = 322/294$), AFAP patients ($n = 13/41$), MYH patients ($n = 17$) and no-mut patients ($n = 32$). Patients not tested for APC mutation ($n = 131$) were excluded. Genotype/phenotype evaluation showed no difference in the number or location of polyps, age at colectomy, presence of cancer, or duodenal polyps. Major differences were found for MYH patients: later age at diagnosis, more cancers, fewer polyps, and more located in the right part of the colon. For phenotype/genotype correlation, patients aged more than 35 years at the time of colectomy and with fewer than 100 polyps had significantly more mutation found on MYH.

Conclusions. This two-way analysis did not show any correlation that might help to identify a subgroup of patients with APC mutation that may be considered attenuated. It is more likely that the MAP syndrome is the real AFAP.

© Society of Surgical Oncology 2009

First Received: 4 August 2008

J. H. Lefevre, MD
e-mail: jeremie.lefevre@sat.aphp.fr

Published online: 24 January 2009

Colorectal cancer (CRC) is the second most frequent cancer in France and the third most frequent cancer in Western countries.[1,2] Inherited factors are thought to play a major role in colorectal cancer as more than 25% of patients with CRC have a relevant family history. Two dominantly inherited syndromes predisposing to CRC have now been described: familial adenomatous polyposis (FAP) and hereditary non polyposis colorectal cancer (HNPCC).[3] The incidence of FAP varies from 1 in 8300 to 1 in 14,000 births.[4] The phenotype of this disease is characterized by the appearance of multiple colonic adenomas during the second decade of life and a life risk of CRC from 30% to 100% by the age of 40.[5] Extracolonic manifestations are also common and include potentially life-threatening manifestations such as duodenal, periampullary and ampullary adenomas, desmoid tumors, cerebral tumors, and other less serious manifestations that may aid in confirming the diagnosis, including congenital hypertrophy of the retinal pigmented epithelium (CHRPE), osteomas, and dermoid cysts.

Following the report in 1992 of a family with an apparently attenuated form of FAP, it has been customary to describe a disease spectrum, from classical to attenuated familial adenomatous polyposis (AFAP).[6] The main features of the AFAP syndrome are: 100 or less colorectal adenomas with very few adenomas of the rectum, later onset of adenomatosis, bowel symptoms and CCR, and a more limited expression of extracolonic manifestations.[6–9] However, a recent review of the literature reports three phenotypes of FAP: profuse, intermediate, and attenuated,[7]

and myriad terms have been used to describe subgroups of FAP (profuse, severe, mild, sparse, or attenuated), which demonstrates the difficulty of precisely defining these subgroups and their specific phenotypes.[7]

More than 825 germ-line mutations have been described in the APC mutation database,[10] and some specific sites of APC mutations have been reported to be linked to the AFAP phenotype.

The mutations located at the 5′ end of the APC gene, within exon 9 and at the 3′ distal end appear to be associated with the AFAP phenotype,[8,11–13] but the precise limits vary between authors: codons 175,[14] 169,[8,11] 167,[15] or 157[16] for the 3′ end and codons 1596,[14] 1581,[15] and 1403[8] for the 5′ end. Moreover, several studies have shown a great variability of mutations located in the AFAP regions of APC.[7] In addition to the uncertain phenotype-genotype correlation, several authors have questioned whether a better understanding of APC mutations would influence therapeutic decision making, again, with varying recommendations. Some advise prophylactic colectomy with ileorectal anastomosis instead of restorative proctocolectomy with ilealpouch anal anastomosis (IPAA) for patients with a mutation before codon 1250,[7,17,18] or for patients with a mutation in codon 0–200 or beyond 1500,[19] while others caution that decisions should be based solely on the clinical findings.[20]

To further complicate matters, a biallelic mutation of MYH, responsible for adenomatous polyposis with an increased risk of CRC, was described in 2002.[21] Subsequent studies have shown that while patients with biallelic mutations of MYH have numerous polyps, they develop neither thousands, nor the extracolonic features of FAP such as osteomas, duodenal polyps, or CHRPE.[3,22] Strikingly, this phenotype (MAP) resembles the AFAP phenotype. Thus, the clinical differences between FAP, AFAP, and MAP patients remain unclear.

The aim of our study was to investigate our population of colonic polyposis patients for evidence of a correlation between genotype and phenotype in adenomatous polyposis with a special focus on the AFAP syndrome.

PATIENTS AND METHODS

Between January 1978 and September 2007, 515 patients were followed for adenomatous polyposis in our institution. Genetic testing for an APC mutation was carried out in 384 patients after DNA extraction from peripheral blood and by sequencing the integrality of gene if a mutation was not found in the mutation cluster region (codon 1250–1550) of APC (GenBank ID: 324). For patients without an APC mutation, a search for the MYH (GenBank ID: 4595) mutation was performed as described previously.[23] Four groups were then identified: patients with an APC mutation localized in sites believed to be linked to AFAP (AFAP patients), patients with a classical APC mutation (CFAP patients), patients with the MYH biallelic mutation (MAP patients), and patients without a discovered mutation (no-mut). Definitions of the mutation sites responsible for AFAP vary between authors. For the purpose of this study we defined two groups of AFAP patients: one with a small range of mutations (before codon 163 or after codon 1596) (SAFAP) and one with a large range of mutations (before codon 175, in exon 9 or after codon 1403) (LAFAP). Patients with relatives treated in another institution and no available familial charts were considered as single patients.

Data of patients in the four groups (AFAP, CFAP, MAP, and no-mut patients) were collected and compared. Age, gender, detailed cancer family history, and surgical procedure were obtained from the patient's chart. Family history was also checked by direct interview. Histological features (number of colonic polyps, size, localization, presence of cancer) were all evaluated by the same pathologist who was blinded of the mutation identified. The genotype-phenotype correlation was also examined in the other direction by looking at the localization of the mutation on APC in the subgroup of patients displaying the most frequent characteristics of "attenuated polyposis": later age of surgery (35 years) and number of polyps (not greater than 100).

Statistical Analysis

Results are presented as median values (interquartile range) for continuous variables and numbers (percentages) for categorical variables. Chi-square or Fisher exact tests were used for categorical variables and *t* test or Wilcoxon test for continuous variables. All tests were two-sided at .05 significance level. Analyses were performed using the Statview computer program (Version 5, 1992–1998, SAS Institute Inc., Cary, NC) and R statistical package (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

The AFAP population was divided into two subgroups: SAFAP ($n = 13$ patients) and LAFAP ($n = 41$ patients). Therefore, depending on the definition of AFAP mutation employed, there were 322 or 294 CFAP patients. In 49 patients no APC mutation was identified, while a biallelic mutation of MYH was found in 17. Four groups of patients were then defined: CFAP patients ($n = 322$ and 294), AFAP patients ($n = 13$ and 41), MYH patients ($n = 17$),

and no-mut patients ($n = 32$). Patients not tested for APC mutation ($n = 131$) were excluded from the study. The patient distribution is shown in Fig. 1.

Mutation Analysis and Family History

The 13 SAFAP patients came from 2 families (6 patients) and 7 were single cases. Seven different mutations were identified in these 13 patients. Of the corresponding 322 CFAP patients, 180 were familial cases occurring in 65 families and 142 were single cases. The 41 LAFAP patients came from 8 families (20 patients) and 22 were single cases, and the corresponding 294 patients with CFAP mutations came from 59 families (166 patients) and 128 were single cases.

Among the 335 APC mutations found, there were 272 point mutations (83.4%), 36 deletions, 14 splicing mutations, and 4 mutations of the promoter. The two most frequent point mutations for CFAP patients were on codon 1309 ($n = 48$, 17.3%) and on codon 1062 ($n = 20$, 7.2%). The 17 MYH biallelic mutations found were observed in 15 families.

The ethical origin of patients was not available in the database, but the vast majority of this cohort was composed of white French patients.

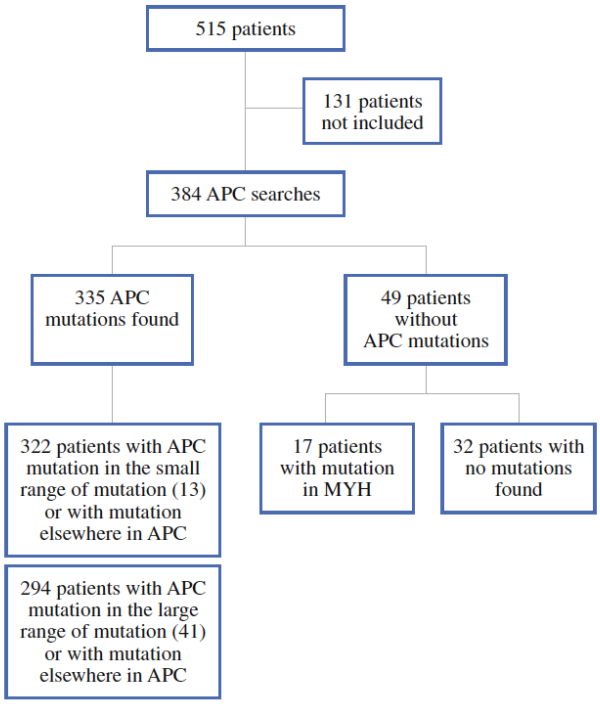


FIG. 1 Patients' distribution. *SAFAP* patient with APC mutation in the small range of mutation, *LAFAP* patient with APC mutation in the large range of mutation, *CFAP* patient with mutation elsewhere in APC, *MYH* patient with mutation in MYH, *No-Mut* patient without any mutation found

Clinical Features

Diagnosis of polyposis, number and localization of polyps, presence of cancer, desmoid tumors, duodenal polyps, gastric fundic polyps, and surgical procedures for each patient group are detailed in Table 1. Statistical analysis for CFAP and AFAP groups of each parameter is shown in Table 2.

Number and Localization of Polyps

The exact number of polyps was known for 177 patients (46.1% of our cohort). For the remaining patients, a precise number was not determined. These patients were not included in the statistical analysis. According to the definition of AFAP syndrome, the number of 100 polyps was used for comparison. No difference was observed between patients with CFAP and SAFAP or LAFAP. The localization of polyps was noted in the histological report of 158 patients (41.1% of the cohort). Again, no difference in localization was observed between the patients with CFAP and SAFAP or LAFAP.

Time of Colectomy

Age at the time of the first colectomy did not differ between AFAP and CFAP patients.

Presence of Upper Gastrointestinal Lesions

Fundic glandular dystrophy was observed more frequently in patients with a mutation in LAFAP regions than in those with a CFAP mutation (65.7% vs. 30.3%, $P < .0001$).

Desmoid Tumors

In the group of LAFAP patients, desmoid tumors were observed more frequently than in the group of CFAP patients (39% vs. 14.3%, $P < .0001$).

Colorectal Cancer

The rate of colorectal cancer observed in the LAFAP or SAFAP groups did not differ from the corresponding CFAP patients.

Surgical Procedure

The rate of colectomy with ileorectal anastomosis was equivalent in the four patient groups. However, a two-stage procedure with conversion of an ileorectal anastomosis to an IPAA was carried out more frequently in CFAP patients rather than in those with LAFAP or SAFAP.

TABLE 1 Characteristics of our cohort of patients regarding the mutation

	General population	Small range		Large range		MYH mutation	No mutation found
		FAP mutation	AFAP mutation	FAP mutation	AFAP mutation		
Number of patients, *n*	384	322	13	294	41	17	32
Gender (male), *n* (%)	177 (49)	159 (49.3)	5 (38.5)	147 (50)	17 (41.4)	12 (70.6)	15 (46.9)
Age at time of first colonic resection, yr	29.4	(± 11.6)	27.9 (± 10.7)	32.5 (± 8.7)	27.7 (± 10.9)	31.0 (± 8.7)	48.0 (± 8.6)
34.0 (± 12.0)							
Number of polyps >100 *n* (%)[a]		90[a] (68.7)	8 (61.5)	79[b] (68.7)	19[c] (65.5)	6 (37.5)[d]	7[e] (41.2)
Location of polyps, right/left/all colon, *N* (%)		17/54/50 (14.0/44.6/41.3)	1/7/1 (11.1/88.8/11.1)	15/51/42 (15.9/47.2/38.9)	3/10/9 (13.6/45.4/40.9)	7/0/3 (70/0/30)	4/5/7 (25/31.2/43.7)
CRC, *n* (%)	87 (22.6)	67 (21.1)	2 (15)	57 (19.4)	12 (29.3)	13 (76.5)	7 (21.9)
Desmoid tumor, *n* (%)	58 (15.1)	55 (17.5)	3 (23.1)	42 (14.3)	16 (39.0)	0 (0)	0 (0)
Gastric glandular polyps, *n* (%)	108 (28.1)	84 (34.7)[f]	6 (46.1)	67 (30.3)[g]	23 (65.7)[h]	2 (14.3)[i]	6 (19.3)[j]
Duodenal polyps, *n* (%)	145 (37.8)	129 (53.5)[k]	4 (31)	115 (52.2)[l]	18 (52.9)[m]	3 (21.4)[i]	9 (29.0)[j]
IRA, *n* (%)	96 (25.1)[n]	76 (23.7)[n]	5 (38.5)	71 (24.3)[n]	10 (24.4)	7 (41.1)	8 (25)
IRA → IPAA, *n* (%)	58 (60.4)	54 (71)	1 (20)	54 (76.0)	1 (10)	1 (14.3)	1 (12.5)

IRA ileorectal anastomosis, *IPAA* ileopouch anal anastomosis

[a] Known for 131 patients; [b] Known for 115 patients; [c] Known for 29 patients; [d] Known for 16 patients; [e] Known for 17 patients; [f] Known for 243 patients; [g] Known for 221 patients; [h] Known for 35 patients; [i] Known for 14 patients; [j] Known for 31 patients; [k] Known for 241 patients; [l] Known for 220 patients; [m] Known for 34 patients; [n] Two patients had a non colonic operation

MYH Patients

The 17 patients with a biallelic mutation on MYH exhibited several significant differences to those with mutations on APC: older age at colectomy (48.0 ± 8.6 vs. 28.2 ± 10.7, $P < .0001$), fewer polyps ($P = .0151$) more frequently located in the right colon ($P < .0001$), no desmoid tumors ($P < .0001$), more colorectal cancer ($P < .0001$), and fewer duodenal polyps ($P = .0148$).

Correlation Phenotype-Genotype

Among the 177 patients with a precise count of polyps, 31 underwent an operation for polyposis after the age of 35. Among those patients with a phenotype of attenuated polyposis, only 5 exhibited a mutation in the attenuated regions of APC. Ten patients had a biallelic mutation on MYH. The breakdown of the patients with regard to phenotype and mutations is described in Fig. 2. Finally, only 3

TABLE 2 Comparison between patients with FAP mutation or AFAP mutation with the two ranges

	FAP mutation vs. AFAP mutation	
	Small range	Large range
Age at the time of the first colectomy	$P = .1339$	$P = .066$
Desmoid tumor	$P = .6029$	$\mathbf{P < .0001}$
Number of polyps (<100/>100)	$P = .5972$	$P = .7429$
Localization of polyps (right/left/all colon)	$P = .1379$	$P = .9841$
Cancer	$P = .6209$	$P = .1561$
Gastric dysplasia	$P = .3940$	$\mathbf{P < .0001}$
Duodenal polyps	$P = .1095$	$P = .9421$
IRA	$P = .2255$	$P = .9916$
IRA → IPAA	$\mathbf{P = .0179}$	$\mathbf{P < .0001}$

IRA ileorectal anastomosis, *IPAA* ileoanal pouch anastomosis

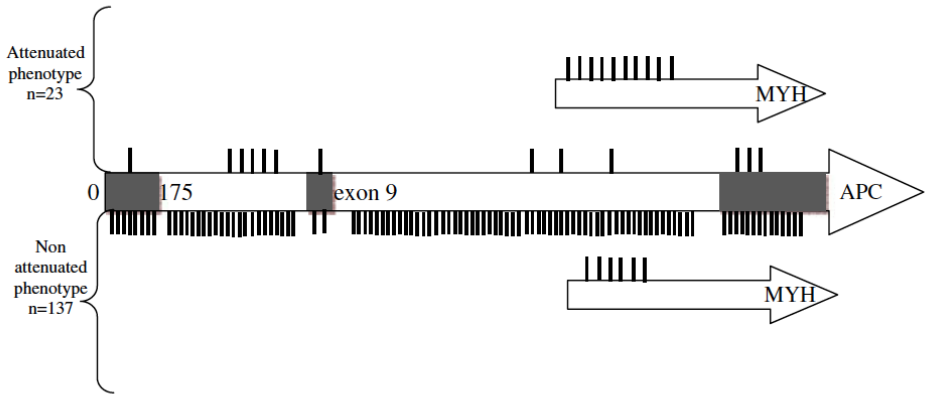


FIG. 2 Repartition of the mutation site depending on the phenotype observed

of the 13 SAFAP patients and only 4 of the 29 LAFAP patients with a known polyp count had an attenuated phenotype. Among the 17 MYH patients, 10 of them had such a phenotype.

DISCUSSION

Adenomatous polyposis is caused by a mutation on the APC gene and was the first syndrome predisposing to colorectal cancer to be described.[24] Even before the identification of the APC gene in 1991,[25] this syndrome was known to be autosomic and dominantly inherited, and its phenotype was well defined, allowing the development of efficient familial surveillance and prophylactic surgery protocols. In 1992, FAP was divided into two entities: classical and attenuated familial adenomatous polyposis,[6] while in 2002, a biallelic mutation of MYH responsible for a recessively inherited adenomatous polyposis syndrome (MAP syndrome) was described. Currently the phenotype/genotype correlation of FAP, AFAP, and MAP is unclear and unhelpful in the clinical management of these patients.

In the literature, most articles about the AFAP syndrome include very few patients. In studies with polyposis patients, the number of patients with a mutation in the AFAP region is low.[18,20] Other articles describe several families.[6,14,26] In our population of polyposis patients from a single center, even with the largest mutation definition, we were able to find only 41 patients (10.7%) with a suspected AFAP mutation. This low rate makes it difficult to reach a precise definition of this syndrome and to establish therapeutic guidelines.

The classical clinical items usually used to define the AFAP syndrome: later age at diagnosis, lower number of polyps (fewer than 100), more frequent localization in the right colon, fewer extracolonic manifestations were compared between CFAP and SAFAP or LAFAP. None of the characteristics was found to be significantly different between these groups. Some differences were, however, observed. Desmoid tumors were more frequently found in LAFAP patients—an observation that renders the term "attenuated" less appropriate. This fact is explained by the high rate of desmoid tumors in patients with a germ-line mutation 3′ to codon 1444.[20,26,27] For patients with supposed AFAP, conservation of the rectum is recommended as rectal polyps are thought to occur less frequently.[7,17,18] In our patient cohort, the rate of rectal conservation and the rate of IPAA were the same in the CFAP group and the LAFAP or SAFAP groups. However, the rate of patients requiring subsequent conversion to IPAA was significantly less than observed in CFAP patients. The presence of desmoid tumors may preclude such conversion. Bearing these two findings in mind, the decision to preserve the rectum in supposed AFAP patients might become more difficult. However, we need to emphasize that the cohort comprised only 10 patients. The limited number of patients makes a meaningful interpretation of the statistics rather difficult. Our findings would suggest that the decision to perform an IRA or IPAA should still be based on the endoscopic findings. The subsequent need for conversion from an IRA may be low if the mutation occurs in the AFAP region of the APC gene.

With regard to the correlation between the phenotype and the location of the mutation on APC, we found that only 5 of the 31 patients with fewer than 100 polyps and an age at surgery greater than 30 years had a mutation on the attenuated regions of APC.

This two-way analysis did not show any correlation that might help to identify a subgroup of patients with the APC mutation that should be treated differently. In contrast with some authors who have proposed the use of mutation localization as a tool for decision making in the management of FAP patients,[7,17–19] we recommend that these decisions be based purely on the clinical features of each patient. Case reports such as the one by Matsuo et al.[28] describing an advanced rectal cancer in a 16-year-old boy from an attenuated polyposis family, indirectly support our stance. Moreover, the variability of expression of APC mutation has already been demonstrated.[29] In a series of 74 patients with the same 5-base pair deletion at codon 1309, a wide variation for age at diagnosis (19–62 years) and rate

of extracolonic manifestations (0–100%) was reported. In the same study, 6 patients had diffuse polyposis while 6 others had only occasional polyps in the right colon. At present, explanations for this high variability in the expression of this syndrome are lacking. The existence of APC modifier genes, the location of the second hit on the wild allele, or even the "three hits" theory (meaning two somatic mutations, including loss or mutation of the germ-line mutant allele) are possible explanations.[30]

Eventually, when confronted with an "attenuated familial adenomatous polyposis" phenotype, MYH mutation should be considered before APC. In fact, the MAP syndrome including the presence of several polyps in the colon but never thousands, later age of diagnosis, upper gastrointestinal manifestations (duodenal polyps and/or fundic gastric hyperplasia), and lack of familial history of polyposis, may be the "true" AFAP syndrome, especially considering that APC mutations are only found in 5–15% of patients with attenuated polyposis.[31,32] One study even reported no APC mutation detected in any of their 6 patients with a suspected AFAP syndrome, indicating that other genetic defects may be the cause of the attenuated phenotype.[33] In our study, patients with MYH mutation were significantly older at the time of surgery and had a lower number of polyps that were more frequently located in the right colon. Moreover, when comparing the phenotype with the localization of the mutation, 10 of the 31 patients with an attenuated phenotype were found to have a biallelic mutation on MYH.

Despite the absence of correlation between the genotype and the phenotype in patients with an APC mutation, the search for the APC mutation remains essential as it may mandate a search for at-risk relatives. If no mutation is found on APC, a search for the MYH mutation should be undertaken. If no MYH mutation is identified, all relatives should be considered to be APC mutant even if the patient has an AFAP phenotype. In conclusion, we consider that adenomatous polyposis should now be separated in two entities: the FAP syndrome with an intrinsic large variability of expression and the MAP syndrome.

ACKNOWLEDGMENT There are no potential conflicts of interests.

REFERENCES

1. Bouvier AM, Remontet L, Jougla E, Launoy G, Grosclaude P, Buémi A, et al. Incidence of gastrointestinal cancers in France. *Gastroenterol Clin Biol*. 2004;28:877–81.
2. Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics. *CA Cancer J Clin*. 1999;49:8–31.
3. Grady WM. Genetic testing for high-risk colon cancer patient. *Gastroenterology*. 2003;124:1574–94.
4. Benamouzig R, Chaussade S, Olschwang S. Polyposis coli: a practical approach. *Gastroenterol Clin Biol*. 2005;29:1035–9.
5. Penna CP, Tiret E, Kartheuser AH, Olschwang S, Parc R. Value of screening of familial adenomatous polyposis for prevention of colorectal cancer. *Gastroenterol Clin Biol*. 1992;16:210–4.
6. Lynch HT, Smyrk T, McGinn T, Lanspa S, Cavalieri J, Lynch J, et al. Attenuated familial adenomatous polyposis (AFAP). A phenotypically and genotypically distinctive variant of FAP. *Cancer*. 1995;76:2427–33.
7. Nieuwenhuis MH, Vasen HFA. Correlations between mutation site in APC and phenotype of familial adenomatous polyposes (FAP): A review of the literature. *Crit Rev Oncol/Hematol*. 2007; 61:153–61.
8. Scott RJ, Meldrum C, Crooks R, Spigelman AD, Kirk J, Tucker K, et al. Familial adenomatous polyposis: more evidence for disease diversity and genetic heterogeneity. *Gut*. 2001;48:508–14.
9. Knudsen AL, Bisgaard ML, Bulow S. Attenuated familial adenomatous polyposis (AFAP). A review of the literature. *Fam Cancer*. 2003;2:43–55.
10. Beroud C, Soussi T. APC gene: database of germline and somatic mutation in human tumors and cell lines. *Nucl Acids Res*. 1996;24:121–4.
11. Spirio L, Olschwang S, Groden J, Robertson M, Samowitz W, Joslyn G, et al. Alleles of the APC gene: an attenuated form of familial polyposis. *Cell*. 1993;75:951–7.
12. Gebert JF, Dupon C, Kadmon M, Hahn M, Herfarth C, von Knebel Doeberitz M, et al. Combined molecular and clinical approaches for the identification of families with familial adenomatous polyposis coli. *Ann Surg*. 1999;229:350–61.
13. Van der Luitj RB, Vassen HF, Tops CM, Breukel C, Fodde R, Meera Khan P. APC mutation in the alternatively spliced region of exon 9 associated with late onset of adenomatous polyposis. *Hum Genet*. 1995;96:705–10.
14. Burt RW, Leppert MF, Slattery ML, Samowitz WS, Spirio LN, Kerber RA, et al. Genetic testing and phenotype in a large kindred with attenuated familial adenomatous polyposis. *Gastroenterology*. 2004;127:444–51.
15. Lipton L, Tomlinson I. The genetics of FAP and FAP-like syndromes. *Fam Cancer*. 2006;5:221–6.
16. Moisio AL, Järvinen H, Peltomäki P. Genetic and clinical characterisation of familial adenomatous polyposis: a population based study. *Gut*. 2002;50:845–50.
17. Vasen HF, van der Luijt RB, Slors JF, Buskens E, de Ruiter P, Baeten CG, et al. Molecular genetic tests as a guide to surgical management of familial adenomatous polyposes. *Lancet*. 1996; 348:433–5.
18. Wu JS, Paul P, McGannon EA, Church JM. APC genotype, polyp number, and surgical options in familial adenomatous polyposis. *Ann Surg*. 1998;227:57–62.
19. Bülow C, Vasen H, Järvinen H, Björk J, Bisgaard ML, Bülow S. Ileorectal anastomosis is appropriate for a subset of patients with familial adenomatous polyposes. *Gastroenterology*. 2000;119: 1454–60.
20. Friedl W, Caspari R, Sengteller M, Uhlhaas S, Lamberti C, Jungck M, et al. Can APC mutation analysis contribute to therapeutic decisions in familial adenomatous polyposes? Experience from 680 FAP families. *Gut*. 2001;48:515–21.
21. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, et al. Inherited variants of MYH associated with somatic G:C → T:A mutations in colorectal tumors. *Nat Genet*. 2002;30:227–32.
22. Sieber OM, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips RK, et al. Multiple colorectal adenomas, classic adenomatous polyposis and germ-line mutations in MYH. *N Engl J Med*. 2003;348:791–9.

23. Lefevre JH, Rodrigue CM, Mourra N, Bennis M, Flejou JF, Parc R, et al. Implication of MYH in colorectal polyposis. *Ann Surg*. 2006;244:874–80.
24. Bussey HJR. *Familial Polyposis Coli*. Baltimore: John Hopkins Univeristy Press 1975.
25. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol*. 2001; 96:2292–3003.
26. Caspari R, Olschwang S, Friedl W, Mandl M, Boisson C, Böker T, et al. Familial adenomatous polyposis: desmoid tumours and lack of ophthalmic lesions (CHRPE) associated with APC mutations beyond codon 1444. *Hum Mol Genet*. 1995;4:337–40.
27. Houlston R, Crabtree M, Phillips R, Crabtree M, Tomlinson I. Explaining differences in the severity of familial adenomatous polyposis and the search for modifier genes. *Gut*. 2001;48:1–5.
28. Matsuo S, Eguchi S, Azuma T, Hidaka M, Yamaguchi S, Hayashi T, et al. Attenuated familial adenomatous polyposis associated with advanced rectal cancer in a 16-year-old boy: report of a case. *Surg Today*. 2001;31:1020–3.
29. Giardiello FM, Krush AJ, Petersen GM, Booker SV, Kerr M, Tong LL, et al. Phenotypic variability of familial adenomatous polyposis in 11 unrelated families with identical APC gene mutation. *Gastroenterology*. 1994;106:1542–7.
30. Sieber OM, Segditsas S, Knudsen AL, Zhang J, Luz J, Rowan AJ, et al. Disease severity and genetic pathways in attenuated familial polyposis vary greatly but depend on the site of the germline mutation. *Gut*. 2006;55:1440–8.
31. Lamlum H, Al Tassan N, Jaeger E, Frayling I, Sieber O, Reza FB, et al. Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q. *Hum Mol Genet*. 2000;9:2215–21.
32. Sieber OM, Lamlum H, Crabtree MD, Rowan AJ, Barclay E, Lipton L, et al. Whole gene APC deletions cause classical familial adenomatous polyposis, but not attenuated polyposis or multiple colorectal adenomas. *Proc Natl Acad Sci USA*. 2000; 99:2954–8.
33. Nilbert M, Fernebro J, Kristofferson U. Novel germline APC mutations in Swedish patients with familial adenomatous polyposis and Gardner syndrome. *Scand J Gastroenterol*. 2000; 35:1200–3.

Article D. Frequent mutation in North-African patients with MUTYH-Associated polyposis. (Clin Genet).

An International Journal of Genetics, Molecular, and Personalized Medicine

Clin Genet 2010
Printed in Singapore. All rights reserved

© 2010 John Wiley & Sons A/S

CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2010.01528.x

Short Report

Frequent mutation in North African patients with *MUTYH*-associated polyposis

Lefevre JH, Colas C, Coulet F, Baert-Desurmont S, Mongin C, Tiret E, Frebourg T, Soubrier F, Parc Y. Frequent mutation in North African patients with *MUTYH*-associated polyposis.
Clin Genet 2010. © John Wiley & Sons A/S, 2010

MUTYH-associated polyposis (MAP) has been characterized as an autosomal recessive disease predisposing to a variable number of colorectal adenomas with a high risk of cancer. Numerous studies have indicated that two missense mutations (Y179C and G396D) account for about 80% of *MUTYH* allelic variants in Europeans. Ethnic and geographic differences in the mutation spectrum have been observed. The aim of this study was to report mutations in patients from North Africa, determine the incidence of the c.1227_1228dup mutation in our cohort of *MUTYH* patients and to evaluate the existence of a founder effect. Within a group of 36 families with MAP, 11 were shown to have a homozygous c.1227_1228dup mutation. These families came from Algeria ($n = 5$), Tunisia ($n = 4$), Morocco ($n = 1$) and Portugal ($n = 1$). Probands belonging to families of North African origin showed a significantly higher frequency of c.1227_1228dup (78.6% *vs* 4.5%, $p < 0.0001$). Haplotype analyses were performed using 10 microsatellite markers surrounding the *MUTYH* gene spanning a region of 4.4 cM. We identified a common haplotype of at least 1.3 cM in all families suggesting a founder effect for this mutation.

Conflict of interest

Nothing to declare.

JH Lefevre[a,b], C Colas[c], F Coulet[c], S Baert-Desurmont[d,e], C Mongin[a], E Tiret[a], T Frebourg[d,e], F Soubrier[c] and Y Parc[a]

[a]Department of Digestive Surgery, Hopital Saint-Antoine AP-HP, University Paris VI (Pierre and Marie Curie), Paris, France, [b]Cancer and Immunogenetics Laboratory, Department of Clinical Pharmacology, University of Oxford, Oxford, UK, [c]Laboratory of Oncogenetics and Angiogenetics, Groupe Hospitalier Pitié-Salpêtrière, University Paris VI (Pierre and Marie Curie), Paris, France, [d]Department of Genetics, University Hospital, Rouen, France, and [e]INSERM U614, Faculty of Medicine, Rouen, France

Key words: adenomatous polyposis – *APC* – colorectal cancer – founder effect – *MUTYH* – *MUTYH*-associated polyposis

Corresponding author: Prof Yann Parc, Department of Digestive Surgery, Hôpital Saint-Antoine AP-HP, University Pierre et Marie Curie Paris VI, 184 rue du Faubourg Saint-Antoine, F-75571 Paris, France.
Tel.: +33 1 49 28 25 30;
Fax: +33 1 49 28 25 48;
e-mail: yann.parc@sat.aphp.fr

Received 4 June 2010, revised and accepted for publication 8 August 2010

MUTYH-associated polyposis (MAP) is a recently identified syndrome responsible for adenomatous polyposis, where patients do not show mutations on *APC* (1–6). *MUTYH* mutations are found at a frequency ranging from 4% to 30% in these patients (7). The identification of the germline mutation is essential because MAP is transmitted as an autosomal recessive disease in contrast to familial adenomatous polyposis (FAP) which is autosomal dominant. Thus, the genetic screening of *MUTYH* is mandatory for patients with polyposis of unrecognized genetic origin.

The two most frequent mutations found in biallelic mutated patients are Y179C and G396D (8–10). These two mutations have been reported to occur at a frequency of approximately 1% in Europeans (8) and are present in up to 70% of MAP patients (10). However, ethnic and geographic differences in the mutation spectrum have been observed: E480X is frequent in Indian patients (5)

and Y179C or G396D have not been found among Japanese patients (11). Patients from North Africa have not been examined yet with respect to these mutations; however, in a previous study on *MUTYH*, we described a family from Algeria with the rare c.1227_1228dup (p.Glu396GlyfsX43) mutation (2). This mutation had been first described by Isidro et al. in Portuguese patients and was found in a patient from the island of Madeira (12). In that study, the c.1227_1228dup mutation was not detected in 100 unrelated healthy individuals (selected at random and not pooled: 50 from Madeira and 50 from mainland Portugal) (12). Moreover, in a large multicentric study of 114 patients from USA, Australia and Canada with monoallelic or biallelic mutations on *MUTYH*, this mutation was not detected (13).

In this study, we identify the recurrent mutation c.1227_1228dup in patients from North Africa, report its incidence and assess the likelihood of this being a founder mutation.

Materials and methods

All patients with adenomatous polyposis or colorectal cancer (CRC) diagnosed before the age of 50 had a consultation with a geneticist (CC). After obtention of an informed consent, patients were screened either for *MUTYH* mutation or for *APC* mutation first, depending on their age, phenotype and family history. All patients with a proven germinal mutation are included in a prospective polyposis database. In this study, we focused on families from North African origin with a biallelic mutation of *MUTYH* found in the proband.

DNA extraction

Lymphocyte DNA was extracted from ethylenediaminetetraacetic acid (EDTA) venous blood samples. Human genomic DNA was prepared from whole blood samples using a semi-automated Extragene® extractor (Genomic Industry, Archamps, France) and a DNA extraction kit (Promega France, Charbonnières-les-Bains, France) according to the manufacturers' standard protocol.

PCR amplification and sequencing of *MUTYH* and *APC*

All *APC* and *MUTYH* exons were amplified by polymerase chain reaction (PCR) in a total volume of 30 µl of reaction mixture containing 100 ng of genomic DNA of each sample, 1X PCR buffer, 2.5 mM $Mgcl_2$, 10 mM dNTP, 3.3 µM of each primers as previously described (14) (Table S1, Supporting Information) and 1U Ampli Taq Gold® DNA polymerase (Applied Biosystems, Foster City, CA). Amplification was performed on a GeneAmp PCR system 9700 (Applied Biosystems). Before sequencing, the positive reactions were cleaned up on a MultiScreen™ PCR 96 well plate (Millipore, Billerica, MA). Sequencing reactions were carried out in forward and reverse orientations using the BigDye Terminator v1.1 cycle sequencing kit (Applied Biosystems). The products of the sequencing reactions are cleaned up using the Sephadex™ G-50 (GE Healthcare, Life Sciences) in a MultiScreen®-HV 96-well filter plate (Millipore, Billerica, MA), then run up on a ABI 3730 DNA sequencer. The resulting sequence data were analyzed with the SEQSCAPE software, version 2.5 (Applied Biosystems) in comparison with the references sequences of human *APC*, *MUTYH* gene (RefSeq: NM 001127511.1 and NM_001128425.1). The description of sequence variants in the *MUTYH* gene follows current human genome variation society (HGVS) recommendations using the longest *MUTYH* transcript, transcript alpha 5, NM_001128425.1, and not the historical transcript alpha 3, NM_001048171.1, with exon 3 alternatively spliced.

Haplotype analysis

Haplotype analyses were performed for all mutation carriers, using 10 microsatellite markers surrounding the *MUTYH* gene: from telomere to centromere D1S447, D1S211, D1S2733, D1S2713, D1S2802, D1S451, D1S2677, *MUTYH*, D1S322, D1S2797, D1S2720 (Fig. 1). Each marker was amplified using fluorescent primers. PCR products were electrophoresed in an ABIPrism 3730 Analyzer (Applied Biosystems) and analyzed with GENEMAPPER version 4.1 software (Applied Biosystems). Amplification primers for microsatellite markers D1S447, D1S211, D1S2733, D1S27 13, D1S2802, D1S451, D1S2677, D1S322, D1S2 797 and D1S2720 are listed in the University of California Santa Cruz genome database (genome.ucsc.edu) (Table S2).

Statistical analysis

Results are presented as mean ± SD (interquartile range) for continuous variables and counts (percentages) for categorical variables. Chi-squared or Fisher's exact tests were used for categorical variables and Student's *t*-test for continuous variables. All tests were two-sided at 0.05 significance level. Analyses were performed using Statview® (SAS Institute Inc., Cary, North Carolina, USA).

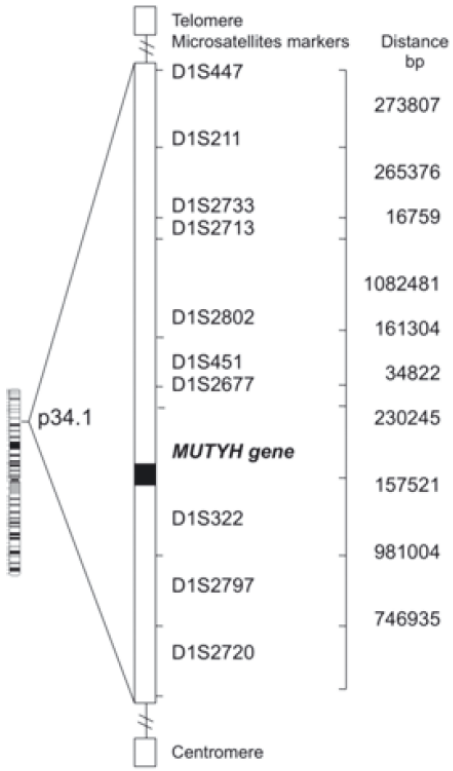


Fig. 1. Microsatellite genotypes around *MUTYH*.

Results

Patients and families

Thirty-six different families from the polyposis database were included in this study after a biallelic *MUTYH* mutation was found in the proband. Among them, 13 families were originally from North Africa [Algeria ($n = 7$), Tunisia ($n = 5$) and Morocco ($n = 1$)) and the remaining 23 were of Caucasian European origin (France ($n = 11$), Portugal ($n = 5$), Belgium ($n = 1$), Italy ($n = 1$), and Turkey ($n = 1$)) or had mixed origins (French/Algeria ($n = 1$), Germany/France ($n = 1$), Belgium/France ($n = 1$), and Armenia/France ($n = 1$)].

Distribution of *MUTYH* mutations in different ethnic groups

Probands for 11 families had a homozygous c.1227_1228dup (p.Glu396GlyfsX43) mutation. These 11 families were originally from Algeria ($n = 5$), Tunisia ($n = 4$), Portugal ($n = 1$) and Morocco ($n = 1$). Another proband was a compound heterozygote (c.1227_1228dup /Y179C) and was of Algerian and French ancestry. Among the 14 families of North African origin (i.e. Algeria, Tunisia, Morocco or mixed origin), 11 probands (78.6%) had a homozygous or heterozygous c.1227_1228dup mutation. By contrast, only one among the 22 families of non-North African origin exhibited this mutation (4.5%) and he was of Portuguese origin. The difference was statistically significant ($p < 0.0001$, Fisher's exact test). Subsequently, the biallelic homozygous c.1227_1228dup mutation was found in four additional subjects from these 11 families. Three of them had polyposis and one had no polyps at 24 years old. All these 15 patients were studied for genotype–phenotype correlations. The control group included 27 patients belonging to 25 families with other biallelic mutations on *MUTYH*. The majority of these mutations were Y169C ($n = 10$ patients, 37%) and G396D ($n = 13$ patients, 48%) in a heterozygous or homozygous state.

Genotype-phenotype correlation

The population's characteristics are detailed in Table 1. No difference in the phenotypic expression of the polyposis was observed in the *MUTYH* biallelic mutated patients between patients with

Table 1. Characteristics of the cohort of 42 patients with a biallelic mutation on *MUTYH*[a]

	c.1227_1228dup ($n = 15$)	Other biallelic mutations on *MUTYH* ($n = 27$)	p
Male	9/15 (60%)	28/27 (66.7%)	0.187
Consanguinity			
No	2/13 (15.4)	16/27 (59.3)	0.0006
Yes	10/13 (76.9)	4/27 (14.8)	
Possible	1/13 (7.7)	7/27 (25.9)	
More than 30 polyps	11/15 (73.3)	23/26 (88.5)	0.215
Colorectal cancer	12/15 (80)	19/27 (70.4)	0.496
Duodenal polyps	3/9 (33.3)	4/17 (23.5)	0.592
Familial history of polyposis	9/14 (64.3)	14/27 (51.9)	0.447
Age at diagnosis	44.3 ± 8.9 [30–62]	44.2 ± 8.8 [24–60]	0.969
Pseudoinheritance transmission	2/11 (18%)	1/25 (4%)	0.1925
Number of polyps	75.5 ± 80.2 [7–266]	75.2 ± 40.0 [15–156]	0.855

[a]Continuous variables are given as mean ± SD [interquartile range].

a c.1227_1228dup homozygous mutation and the patients with other *MUTYH* mutations (Table 1). However, patients with a c.1227_1228dup homozygous mutation were more probably to have a history of consanguinity in their medical records.

Haplotype analysis

Ten microsatellites were analyzed in one member of each of the 11 families with a *MUTYH* c.1227_1228dup biallelic mutation. *MUTYH*-associated haplotypes in families carrying this mutation are shown in Table 2. Every family had a similar haplotype between markers D1S2802 and D1S2797 surrounding the *MUTYH* gene. The common haplotype had a length of at least 1.3 cM (Fig. 1).

Discussion

MUTYH-associated polyposis is a syndrome which accounts for almost one third of patients with adenomatous polyposis and no *APC* mutations (2). A biallelic mutation is also found in 0.4% to 1% of population-based CRC (13). The first case was described by Al-Tassan et al. who discovered two mutations in *MUTYH*, Y179C and G396D in a Welsh family with CRC and multiple adenomas (1). Currently, an overview of the most frequent mutations in *MUTYH* can be found at http://www.LOVD.nl/MUTYH and http://www.insight-group.org (15). The increasing number of publications on *MUTYH* has unraveled ethnic and geographic differences in the mutation frequency of this gene. Among European populations, the two missense mutations Y169C and G396D account for almost 90% of all *MUTYH* mutations in individuals of Northern European ancestry (13). Some studies have even limited the screening to just these two common mutations (16–18). On the other hand, these mutations have not been identified in Asian patients with adenomatous polyposis (15, 19). Typical mutations found in the Japanese population include R245C and c.934-2A>G (11). Moreover, E480X is the only mutation identified among Indian patients to date (3, 5, 20).

Patients from North Africa have not been examined before, with the exception of a previous study on *MUTYH*, where we described an Algerian family with the rare c.1227_1228dup mutation (2). This homozygous mutation had already been reported in three Portuguese patients (12) and in a Dutch patient of Moroccan origin (21). In this study, we report the largest group of patients showing a biallelic c.1227_1228dup mutation. The fact that all patients in these previous studies as well as in our study were homozygous is

Table 2. Haplotype of the 10 families with a 1227_1228dup mutation found on *MUTYH*

		Theoretical size	Family 1	Family 2	Family 3	Family 4	Family 5	Family 6	Family 7	Family 8	Family 9	Family 10	Family 11
MS1	D1S447	123–141	126/126	128/130	124/124	130/130	130/130	130/130	123/130	130/130	123/130	130/130	130/130
MS2	D1S211	170–198	185/185	166/181	183/183	166/166	166/166	166/166	166/181	166/166	166/181	166/166	166/166
MS3	D1S2733	107–121	113/113	105/111	105/105	111/111	111/111	111/111	111/111	111/111	105/111	105/111	111/111
MS4	D1S2713	227–279	262/262	252/264	260/264	264/264	264/264	264/264	266/274	264/268	264/264	264/264	264/264
MS5	D1S2802	192–202	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193	193/193
MS6	D1S451	174–188	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175	175/175
MS7	D1S2677	135–153	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139	139/139
MS8	D1S322	90	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106	106/106
MS9	D1S2797	144–180	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162	162/162
MS10	D1S2720	235–245	234/234	232/238	232/232	234/234	232/234	236/236	234/236	234/234	238/242	237/237	238/238

probably secondary to the high rate of consanguinity in North African countries. Many studies have shown that consanguineous marriages were deeply rooted in Arab and Muslim populations (22–24). This characteristic was the only significant difference between patients with c.1227_1228dup and patients with other biallelic mutations on *MUTYH*. The frequent homozygosity is also due to the frequency of the mutation, which remains to be precisely determined in North African populations.

Similar genotypes for five microsatellites around *MUTYH* were observed in 11 unrelated families therefore showing the existence of a common haplotype extending for approximately 1.3 cM and the likelihood that a founder effect has occurred. The origin of this founder mutation is not known, and large populations from different parts of Africa should be tested for this mutation in order to map and date the event.

This observation underscores the importance of identifying the country of origin of a particular mutation and understanding population migratory movements. It may also have major public health consequences for genetic counseling in population from North Africa as for French population, with regards to the mutation frequency, which remains to be estimated, and the high rate of consanguineous unions. Finally, from a clinical point of view: patients of North African origin with an attenuated polyposis should have *MUTYH* exon 13 sequenced as the first step of their genetic diagnosis.

Supporting Information

The following Supporting information is available for this article:

Table S1. *MUTYH* primers for PCR.

Table S2. Microsatellite primers for haplotype analysis.

Additional Supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgement

The authors would like to particularly thank Carolina Bonilla for editing.

References

1. Al-Tassan N, Chmiel NH, Maynard J et al. Inherited variants of MYH associated with somatic G:C–>T:A mutations in colorectal tumors. Nat Genet 2002: 30: 227–232.
2. Lefevre JH, Rodrigue CM, Mourra N et al. Implication of MYH in colorectal polyposis. Ann Surg 2006: 244: 874–879; discussion: 879–880.
3. Jones S, Emmerson P, Maynard J et al. Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C–>T:A mutations. Hum Mol Genet 2002: 11: 2961–2967.
4. Lipton L, Halford SE, Johnson V et al. Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway. Cancer Res 2003: 63: 7595–7599.
5. Sampson JR, Dolwani S, Jones S et al. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. Lancet 2003: 362: 39–41.
6. Sampson JR, Jones S, Dolwani S et al. MutYH (MYH) and colorectal cancer. Biochem Soc Trans 2005: 33: 679–683.
7. Gismondi V, Meta M, Bonelli L et al. Prevalence of the Y165C, G382D and 1395delGGA germline mutations of the MYH gene in Italian patients with adenomatous polyposis coli and colorectal adenomas. Int J Cancer 2004: 109: 680–684.
8. Sieber OM, Lipton L, Crabtree M et al. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med 2003: 348: 791–799.
9. Kairupan CF, Meldrum CJ, Crooks R et al. Mutation analysis of the MYH gene in an Australian series of colorectal polyposis patients with or without germline APC mutations. Int J Cancer 2005: 116: 73–77.
10. Aretz S, Uhlhaas S, Goergens H et al. MUTYH-associated polyposis: 70 of 71 patients with biallelic mutations present with an attenuated or atypical phenotype. Int J Cancer 2006: 119: 807–814.
11. Miyaki M, Iijima T, Yamaguchi T et al. Germline mutations of the MYH gene in Japanese patients with multiple colorectal adenomas. Mutat Res 2005: 578: 430–433.
12. Isidro G, Laranjeira F, Pires A et al. Germline MUTYH (MYH) mutations in Portuguese individuals with multiple colorectal adenomas. Hum Mutat 2004: 24: 353–354.
13. Cleary SP, Cotterchio M, Jenkins MA et al. Germline MutY homologue mutations and colorectal cancer: a multisite case-control study. Gastroenterol 2009: 136: 1251–1260.
14. Bougatef K, Marrakchi R, Moussa A et al. First genetic analysis in Tunisian familial adenomatous polyposis probands. Oncol Rep 2008: 19: 1213–1218.
15. Poulsen ML, Bisgaard ML. MUTYH associated polyposis (MAP). Curr Genomics 2008: 9: 420–435.
16. Enholm S, Hienonen T, Suomalainen A et al. Proportion and phenotype of MYH-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients. Am J Pathol 2003: 163: 827–832.
17. Croitoru ME, Cleary SP, Di Nicola N et al. Association between biallelic and monoallelic germline MYH gene mutations and colorectal cancer risk. J Natl Cancer Inst 2004: 96: 1631–1634.
18. Peterlongo P, Mitra N, Chuai S et al. Colorectal cancer risk in individuals with biallelic or monoallelic mutations of MYH. Int J Cancer 2005: 114: 505–507.
19. Kim DW, Kim IJ, Kang HC et al. Germline mutations of the MYH gene in Korean patients with multiple colorectal adenomas. Int J Colorectal Dis 2007: 22: 1173–1178.
20. Ajith Kumar VK, Gold JA, Mallon E et al. Sebaceous adenomas in an MYH associated polyposis patient of Indian (Gujarati) origin. Fam Cancer 2007.
21. Nielsen M, Franken PF, Reinards TH et al. Multiplicity in polyp count and extracolonic manifestations in 40 Dutch patients with MYH associated polyposis coli (MAP). J Med Genet 2005: 42: e54.
22. Bener A, Hussain R. Consanguineous unions and child health in the state of Quatar. Paediatr Perinat Epidemiol 2006: 20: 372–378.
23. Bittles A. Consanguinity and its relevance to clinical genetics. Clin Genet 2001: 60: 89–98.
24. Rajab A, Patton MA. A study of consanguinity in the sultanate of Oman. Ann Hum biol 2000: 27: 321–326.

Article E. Cyclin D1 rare variants in UK multiple adenoma and early-onset colorectal cancer patients. (J Hum Genet)

Journal of Human Genetics (2011) 56, 58–63
© 2011 The Japan Society of Human Genetics All rights reserved 1434-5161/11 $32.00
www.nature.com/jhg

ORIGINAL ARTICLE

Cyclin D1 rare variants in UK multiple adenoma and early-onset colorectal cancer patients

Carolina Bonilla[1,2,6], Jérémie H Lefèvre[1,2,3], Bruce Winney[1], Elaine Johnstone[1], Susan Tonks[1], Chrystelle Colas[3,4], Tammy Day[1], Katarzyna Hutnik[1], Abdelhamid Boumertit[1], Rachel Midgley[1], David Kerr[1,5], Yann Parc[3] and Walter F Bodmer[1,2]

We examined the influence that rare variants and low-frequency polymorphisms in the cancer candidate gene *CCND1* have on the development of multiple intestinal adenomas and the early onset of colorectal cancer. Individuals with <100 multiple polyps and patients with colorectal cancer diagnosed before 50 years of age were recruited in UK, and screened for sequence changes in the coding and regulatory regions of *CCND1*. A set of about 800 UK control individuals was genotyped for the variants discovered in the cases. Variants in the promoter, intron-exon boundaries and untranslated regions of the *CCND1* gene had higher frequencies in cases than in controls. Five of these variants were typed in a set of French multiple adenoma and early-onset patients, who also showed higher allele frequencies than UK controls. When pooled together, variants with frequencies lower than 1% conferred an increased risk of disease that was significant in the multiple adenoma group (odds ratio (OR) 2.2; 95% confidence interval, 1.1–4.4; *P*=0.03). Most variants had a putative functional effect when assessed *in silico*. We conclude that rare variants of *CCND1* are risk factors for colorectal cancer, with considerably larger effects than common polymorphisms, and as such should be systematically evaluated in susceptibility studies.
Journal of Human Genetics (2011) **56,** 58–63; doi:10.1038/jhg.2010.144; published online 25 November 2010

Keywords: colorectal cancer; cyclin D1; multiple adenomas; rare variants; UK

INTRODUCTION

The *CCND1* gene is located on chromosome 11q13 and encodes the protein cyclin D1 that regulates cell cycle progression from G1 to S phase during cell division through its interactions with the cyclin-dependent kinases.[1] Increased cyclin D1 expression has been reported as an early event in colorectal tumorigenesis[2] and has been observed in other cancer types, including prostate, breast, lung and endometrial carcinomas.[3,4] Prior work has shown, however, that *CCND1* is not essential for the development of colorectal cancer, although it may act as a modifier of disease severity.[5]

Most association studies of *CCND1* have so far focused on the common and functionally significant G870A (P241P, rs9344) polymorphism, which affects splicing by eliminating a donor site at the end of exon 4. However, results correlating this polymorphism with cancer risk have been inconsistent. A recent meta-analysis of 60 published case-control studies has shown that, overall, individuals with the GA or AA genotype exhibited a 1.1- to 1.2-fold increased risk of developing cancer compared with individuals with the GG genotype.[6] With respect to colorectal cancer in particular, subjects with hereditary nonpolyposis colorectal cancer who carry the rs9344 polymorphism have been reported to acquire the disease at an earlier age.[7–9] Also, carriers of the A allele appear to be more frequent among individuals who developed non-syndromic colorectal cancer before the age of 60,[10] subjects with familial colorectal cancer[11,12] and affected women.[13,14] Other reports, however, do not subscribe to rs9344 genotype being a modifier of the colorectal cancer phenotype.[15,16]

Given the low odds ratios (OR) associated with common variants, such as rs9344, we have argued that genetic risk factors underlying complex diseases are more likely to be due to functionally relevant rare variants with moderate penetrance that will considerably increase susceptibility and will, therefore, sometimes justify prophylactic interventions.[17–19] We defined rare variants as those having higher frequencies than rare severe effect, clearly familial, mutations but lower frequencies than polymorphisms. Thus, rare variants will generally be in the frequency range between 0.1 and 1%. We consider low-frequency variants to be those with frequencies between 1 and 5%, which are not normally used in standard association studies. Following the strategy we previously proposed,[19] we have screened the regulatory and coding regions of *CCND1* in individuals with multiple

[1]Department of Clinical Pharmacology, University of Oxford, Headington, Oxford, UK; [2]Weatherall Institute of Molecular Medicine (WIMM), John Radcliffe Hospital, University of Oxford, Headington, Oxford, UK; [3]Department of Digestive Surgery, Hôpital Saint-Antoine, AP-HP, University Paris VI (Pierre and Marie Curie), Paris, France; [4]Laboratory of Oncogenetics, Hôpital de la Pitié-Salpétrière, AP-HP, University Paris VI (Pierre and Marie Curie), Paris, France; and [5]Sidra Medical Research Center, Doha, Qatar
[6]Current address: School of Social and Community Medicine, University of Bristol, Bristol BS8 2PS, UK
Correspondence: W Bodmer, Weatherall Institute of Molecular Medicine (WIMM), University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK.
E-mail: walter.bodmer@hertford.ox.ac.uk
Received 26 August 2010; revised 6 October 2010; accepted 26 October 2010; published online 25 November 2010

adenomas and patients with early-onset colorectal cancer recruited in the UK clinics. A few of the variants found this way were examined in a group of similarly ascertained French patients. We sought to assess the impact that collections of variants with gene frequencies lower than 1%, and between 1 and 5% have on the onset and progression of colorectal cancer, and what role, if any, the rs9344 polymorphism has in the pathogenesis of this disease.

SUBJECTS AND METHODS

Subjects

The UK patient group consisted of 112 individuals with 3–100 histologically proven synchronous or metachronous adenomatous polyps,[18] and 44 individuals with colorectal cancer diagnosed before 50 years of age. A total of 38 individuals with early-onset disease were obtained through the VICTOR clinical trial, a Phase III double-blind placebo controlled study of rofecoxib in Dukes stage B or C colorectal cancer patients following potentially curative therapy, whereas the remaining six were recruited through the John Radcliffe and Churchill hospitals' gastrointestinal clinics. With the exception of one Black Caribbean and one Indian individual, ethnic origin was White British for all UK patients for whom information was available. Non-white individuals were excluded from further analysis. No patient fulfilled the criteria for familial adenomatous polyposis, autosomal recessive *MYH*-associated polyposis or hereditary nonpolyposis colorectal cancer on clinical grounds.[18] Some of these patients had already been screened for germline mutations in the *APC* and *MYH* genes during previous studies.[20,21]

In addition, we collected samples from 131 French patients, 75 with multiple adenomas and 56 with early-onset colorectal cancer, who were recruited in the Department of Digestive Surgery at the Hôpital Saint-Antoine in Paris. All patients who underwent a colectomy or total coloproctectomy for colorectal cancer or polyposis were selected for the study *a priori*. Those diagnosed with colorectal cancer before the age of 50 years or with more than three polyps detected after 2005, were referred for a consultation with the geneticist. Immunohistochemical staining to determine loss of expression of the genes *MLH1* and *MSH2* and microsatellite status was performed for all patients with early-onset colorectal cancer. Sequencing of the entire *MYH* and *APC* genes was carried out in patients with multiple adenomatous polyps. Only patients with no evidence of hereditary nonpolyposis colorectal cancer, autosomal recessive *MYH*-associated polyposis or familial adenomatous polyposis were included in the current study. No ethnic identification was available for the French patients.

All UK and French cases had histological confirmation of adenomatous polyps, but the precise number of polyps was not determined for all of them. For 24 UK and 14 French adenoma patients, only 'multiple' was recorded.

Within both the UK and French patient groups, individuals with attenuated familial adenomatous polyposis may be included, as they were not purposefully eliminated from the study.

Controls comprised of 866 individuals, collected in 10 different regions across the UK, as part of the People of the British Isles study (see website link below), and were unselected with respect to disease status.

Blood samples from cases and controls and clinicopathological information from patients were collected with individual informed consent and local ethical committee approvals.

DNA extraction and processing

Genomic DNA for patient samples was extracted from peripheral venous blood using standard techniques. The People of the British Isles study control blood samples were transported at room temperature to the laboratory, where the peripheral blood lymphocytes were separated under sterile conditions[22] within 2 days of collection. DNA was prepared from the 10 ml blood residue, remaining after sterile separation using either magnetic beads (GeneCatcher, Invitrogen, Carlsbad, CA, USA) or spin columns (Qiagen, Valencia, CA, USA). DNA concentration was determined using PicoGreen[23] and normalized for genotyping to 25 ng μl^{-1}. Samples from UK cases underwent whole genome amplification because of limited volumes and amounts of genomic DNA. We used the Repli-g Mini kit (Qiagen) that implements a multiple displacement amplification reaction to generate up to 10 µg of DNA per 50 µl reaction from a starting amount of at least 10 ng of genomic template. Genomic DNA from French cases and UK controls was used for genotyping.

PCR amplification

All exons, 5′untranslated region (UTR), 3′UTR, intron-exon boundaries and about 1.5 kb of the *CCND1* promoter were screened, covered by a total of 21 PCR fragments (Supplementary Table 1). DNA amplification was carried out in 50 µl reactions with a final concentration of 1X PCR Gold buffer, 200 µM dNTPs and 0.5 µM of each primer. AmpliTaq Gold (2 U) (Applied Biosystems, Foster City, CA, USA), 1.5–2.5 mM of $MgCl_2$ and 20–50 ng of genomic or whole-genome amplified DNA were used per reaction. Cycling conditions basically consisted of an initial denaturing step at 95 °C for 10 min, followed by 35 cycles at 95 °C for 25 s; annealing temperature 55 °C–65 °C, 35 s; 72 °C, 30 s; and a final elongation step at 72 °C for 5 min. Agarose gels (2%) were run to verify the successful amplification of each fragment.

Mutation analysis

Mutation screening in UK patients was performed using a WAVE DNA fragment analysis dHPLC system with UV detection (Transgenomic, Omaha, NE, USA). Temperature gradients were designed using the WaveMaker software (Transgenomic) to obtain fragment-melting profiles. PCR products, of sizes between 234 and 522 bp, were denatured for 5 min at 94 °C, then gradually reannealed by decreasing the temperature to 25°C for 30 min to form homo and/or heteroduplexes. PCR products were subsequently eluted through an acetonitrile gradient at 0.9 ml min^{-1} over 6.8 min, at one or two different temperatures, selected according to their melting profiles. The column mobile phase consisted of buffer A, a 0.1 M triethylammonium acetate solution at pH 7, and buffer B, a 0.1 M triethylammonium acetate solution containing 25% acetonitrile at pH 7. The retention time of the eluate was registered by the ultraviolet detector at 260 nm. Under these conditions, invariant DNA fragments elute as a single peak, whereas variant fragments, which contain mixtures of homoduplex and heteroduplex DNAs, elute as two to four peaks or as a single peak with a shoulder.

DNA sequencing and genotyping

Cases with heteroduplex peaks, and several individuals with only homoduplex peaks were sequenced to identify the genetic variants. Each PCR product was purified following the EXO-SAP protocol and submitted to the Weatherall Institute of Molecular Medicine central facility for direct sequencing. Whenever possible, both strands were sequenced. Sequences were analyzed using Sequence Scanner version 1 (Applied Biosystems) and compared with the *CCND1* GenBank entries NM_053056 and NT_167190.

All newly discovered variants were verified by genotyping the UK cases. Controls were then genotyped for variants identified and validated in patients. Two variants (CCND1–3 and 30) were genotyped in a subset of only 222 People of the British Isles study controls. Because the screening was not fully successful for exon 2 and the distal 5′-upstream fragments, we selected 12 additional *CCND1* variants from the single nucleotide polymorphism database (dbSNP) ($\leqslant$3% minor allele frequency in HapMap European population, when reported) located in these regions, and typed them in both cases and controls. French patients were genotyped for a subset of five variants: two promoter variants (CCND1–3, 7), one 5′UTR variant (CCND1–19), one coding SNP (CCND1–21) and one 3′UTR SNP (CCND1–30). Genotyping of all variants, including rs9344, was done using the Sequenom MassArray technology, namely matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and the iPLEX Gold assay (Sequenom Inc., San Diego, CA, USA).

Statistical analysis

Hardy–Weinberg equilibrium was tested for in controls using a χ^2 goodness-of-fit test. Variant differences between cases and controls were assessed using Fisher's exact test. Two-sided *P*-values below 0.05 were considered statistically significant. Heterogeneity between UK multiple adenoma and early-onset cases was evaluated with a contingency table approach implemented in the program PowerMarker version 3.25 (Liu and Muse[24]). Odds ratios and 95% confidence interval were estimated with the software package SPSS version 16.0 (SPSS, Chicago, IL, USA).

Linkage disequilibrium patterns across *CCND1* were examined using the program Haploview version 4.1,[25] and by counts in 2×2 tables and their statistical assessment.

Functional *in silico* analysis of *CCND1* variants to identify changes in transcription factor binding and alternative splicing sites were carried out using the web programs AliBaba 2.1 and Human Splicing Finder,[26] respectively. The effect of 5′UTR and 3′UTR variants was analyzed using the online database UTRdb. Additionally, we used the GenEpi toolbox[27] to examine conservation across species, as well as disruption of microRNA target sequences in the 3′UTR.

RESULTS

The characteristics of both sets of patients and controls are summarized in Table 1 All *CCND1* variants were in Hardy–Weinberg equilibrium in controls. A list of all variants, either first identified by dHPLC and/or selected from dbSNP, is given in Supplementary Table 2, and patient and control gene frequency data are given in Supplementary Table 3.

Stratifying the cases into multiple adenoma and early-onset groups (see Table 2) suggested that there might be some differences in the variant frequencies between the UK and French groups, notably for the CCND1–3 variant. Heterogeneity test results showed that there were no significant differences in allele frequencies per locus between the UK multiple adenoma and early-onset subgroups. This justifies combining the sets of UK patients for an assessment of overall association patterns.

Rare, low frequency and common variants

Overall 20 variants were typed in UK cases and controls, of which 10 were detected through mutation analysis and 10 were selected directly from dbSNP. Four additional variants were found by dHPLC screening, but could not be validated because of technical issues with the Sequenom assays (CCND1–23, 24, 25, 26). Similarly, two further dbSNP variants were examined in patients and were not subsequently analyzed in controls because they could not be successfully typed (CCND1–5, 6). There were three common variants detected, with gene frequencies equal to or higher than 10% in the HapMap European population (CEU), one of which was rs9344 (CCND1–21, 23, 30). The other two were not included in the analysis. Another common polymorphism located in exon 5, rs7177, was excluded as well. Of all the variants successfully genotyped in patients and controls, five sites were invariant in both and these were also excluded from analysis (CCND1–1, 2, 11, 14, 16). Three variants were monomorphic only in controls (CCND1–19, 22, 28), and one was invariant only in cases (CCND1–15), and these were included in the analysis, which therefore involved 14 variants in total, including rare, low frequency and common variants. In all, 6 out of the 14 variants uncovered by dHPLC screening of the UK samples were not present in dbSNP (CCND1–17, 18, 22, 24, 25, 26). Three variants (CCND1–13, 20, 29) had reported CEU gene frequencies between 3 and 5%, and two variants (CCND1–16, 27) had CEU gene frequencies lower than 1% (see Supplementary Table 2 for more details about the variants detected, genotyped and subsequently analyzed).

Association analysis

In an analysis with relatively small numbers, it is not expected that individual variants will show significant effects unless they are associated with large ORs. Nevertheless, one rare variant (CCND1–7) did give a significant *P*-value for the difference between UK cases (combining the early onset and multiple adenoma sets) and controls (OR 3.7; 95% confidence interval 1.2–11.8; $P=0.03$) (Table 3).

Table 1 Characteristics of patient and control samples

	N	*Mean age (years)*	*Male:female*	*Mean no. of polyps*
UK multiple adenomas	112	59[a]	68:20[b]	11[b]
UK early onset	42	41	24:18	NA
French multiple adenomas	75	51[c]	44:31	26[d]
French early onset	56	40	24:32	NA
PoBI controls	866	62	478:382[e]	NA

Abbreviations: NA, not applicable; PoBI, People of the British Isles study.
[a]Missing data for 33 individuals.
[b]Missing data for 24 individuals.
[c]Missing data for three individuals.
[d]Missing data for 14 individuals.
[e]Missing data for six individuals.

Table 2 Number of patients carrying rare and low-frequency variants

ID	*dbSNP*	*UK m.a.*	*UK e.o.*	*French m.a.*	*French e.o.*
CCND1–3	rs36225071	0/109	0/41	1/75	2/56
CCND1–7	rs36225069	3/109	2/42	1/75	1/55
CCND1–10	rs36225073	6/104	2/42		
CCND1–13	rs3212859	5/101	2/41		
CCND1–15	rs954619	0/109	0/42		
CCND1–17	NA	1/108	0/42		
CCND1–18	NA	4/107	0/42		
CCND1–19	rs55911137	1/77	0/30	0/75	0/55
CCND1–20	rs3862792	5/100	2/42		
CCND1–22	NA	1/109	0/42		
CCND1–27	rs3212906	1/104	1/42		
CCND1–28	rs55666306	1/103	0/42		
CCND1–29	rs3212907	9/105	1/42		

Abbreviations: dbSNP, single nucleotide polymorphism database; e.o., early-onset disease; m.a., multiple adenoma; NA, not available.

Table 3 Carrier frequency of rare and low-frequency variants in UK cases and controls

ID	*dbSNP*	*Major/minor allele*	*Cases*[b]	*Controls*[b]	*P-value*
CCND1–3[a]	rs36225071	G/C	0/150	1/222	0.60
CCND1–7	rs36225069	A/T	5/151	7/779	0.03
CCND1–10	rs36225073	G/A	8/146	35/743	0.67
CCND1–13	rs3212859	G/T	7/142	34/749	0.83
CCND1–15	rs954619	C/T	0/151	1/747	1.00
CCND1–17	NA	T/C	1/150	11/736	0.70
CCND1–18	NA	T/C	4/149	12/730	0.33
CCND1–19	rs55911137	G/C	1/107	0/732	0.13
CCND1–20	rs3862792	C/T	7/142	35/743	0.83
CCND1–22	NA	G/C	1/151	0/750	0.17
CCND1–27	rs3212906	T/A	2/146	9/720	1.00
CCND1–28	rs55666306	G/A	1/145	0/749	0.16
CCND1–29	rs3212907	C/T	10/147	70/747	0.43

Abbreviations: dbSNP, single nucleotide polymorphism database; NA, not available.
[a]Typed in only 222 controls.
[b]Number of cases or controls carrying the minor allele/total number of individuals successfully genotyped.

When we examined all rare variants together there was a significant increase in risk (OR⩾2) for the UK multiple adenoma group as well as for all UK cases (Table 4). There was also an increase in risk for the

early-onset group considered on its own, but it did not reach significance presumably because of the small sample size.

There is strong linkage disequilibrium (LD) in the region between CCND1–10, 13 and 20 ($r^2 > 0.8$ for each pair of variants), with the minor alleles all-present in one haplotype (ATT), except for a single patient who carried the haplotype ATC. Therefore, these lower frequency variants were combined and analyzed as a single haplotype instead of independent variants. The aggregation of the CCND1–10/13/20 haplotype with the other nine variants with frequencies equal to or less than 5%, in the analysis, yielded lower ORs and a non-significant effect (Supplementary Table 4), as did a separate test on only the low frequency (1–5%) variants (Supplementary Table 5).

Table 4 Number of UK individuals with and without rare variants

UK samples	N	*⩾1 rare variants*	*No rare variants*	*OR*	*95% CI*	P-*value*
Multiple adenomas	107	12	95	2.2	1.1–4.4	0.03
Early onset	42	3	39	1.4	0.4–4.6	0.50
All UK cases	149	15	134	2.0	1.1–3.7	0.04
Controls	743	40	703	Reference		

Abbreviations: CI, confidence interval; OR, odds ratio.
Eight rare variants were used in this analysis (CCND1–7, 15, 17, 18, 19, 22, 27 and 28). To take into account some variation in the numbers of individuals typed for different variants, the relevant total number of individuals was estimated by the harmonic mean of N for all variants considered.

Functional analysis

We investigated the potential functional consequences of variants using *in silico* approaches. Promoter variants were examined for changes in transcription factor binding sites based on the matrices compiled by the program AliBaba 2.1. Several variants affect Sp1 sites (that is, CCND1–7, 10, 17), whereas CCND1–3 alters a CACCC box, which is recognized by transcription factors of the Sp/Krüppel-like Factor family. CCND1–13 creates a Cdx2 site. CCND1–18 has no apparent effect on transcription factor binding sites. For variants within the coding region, we examined the creation and/or elimination of alternative splicing sites. The rare C allele at CCND1–22 is predicted by the software tool 'Human Splicing Finder' to disrupt an acceptor site, generate a splicing enhancer site and also break up a splicing silencer site. In addition, it is well-known that the common variant rs9344 A allele modifies a splicing donor site and predisposes for an alternatively spliced transcript of the cyclin D1 protein (cyclin D1b).[28] In transcript b, no splicing occurs at the exon 4-intron 4 boundary, and exon 5 is missing.[1] Most variants found in 3′UTRs of the patients are reported to be within conserved blocks, as are the majority of the variants in the promoter regions and the exons. However, no 3′UTR variant altered a miRNA target site. There are no changes in the 5′UTR caused by CCND1–19 according to UTRdb, but 'Human Splicing Finder' indicates the potential creation of a splicing enhancer site (see Table 5 for a summary of the variants' properties).

Because this functional analysis was done *in silico*, there remains the possibility that the predicted effects are not real and may need to be confirmed by direct experimental analysis.

Table 5 Presumed effect of *CCND1* variants as determined by *in silico* tools

ID	*Minor allele*	*Gene position*	*Effect*	Conserved
CCND1–3	C	–2140	–CACCC-bi	Yes
CCND1–7	T	–1946	–Sp1/+Sp1	Yes
CCND1–10	A	–1806	+Sp1	Yes
CCND1–13	T	–1375	+Cdx2	Yes
CCND1–15	T	–1348	–CDP-CR1	Yes
CCND1–17	C	–1004	–Sp1	NA
CCND1–18	C	–457	No change	NA
CCND1–19	C	–82	New ESE	NA
CCND1–22	C	Intron 3	Acceptor site broken/ new ESE/ESS broken	NA
CCND1–20	T	Exon 4	Synonymous	Yes
CCND1–21	A	Exon 4	New acceptor site/donor site eliminated	Yes
CCND1–27	A	3′UTR		Yes
CCND1–28	A	3′UTR		NA
CCND1–29	T	3′UTR		Yes
CCND1–30	G	3′UTR		Yes

Abbreviations: ESE, exonic splicing enhancer; ESS, exonic splicing silencer; NA, not applicable; UTR, untranslated region.
CACCC, Sp1, Cdx2, CDP-CR1 are the transcription factors.
+/– denotes gain or loss of corresponding transcription factor binding site.

rs9344 G/A and rare variants

The frequency of the rs9344 A allele was higher in UK and French cases as compared with controls, although not significantly so (Supplementary Table 3). French early-onset cases showed the highest frequency (0.53), which was, however, comparable to that reported for the CEU population. UK samples had lower frequencies than CEU, in agreement with findings for the British 1958 Birth Cohort (frequency of allele A=0.44).

Linkage disequilibrium analysis using Haploview indicates that rs9344 does not appear to be in strong LD with any rare or low frequency variant, and the correlation with the haplotype harboring the variants CCND1–10, 13 and 20, is also fairly weak (r^2=0.02). However, examination of 2×2 tables reveals that LD between the CCND1–10/13/20 haplotype and rs9344 is nevertheless highly significant (Table 6). Given that it has been demonstrated that this polymorphism is not independently predictive of cancer and that additional events may be necessary to induce cyclin D1b production,[28] we examined whether the presence of rare variants along the *CCND1* gene is in any way associated with genotype at this position. We found that patients who carry at least one rare variant are more likely to also carry the rs9344 G allele than cases who do not harbor rare variants. A similar finding was obtained in controls although the effect was somewhat smaller (Supplementary Table 6). We believe, however, that this association is much more likely to be due to the effect of LD than to an actual relationship with disease.

DISCUSSION

We have screened the coding and regulatory regions of the oncogene *CCND1*, a well-known candidate for colorectal cancer susceptibility, in about 150 UK individuals with multiple intestinal polyps or early-onset colorectal cancer. In total, 14 variants were selected for analysis, including nine rare variants (MAF<1%), four lower frequency variants (MAF, 1–5%) and one common variant (MAF>5%). Three of the rare variants had not been reported elsewhere (CCND1–17, 18, 22) and only CCND1–18, was detected in more than one affected subject. Three of the rare variants were not found in a set of over 800 UK controls (CCND1–19, 22, 28). Four variants were also typed in a set of French patients with similar disease phenotypes

Table 6 Linkage disequilibrium values between haplotypes at loci CCND1–10, 13, 20 and alleles at CCND1–21 (rs9344)

CCND1–10, 13, 20\CCND1–21	GG	*GA/AA*	P-*value*
Haplotype ATT	19	14	0.001
Other haplotypes	199	489	

r^2=0.02, r=0.14, OR 3.3, 95% CI, 1.6–6.8.

to those of the UK cases. One of these variants (CCND1–19) was not present in any of the French cases, whereas another one (CCND1–3) was present in French but not in UK patients. These results indicate that current catalogs of genetic variation are not exhaustive enough to represent low frequency variation among patients, which are also often likely to be population specific. However, more information on the genetic structure of the French patient sample is necessary, as well as a set of French controls, to extract significant conclusions about the impact of *CCND1* rare variants on colorectal disease in this population.

Our findings show that, when all the rare variants are combined, there is a significant increase in the risk of colorectal disease with an OR of about 2. The effect appears to be stronger for developing multiple adenomas than for early-onset disease. In one case, CCND1–7, the effect was significant even for a single variant, with an OR greater than 3 for all UK cases.

When the data for all variants with frequencies lower than 5% (that is, rare and low frequency variants) were combined, the association was not significant. This suggests, as might be expected, that the lower the frequency of a disease-associated variant, the higher the OR is likely to be.

Nearly all the variants analyzed have putatively recognizable functional effects or lie within a conserved sequence block. However, no functional effects have yet been directly confirmed experimentally. CCND1–7 is a putative regulatory variant that alters two Sp1 binding sites. The Sp1 sites in the *CCND1* promoter are highly conserved,[29] and are required for the transcriptional activation of *CCND1* following mitogenic stimulation.[30,31]

It is interesting to note that, out of the 11 variants with MAF $<5\%$ identified in our study by chromatographic screening, 8 are located in non-coding regions of the gene, 6 of them in the 3′UTR. It has recently been established that genomic modifications of the *CCND1* 3′UTR in mantle cell lymphoma tumors produced mRNAs with truncated 3′UTRs that have considerably longer half lives than those of the full-length mRNAs.[32] These transcripts are shorter versions of cyclin D1a and are not the alternatively spliced cyclin D1b isoform. Genetic changes that generate such transcripts include deletions of part of the 3′UTR or point mutations that create novel polyadenylation signals.[32] Wiestner and colleagues believe that these alterations are somatic, although they did not examine germline DNA. Other authors[33,34] have also suggested the involvement of 3′UTR changes, including polymorphisms and rare deletions, as a cause of increased expression of *CCND1* in cancer. In addition, loss of microRNA target sites as a result of 3′UTR shortening can lead to pathogenic overexpression of the protein.[35] In our *in silico* analysis, however, none of the variants studied was suggested to modify a microRNA binding sequence.

We have also explored the putative relationship between a common variant that has been repeatedly associated with colorectal cancer, though inconsistently, and the presence of rare variants with potentially functional consequences. It could be the case that common and rare variants in the same gene or in different genes along the same pathway interact and modify each other's effects to produce the phenotype.[19,36]

A search for possible combined effects of the common variant rs9344 and one or more of the rare or low-frequency variants is almost certainly likely to be confounded by weak LD between rs9344 and these variants, and so would require a much larger sample size. This is indicated by the very significant, but low level of LD between rs9344 and the CCND1–10/13/20 haplotype.

Our analysis of variation at the *CCND1* gene has added to the evidence for the importance of rare variants as determinants of disease susceptibility. It has also shown how a moderate size study of rare variants in a candidate gene can reveal effects that are of clearly greater biological significance than very much larger whole-genome association studies (GWAS) of common variants. Thus, in these large case-control association studies only variants with frequencies higher than 5% are examined and the vast majority of significantly associated variants have been shown to give rise to very modest risk increases, generally with ORs not more than about 1.2. The case for rare variants has now been extensively discussed, both from observed data[17–19,37] and on the basis of theoretical considerations.[38–41] More recent studies have demonstrated additional rare variant influences on the pathogenesis of a variety of complex diseases and traits such as type 1 diabetes, colorectal cancer, plasma lipoprotein levels and neurological disorders.[42–49]

Screening candidate genes in groups of patients for germline variation is the first step in unraveling rare variation, a step that is already being made much easier by the increasing accessibility of next-generation sequencing technologies. However, functional studies of the most interesting variants must follow closely. Cyclin D1 is a regulator of the entrance into the G1 phase of the cell cycle, and has been considered a promising predictive and prognostic biomarker for a number of cancers. Nevertheless, there have been few assessments of *CCND1* levels of variation, with studies mostly focusing on the analysis of the functional polymorphism rs9344. More exhaustive studies that include rare and low frequency variants as well as evaluate regulatory regions are necessary if *CCND1* is to be effectively used in the clinic.

ACKNOWLEDGEMENTS

We would like to thank all individuals who contributed samples for this study. We are also grateful to Shazad Asraf, Oliver Sieber, and Haitao Wang for making samples available, and to Jennifer Wilding for helpful comments. We thank the patients who participated in VICTOR, without whose help the study would not have been possible. The trial was enabled by an educational study grant from Merck, which also supplied the study drug. VICTOR was endorsed by the UK National Cancer Research Institute (NCRI) and Cancer Research Campaign Clinical Trials Awards and Advisory Committee (CTAAC), allowing support from the UK National Cancer Research Network. The University of Birmingham was the centre for randomization and for statistics initially before these moved to the Universities of Oxford and Warwick, respectively. We also thank VICTOR trial staff: A Stanley, J Stokes, F Duchesne, K Reed, S Pendlebury, R Duvvuri, I Kennedy, G Davis, J Birtwistle, L Blair, A Bange, K Bell, L Dunham, R Deacon, A Mellor, B Mikolajczyk, M Davoudianfar and A Ferch; and VICTOR laboratory support staff: K Elliot, M Presz, H Wang; Clinical Coordinator: D Rea; The VICTOR Trial Advisory Group and the VICTOR Data and Safety Monitoring Committee, and all nurses and medical staff who participated in this trial. CB was Cancer Research UK Julia Bodmer Fellow. This work was made possible by a Cancer Research UK program grant, a Wellcome Trust grant and a Serrousi Foundation grant to WFB.

URLs

AliBaba 2.1 http://www.gene-regulation.com/pub/programs/alibaba2/index.html
British Birth Cohort http://www.b58cgene.sgul.ac.uk/index.php

dbSNP http://www.ncbi.nlm.nih.gov/projects/SNP/
GenEpi toolbox http://genepi_toolbox.i-med.ac.at/
Human Splicing Finder (HSF) http://www.umd.be/HSF/
People of the British Isles study (PoBI) http://www.peopleofthebritishisles.org
UTRdb http://utrdb.ba.itb.cnr.it/

1 Howe, D. & Lynas, C. The cyclin D1 alternative transcripts [a] and [b] are expressed in normal and malignant lymphocytes and their relative levels are influenced by the polymorphism at codon 241. *Haematologica.* **86,** 563–569 (2001).
2 Arber, N., Hibshoosh, H., Moss, S. F., Sutter, T., Zhang, Y., Begg, M. *et al.* Increased expression of cyclin D1 is an early event in multistage colorectal carcinogenesis. *Gastroenterology.* **110,** 669–674 (1996).
3 McKay, J. A., Douglas, J. J., Ross, V. G., Curran, S., Murray, G. I. & Cassidy, J. Cyclin D1 protein expression and gene polymorphism in colorectal cancer.. *Int. J. Cancer.* **88,** 77–81 (2000).
4 Ashton, K. A., Proietto, A., Otton, G., Symonds, I., McEvoy, M., Attia, J. *et al.* The influence of the Cyclin D1 870 G>A polymorphism as an endometrial cancer risk factor. *BMC Cancer.* **8,** 272 (2008).
5 Wilding, J., Straub, J., Bee, J., Churchman, M., Bodmer, W., Dickson, C. *et al.* Cyclin D1 is not an essential target of beta-catenin signaling during intestinal tumorigenesis, but it may act as a modifier of disease severity in multiple intestinal neoplasia (Min) mice. *Cancer Res.* **62,** 4562–4565 (2002).
6 Pabalan, N., Bapat, B., Sung, L., Jarjanazi, H., Francisco-Pabalan, O. & Ozcelik, H. Cyclin D1 Pro241Pro (CCND1-G870A) polymorphism is associated with increased cancer risk in human populations: a meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* **17,** 2773–2781 (2008).
7 Kong, S., Amos, C. I., Luthra, R., Lynch, P. M., Levin, B. & Frazier, M. L. Effects of cyclin D1 polymorphism on age of onset of hereditary nonpolyposis colorectal cancer. *Cancer Res.* **60,** 249–252 (2000).
8 Talseth, B. A., Ashton, K. A., Meldrum, C., Suchy, J., Kurzawski, G., Lubinski, J. *et al.* Aurora-A and cyclin D1 polymorphisms and the age of onset of colorectal cancer in hereditary nonpolyposis colorectal cancer. *Int. J. Cancer.* **122,** 1273–1277 (2008).
9 Chen, J., Etzel, C. J., Amos, C. I., Zhang, Q., Viscofsky, N., Lindor, N. M. *et al.* Genetic variants in the cell cycle control pathways contribute to early onset colorectal cancer in Lynch syndrome. *Cancer Causes Control.* **20,** 1769–1777 (2009).
10 Kong, S., Wei, Q., Amos, C. I., Lynch, P. M., Levin, B., Zong, J. *et al.* Cyclin D1 polymorphism increased risk of colorectal cancer at young age. *J. Natl. Cancer Inst.* **93,** 1106–1108 (2001).
11 Porter, T. R., Richards, F. M., Houlston, R. S., Evans, D. G. R., Jankowski, J. A., Macdonald, F. *et al.* Contribution of cyclin d1 (CCND1) and E-cadherin (CDH1) polymorphisms to familial and sporadic colorectal cancer. *Oncogene.* **21,** 1928–1933 (2002).
12 Grünhage, F., Jungck, M., Lamberti, C., Berg, C., Becker, U., Schulte-Witte, H. *et al.* Association of familial colorectal cancer with variants in the E-cadherin (CDH1) and cyclin D1 (CCND1) genes. *Int. J. Colorectal Dis.* **23,** 147–154 (2008).
13 Schernhammer, E. S., Tranah, G. J., Giovannucci, E., Chan, A. T., Ma, J., Colditz, G. A. *et al.* Cyclin D1 A870G polymorphism and the risk of colorectal cancer and adenoma. *Br. J. Cancer.* **94,** 928–934 (2006).
14 Tan, X.- L., Nieters, A., Kropp, S., Hoffmeister, M., Brenner, H. & Chang-Claude, J. The association of cyclin D1 G870A and E-cadherin C-160A polymorphisms with the risk of colorectal cancer in a case control study and meta-analysis. *Int. J. Cancer.* **122,** 2573–2580 (2008).
15 Bala, S. & Peltomaki, P. Cyclin D1 as a genetic modifier in hereditary nonpolyposis colorectal cancer. *Cancer Res.* **61,** 6042–6045 (2001).
16 Kruger, S., Engel, C., Bier, A., Mangold, E., Pagenstecher, C., Von Knebel Doeberitz, M. *et al.* Absence of association between cyclin D1 (CCND1) G870A polymorphism and age of onset in hereditary nonpolyposis colorectal cancer. *Cancer Lett.* **236,** 191–197 (2006).
17 Bodmer, W. Familial adenomatous polyposis (FAP) and its gene, APC. *Cytogenet. Genome Res.* **86,** 99–104 (1999).
18 Fearnhead, N. S., Wilding, J. L., Winney, B., Tonks, S., Bartlett, S., Bicknell, D. C. *et al.* Multiple rare variants in different genes account for multifactorial inherited susceptibility to colorectal adenomas. *Proc. Natl Acad. Sci. USA.* **101,** 15992–15997 (2004).
19 Bodmer, W. & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat. Genet.* **40,** 695–701 (2008).
20 Frayling, I. M., Beck, N. E., Ilyas, M., Dove-Edwin, I., Goodman, P., Pack, K. *et al.* The APC variants I1307 K and E1317Q are associated with colorectal tumors, but not always with a family history. *Proc. Natl Acad. Sci. USA.* **95,** 10722–10727 (1998).
21 Lamlum, H., Al Tassan, N., Jaeger, E., Frayling, I., Sieber, O., Reza, F. B. *et al.* Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q. *Hum. Mol. Genet.* **9,** 2215–2222 (2000).
22 Bodmer, W. F. & Bodmer, J. G. Cytofluorochromasia for HLA A, B, C and DR typing. In: *NIAID Manual of Tissue Typing Techniques 1979–1980*, Vol. 80. NIH Publication: Bethesda, MD, pp 545–546.
23 Ahn, S. J., Costa, J. & Rettig Emanuel, J. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. *Nucleic Acids Res.* **24,** 2623–2625 (1996).
24 Liu, K. & Muse, S. V. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics.* **21,** 2128–2129 (2005).
25 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* **21,** 263–265 (2005).
26 Desmet, F.- O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M. & Béroud, C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37,** e67 (2009).
27 Coassin, S., Brandstätter, A. & Kronenberg, F. Lost in the space of bioinformatic tools: a constantly updated survival guide for genetic epidemiology. The GenEpi Toolbox. *Atherosclerosis.* **209,** 321–335 (2010).
28 Comstock, C. E. S., Augello, M. A., Benito, R. P., Karch, J., Tran, T. H., Utama, F. E. *et al.* Cyclin D1 splice variants: polymorphism, risk, and isoform-specific regulation in prostate cancer. *Clin. Cancer Res.* **15,** 5338–5349 (2009).
29 Eto, I. Molecular cloning and sequence analysis of the promoter region of mouse cyclin D1 gene: implication in phorbol ester-induced tumour promotion. *Cell Prolif.* **33,** 167–187 (2000).
30 Nagata, D., Suzuki, E., Nishimatsu, H., Satonaka, H., Goto, A., Omata, M. *et al.* Transcriptional activation of the cyclin D1 gene is mediated by multiple cis-elements, including SP1 sites and a cAMP-responsive element in vascular endothelial cells. *J. Biol. Chem.* **276,** 662–669 (2001).
31 Bartusel, T., Schubert, S. & Klempnauer, K.- H. Regulation of the cyclin D1 and cyclin A1 promoters by B-Myb is mediated by Sp1 binding sites. *Gene.* **351,** 171–180 (2005).
32 Wiestner, A., Tehrani, M., Chiorazzi, M., Wright, G., Gibellini, F., Nakayama, K. *et al.* Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. *Blood.* **109,** 4599–4606 (2007).
33 Sathyan, K. M., Nalinakumari, K. R., Abraham, T. & Kannan , S. CCND1 polymorphisms (A870G and C1722G) modulate its protein expression and survival in oral carcinoma. *Oral Oncol.* **44,** 689–697 (2008).
34 Hosokawa, Y., Suzuki, R., Joh, T., Maeda, Y., Nakamura, S., Kodera, Y. *et al.* A small deletion in the 3′-untranslated region of the cyclin D1/PRAD1/bcl-1 oncogene in a patient with chronic lymphocytic leukemia. *Int. J. Cancer.* **76,** 791–796 (1998).
35 Deshpande, A., Pastore, A., Deshpande, A. J., Zimmermann, Y., Hutter, G., Weinkauf, M. *et al.* 3′UTR mediated regulation of the cyclin D1 proto-oncogene. *Cell Cycle.* **8,** 3584–3592 (2009).
36 Felix, R., Bodmer, W., Fearnhead, N. S., van der Merwe, L., Goldberg, P. & Ramesar, R. S. GSTM1 and GSTT1 polymorphisms as modifiers of age at diagnosis of hereditary nonpolyposis colorectal cancer (HNPCC) in a homogeneous cohort of individuals carrying a single predisposing mutation. *Mut. Res.* **602,** 175–181 (2006).
37 Cohen, J. C., Kiss, R. S., Pertsemlidis, A., Marcel, Y. L., McPherson, R. & Hobbs, H. H. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science.* **305,** 869–872 (2004).
38 Pritchard, J. K. & Cox, N. J. The allelic architecture of human disease genes: common disease-common variant..or not? *Hum. Mol. Genet.* **11,** 2417–2423 (2002).
39 Schork, N. J., Murray, S. S., Frazer, K. A. & Topol, E. J. Common vs. rare allele hypotheses for complex diseases. *Curr. Opin. Genet. Dev.* **19,** 212–219 (2009).
40 Goldstein, D. B. Common genetic variation and human traits. *New Eng. J. Med.* **360,** 1696–1698 (2009).
41 Eyre-Walker, A. Evolution in health and medicine Sackler colloquium: genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proc. Natl Acad. Sci. USA.* **107,** 1752–1756 (2010).
42 Nejentsev, S., Walker, N., Riches, D., Egholm, M. & Todd, J. A. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science.* **324,** 387–389 (2009).
43 Kaufmann, A., Vogt, S., Uhlhaas, S., Stienen, D., Kurth, I., Hameister, H. *et al.* Analysis of rare APC variants at the mRNA level: six pathogenic mutations and literature review. *J. Mol. Diagn.* **11,** 131–139 (2009).
44 Azzopardi, D., Dallosso, A. R., Eliason, K., Hendrickson, B. C., Jones, N., Rawstorne, E. *et al.* Multiple rare nonsynonymous variants in the adenomatous polyposis coli gene predispose to colorectal adenomas. *Cancer Res.* **68,** 358–363 (2008).
45 Mefford, H. C., Cooper, G. M., Zerr, T., Smith, J. D., Baker, C., Shafer, N. *et al.* A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res.* **19,** 1579–1585 (2009).
46 Burnett, J. R. & Hooper, A. J. Common and rare gene variants affecting plasma LDL cholesterol. *Clin. Biochem. Rev.* **29,** 11–26 (2008).
47 Cohen, J. C., Pertsemlidis, A., Fahmi, S., Esmail, S., Vega, G. L., Grundy, S. M. *et al.* Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc. Natl Acad. Sci. USA.* **103,** 1810–1815 (2006).
48 Walsh, T., McClellan, J. M., McCarthy, S. E., Addington, A. M., Pierce, S. B., Cooper, G. M. *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science.* **320,** 539–543 (2008).
49 Bearden, C. E., Glahn, D. C., Lee, A. D., Chiang, M.- C., Van Erp, T. G. M., Cannon, T. D. *et al.* Neural phenotypes of common and rare genetic variants. *Biol. Psychol.* **79,** 43–57 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

Article F. Role of rare variants in undetermined multiple adenomatous polyposis and early onset colorectal cancer. (En correction J Hum Genet)

Journal of Human Genetics (2012), 1–8
© 2012 The Japan Society of Human Genetics All rights reserved 1434-5161/12 $32.00
www.nature.com/jhg

ORIGINAL ARTICLE

Role of rare variants in undetermined multiple adenomatous polyposis and early-onset colorectal cancer

Jérémie H Lefevre[1,2,3,7], Carolina Bonilla[1,2,7,8], Chrystelle Colas[3,4], Bruce Winney[1], Elaine Johnstone[5], Susan Tonks[1], Tammy Day[1], Katarzyna Hutnik[1], Abdelhamid Boumertit[1], Florent Soubrier[4], Rachel Midgley[1], David Kerr[6], Yann Parc[3] and Walter F Bodmer[1,2]

Some 15–20% of multiple adenomatous polyposis have no genetic explanation and 20–30% of colorectal cancer (CRC) cases are thought to be due to inherited multifactorial causes. Accumulation of deleterious effects of low-frequency dominant and independently acting variants may be a partial explanation for such patients. The aim of this study was to type a selection of rare and low-frequency variants (<5%) to elucidate their role in CRC susceptibility. A total of 1181 subjects were included (866 controls; 315 cases). Cases comprised UK ($n=184$) and French ($n=131$) patients with MAP ($n=187$) or early-onset CRC ($n=128$). Seventy variants in 17 genes were examined in cases and controls. The effect of the variant effect on protein function was investigated *in silico*. Out of the 70 variants typed, 36 (51%) were tested for association. Twenty-one variants were rare (minor allele frequency (MAF) <1%). Four rare variants were found to have a significantly higher MAF in cases (EXO1-12, MLH1-1, CTNNB1-1 and BRCA2-37, $P<0.05$) than in controls. Pooling all rare variants with a MAF <0.5% showed an excess risk in cases (odds ratio = 3.2; 95% confidence interval = 1.1–9.5; $P=0.04$). Rare variants are important risk factors in CRC and, as such, should be systematically assayed alongside common variation in the search for the genetic basis of complex diseases.

Journal of Human Genetics advance online publication, 9 August 2012; doi:10.1038/jhg.2012.99

Keywords: adenomatous polyposis; APC; colorectal neoplasia; rare variant

INTRODUCTION

Familial adenomatous polyposis (FAP) and *MUTYH*-associated polyposis (MAP) are two inherited syndromes that show a high incidence of adenomatous polyps and an elevated risk of developing colorectal cancer (CRC). They account for a small fraction of CRC, <4%.[1] However, despite the fact that these two syndromes are caused by deleterious highly penetrant mutations in *APC* (GeneID 324)[2] and *MUTYH* (GeneID 4595),[3] around 15–20% of patients with polyposis exhibit no known genetic risk factors. This is especially so for multiple polyposis patients who carry between 3 and 100 adenomatous polyps. In addition, 20–30% of CRC is thought to be due to inherited multifactorial causes.[4] In the absence of identification of a new deleterious mutation, CRC may in part be due to the summation of the deleterious effects of a series of low-frequency dominant and independently acting variants of a variety of different genes, each, conferring a moderate but readily detectable increase in relative risk.[4] This 'rare variant' hypothesis was based upon the observation by Frayling *et al.*[5] of the APC I1307K and E1317Q variants in patients with multiple adenomas. The I1307K variant is found in the Ashkenazi Jewish population at a frequency of ~6–7%, whereas it is absent from non-Jewish populations, and confers an increased risk of multiple adenomas and CRC.[5] This variant implies an amino (isoleucine to lysine) substitution in a region involved in protein binding, leading to a mild dominant-negative effect. The E1317Q variant substitution may also affect the function of the APC protein presumed to translate into a slight but definitive advantage for the growth of a tumor.[6]

Following these observations, other rare variants have been tested. The candidate variants were selected because of their known involvement in sporadic or hereditary CRC or adenomas. Fearnhead *et al.*[7]

[1]Department of Clinical Pharmacology, Cancer and Immunogenetics Laboratory, University of Oxford, Oxford, UK; [2]Weatherall Institute of Molecular Medicine (WIMM), John Radcliffe Hospital, University of Oxford, Oxford, UK; [3]Department of Digestive Surgery, Hôpital Saint-Antoine AP-HP, University Paris VI (Pierre and Marie Curie), Paris, France; [4]Laboratory of Angiogenetics and Oncogenetics, Hôpital Pitié-Salpétrière AP-HP, University Paris VI (Pierre and Marie Curie), Paris, France; [5]Department of Clinical Pharmacology, Translational Oncology Group, Oxford, UK and [6]Nuffield Department of Clinical Laboratory Sciences, John Radcliffe Hospital, University of Oxford, Oxford, UK
[7]These authors contributed equally to this work.
[8]Current address: School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK.
Correspondence: Professor WF Bodmer, Department of Digestive Surgery, Hopital Saint-Antoine AP-HP, University Paris VI (Pierre and Marie Curie), 184, rue du Fbg Saint-Antoine, Paris 75012, France.
E-mail: walter.bodmer@hertford.ox.ac.uk
Received 26 April 2012; revised 4 July 2012; accepted 13 July 2012

observed a cumulative effect of 13 rare variants on five different genes in a cohort of 124 patients with adenomatous polyposis with an overall odds ratio (OR) of 2.2 ($P=0.0001$) when compared with a control set. Because of this publication, several variants in different CRC-susceptibility genes, such as *hMLH1* and *hMSH6*, have been reported to increase the risk of CRC but not cause Lynch syndrome,[8,9] whereas *CHEK2* confers a higher CRC risk in hereditary non polyposis CRC (HNPCC)/HNPCC-related families.[10]

Rare variants are defined by a minor allele frequency (MAF) $<1\%$ in the general population and are unlikely to be identified by genome-wide association studies due to their low frequency and small contribution to the overall susceptibility of a disease.[4] Only variants with a frequency $>5\%$ are detected in these large case–control association studies. Rare variants are best identified in studies with selected cases and candidate genes already known to be likely to be functionally relevant.[1,5] Patients with early-onset CRC (before the age of 50) and multiple polyposis (3–100 polyps) with no known mutations in *APC* or *MYH* are ideal candidates to demonstrate an elevated predisposition to disease due to the accumulation of rare variants, as they are likely to involve inherited susceptibility.

The aim of this study was therefore to type a selection of rare (MAF $<1\%$) and low-frequency variants (MAF 1–5%) in a relatively large set of patients with undetermined multiple polyposis (3–100 polyps) or early-onset CRC (diagnosed before 50 years of age) in order to elucidate the wider role of such variants in CRC susceptibility.

MATERIALS AND METHODS

A total of 315 cases and 866 controls, 1181 subjects in all, were included in this study. Collection of blood samples from cases and controls and clinicopathological information from patients were undertaken with appropriate individual informed consent and local ethical committee approvals.

Controls

The controls comprised 866 individuals collected in 10 different regions across the United Kingdom as part of the People of the British Isles study[11] (see below) and were unselected with respect to disease status.

Cases

The UK patient group consisted of 112 individuals with 3–100 histologically proven synchronous or metachronous adenomatous polyps and 72 individuals with CRC diagnosed before 50 years of age. Sixty-three individuals with early-onset disease were obtained through the VICTOR clinical trial, a phase III double-blind placebo-controlled study of rofecoxib in Dukes stage B or C CRC patients following potentially curative therapy, whereas the remaining nine cases were recruited through the John Radcliffe and Churchill hospitals' gastrointestinal clinics. With the exception of one Black Caribbean and one Indian individual, ethnic origin was White British for all UK patients for whom information was available. Non-white individuals were excluded from further analysis. No patient fulfilled the criteria for FAP, autosomal recessive MAP or HNPCC on clinical grounds. Some of these patients had already been screened for germline mutations in the *APC* and *MYH* genes in previous studies.[5,12]

We also collected samples from 131 French patients, including 75 with multiple adenomas and 56 with early-onset CRC, who were recruited in the Department of Digestive Surgery at the Hospital Saint-Antoine in Paris using the criteria described above. Cases were selected from those who underwent a colectomy or total coloproctectomy for CRC or polyposis. Patients diagnosed with CRC before the age of 50 or with more than three polyps detected after 2005 were referred for a consultation with the geneticist. Immunohistochemical staining to determine loss of expression of the genes *MLH1* and *MSH2* and microsatellite status was performed for all patients with a CRC diagnosed before the age of 50. Microsatellite instability was confirmed using PCR as already described.[13] Sequencing of the entire *MUTYH* and *APC* genes was carried out in patients with more than three adenomatous polyps.[14] Only patients with no indication of HNPCC, MAP or FAP were included in this study. No ethnic identification was available for the French patients.

All the UK and French cases had histological confirmation of adenomatous polyps, but not all of them had the precise number of polyps determined. For 24 UK and 14 French adenoma patients, only 'multiple' was recorded. Within both the UK and French patient groups, individuals with attenuated FAP may be included, as they were not purposely eliminated from the study.

Variant selection

Rare and low-frequency variants were chosen based on prior literature reports that suggested a putative association with CRC, and with features related to colorectal disease, gastric cancer or other cancers (mostly of the breast and prostate). Variants in the *BRCA* genes were specifically selected from those that were classified either as non-pathogenic or as of unknown significance (see Breast Cancer Information Core database). Common variants, such as *CDH1* rs16260, *MTHFR* rs1801133 and *TP53* rs1042522, were genotyped because it has been suggested that they are associated with several types of cancer, including CRC.[15–19]

DNA extraction and processing

Genomic DNA was extracted from patients' peripheral venous blood using the standard techniques. The People of the British Isles control blood samples were transported at room temperature to the laboratory, where the peripheral blood lymphocytes were separated under sterile conditions within 2 days of collection. DNA was prepared from the 10-ml blood residue remaining after sterile separation using either magnetic beads (GeneCatcherTM; Invitrogen, Carlsbad, CA, USA) or spin columns (Qiagen, Valencia, CA, USA). DNA concentration was determined using Pico Green[20] and normalized for genotyping to $25\,ng\,\mu l^{-1}$. Samples from the UK cases underwent whole-genome amplification because of limited volumes and amounts of genomic DNA. We used the Repli-g Mini kit (Qiagen), which implements a multiple-displacement amplification reaction to generate up to $10\,\mu g$ of DNA per $50\,\mu l$ reaction from a starting amount of at least 10 ng of genomic template. Genomic DNA from French cases and UK controls was used for genotyping.

Genotyping

We examined 70 variants, in cases and controls, in the following cancer candidate genes: *APC, AXIN1, AXIN2, BRCA1, BRCA2, CDH1, CHEK2, CTNNB1, EPHB2, EXO1, MLH1, MLH3, MSH2, MTHFR, PMS2, SMAD4* and *TP53*, which were selected based on their involvement in familial cancers and the presence of somatic mutations in cancer. Of these, 16 variants were genotyped in a subset of only 227 controls. The complete list of variants analyzed is given in Supplementary Table 1.

Four variants were genotyped using restriction fragment length polymorphism analysis (CDH1-2, CHEK2-1, CTNNB1-1 and MSH2-8). Variants APC-10 and APC-11 (that is, I1307K and E1317Q) were typed using allele-specific PCR.[7] For details of primers, enzymes and fragment sizes for these see Supplementary Table 2. Primers and conditions for all other variants are available upon request. Genotyping of the remaining variants was done using the Sequenom MassArray technology, namely matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and the iPLEX Gold assay (Sequenom Inc., San Diego, CA, USA).[21]

Statistical analysis

Hardy–Weinberg equilibrium was assessed using an exact test implemented in the program PLINK v.1.07.[22] Case–control association analyses were also conducted with PLINK. Two-sided *P*-values were calculated using Fisher's exact test and those <0.05 (with no multiple comparison correction) were considered statistically significant for an initial analysis. Combined ORs were estimated using the Mantel–Haenszel test.[23,24]

Functional *in silico* analysis

We used the web-based programs PolyPhen-2 and SNPs&GO to predict the effect of nonsynonymous variants on protein function.[25,26] FastSNP and F-SNP were similarly used for noncoding variants.[27,28]

RESULTS

Populations

Clinical characteristics of the patients and controls are shown in Table 1.

Variant selection

Among the 70 variants examined, 24 (34%) were monomorphic in both the UK cases and controls and therefore not useful for analysis (Supplementary Table 1). Of the remaining 46, 10 were monomorphic in cases only and 5 were monomorphic only in controls. Thirty-one variants were considered rare having a control-population MAF $<1\%$, seven were low-frequency variants (that is, MAF between 1 and 5%) and eight were common polymorphisms (that is, MAF $>5\%$). If we define variant class based on the combined MAF in cases and controls as recently suggested,[29,30] only one variant (APC-17) changes categories, going from being a low-frequency variant to a rare one.

No variant was out of Hardy–Weinberg equilibrium in the control population at a Bonferroni-corrected *P*-value of $\leqslant 0.001$ (0.05/46). Two variants were in Hardy–Weinberg disequilibrium in controls (TP53-1 and BRCA1-6, $P<0.05$) and three in patients (EPHB2-3, EXO1-12 and BRCA1-22, $P<0.05$) if a correction for multiple testing was not applied.

Association analysis

UK cases vs controls. When comparing UK cases with controls, four rare variants were found to have a significantly higher MAF among the patients (EXO1-12, MLH1-1, CTNNB1-1 and BRCA2-37, $P<0.05$; Table 2). Variant EXO1-12 was more frequent in individuals with cancer than in those with adenomas, as opposed to MLH1-1, CTNNB1-1 and BRCA2-37, which were only present in the multiple adenoma cases (Table 2). Results close to significance were also seen for rare variant EPHB2-3 and common variant CDH1-2 ($P=0.07$ for both), although the CDH1-2 A allele appeared to protect against disease. When analyzing carrier frequencies instead of allele frequencies, only BRCA2-37 was significant, with an OR of 4.1 (1.2–14.3; $P=0.05$; Table 3). Pooling together all the rare variants showed that the proportion of patients carrying rare variants was higher than the proportion of control carriers, regardless of whether the full set or a subset of controls was used (Table 3). The combined OR, obtained by merging OR 1 (effect of variants typed in the full set of controls) and OR 2 (effect of variants typed in the smaller set of controls) with the Mantel–Haenszel test, was 1.2 (95% confidence interval (CI) 0.8–1.8, $P=0.42$). This effect became much stronger when only variants with a MAF $<0.5\%$ were tested (combined OR 1.8; 95% CI 1.0–3.1; $P=0.05$; Table 3). On the other hand, the analysis of pooled low-frequency variants showed a protective but nonsignificant effect (combined OR 0.8; 95% CI 0.5–1.1; $P=0.18$; Supplementary Table 3). When APC-17 switched categories, from low frequency to rare variant, results changed slightly. For rare variants with MAF $<1\%$, combined OR = 1.1 (95% CI 0.8–1.7; $P=0.54$); for low-frequency variants, combined OR = 0.8 (95% CI 0.5–1.2; $P=0.24$). Also, two variants (MLH3-1 and CHEK2-1) do not make the 0.5% cutoff when assessed from the frequency in the combined set of cases and controls. Taking them out of the analysis of variants with MAF $<0.5\%$ yields a combined OR of 1.6 (95% CI 0.8–2.9; $P=0.17$).

UK multiple adenoma patients vs early-onset CRC patients. Analysis by disease group (that is, multiple adenoma vs early-onset patients) of all variants with frequencies $<0.5\%$ revealed an increase in susceptibility to disease for carriers of rare variants, especially among multiple adenoma patients (combined OR 1.9; 95% CI 1.0–3.5; $P=0.05$; Table 4). Individually, the carrier frequency for BRCA2-37 in multiple polyp cases was significantly higher than that of controls, whereas this variant was absent among early-onset patients. MLH3-1, on the other hand, showed a significantly higher carrier frequency in the early-onset group as compared with controls, whereas BRCA2-27 was only detected in individuals with early-onset disease (Table 4). Overall, out of the 31 variants with MAF $<1\%$, 14 were present in multiple adenoma patients only, whereas 3 (4, if counting APC-17) were present only in early-onset cases. The difference was significant whether APC-17 is included or not ($P<0.05$), although the smaller number of early-onset CRC patients might be introducing bias.

There were significant allele-frequency differences between individuals with multiple adenomas and those with early-onset CRC in two variants in *MLH3* (one rare (MLH3-1) and one low frequency (MLH3-5)) and in a common variant in *CDH1* (CDH1-5) (Table 2). In these three instances, the allele frequency among early-onset patients was higher than among individuals with multiple adenomatous polyps.

Comparison between UK and French samples

Twenty-two of the variants genotyped in UK patients (16 rare, 4 low frequency and 2 common variants) were also examined in French subjects affected by either multiple polyps or early-onset CRC, recruited using the same set of criteria employed in the United Kingdom (Table 5). Three rare variants (MSH2-8, APC-10 and BRCA2-48) were absent from both, the UK and the French population. Eight variants that were detected in UK cases were not found in French patients (EPHB2-1, EPHB2-4, EPHB2-7, EXO1-4, CTNNB1-1, BRCA2-35, BRCA2-37 and CHEK2-1). On the other hand, no variant identified in French patients was missing from the UK sample (cases and/or controls). There was no significant difference between the French and UK cases with respect to the overall number of rare variants with MAF $<1\%$, but UK patients show an excess of rare variants with MAF $<0.5\%$ compared with this set of French patients ($P=0.02$; Table 6).

In silico analysis of functional effects

We investigated the putative effects of nonsynonymous variants with the programs PolyPhen-2 and SNPs&GO. Based on PolyPhen-2, 11 variants were classified as probably damaging, 6 as possibly damaging and 22 as benign, whereas according to SNP&GO, there were 18 disease variants and 21 neutral variants. Although these numbers

Table 1 Case and control sample description

	N	Mean age (years)	Male:female	Mean no. of polyps	CCR
UK multiple adenomas	112	59[a]	68:20[b]	11[b]	9
UK early onset	70	42	38:31[c]	n/a	
French multiple adenomas	75	51[d]	44:31	26[e]	34
French early onset	56	40	24:32	n/a	
PoBI controls	866	62	478:382[f]	n/a	n/a

Abbreviation: PoBI, People of the British Isles; n/a, not applicable.
Missing data for: [a]33, [b]24, [c]1, [d]3, [e]14, [f]6 individuals.

Table 2 Variants analyzed in UK cases and controls

Id	*Variant*	*dbSNP*	*Major/minor allele*	*MAF cases*	*MAF controls*	P-*value*[a]	*MAF multiple adenomas*	*MAF early onset*	P-*value*[b]	*PolyPhen-2/FastSNP*	*SNPs&GO*
MTHFR-1	A222V	rs1801133	C/T	0.338	0.336	0.95	0.337	0.342	0.94	Probably damaging	Disease
EPHB2-1	R80H	n/a	G/A	0.003	0.001	0.19	0.005	0.000	0.53	Probably damaging	Disease
EPHB2-3	I361V	rs56180036	A/G	0.007	0.001	0.07	0.010	0.000	0.37	Benign	Neutral
EPHB2-4	R568W	n/a	C/T	0.003	0.001	0.19	0.000	0.012	0.11	Probably damaging	Disease
EPHB2-7	M883V	n/a	A/G	0.003	0.000	0.22	0.005	0.000	0.53	Possibly damaging	Neutral
EXO1-2	E109K	n/a	G/A	0.000	0.001	0.66	0.000	0.000	n/a	Possibly damaging	Neutral
EXO1-12	D249N	rs61750993	G/A	0.020	0.007	**0.03**	0.018	0.024	0.74	Probably damaging	Neutral
EXO1-4	L410R	n/a	T/G	0.004	0.000	0.18	0.005	0.000	0.54	Probably damaging	Neutral
EXO1-10	G759E	rs4150001	G/A	0.003	0.009	0.38	0.000	0.013	0.11	Benign	Neutral
MLH1-1	G22A	rs41295280	G/C	0.003	0.000	**0.03**	0.005	0.000	0.53	Probably damaging	Neutral
CTNNB1-1	N287S	rs35288908	A/G	0.003	0.000	**0.05**	0.005	0.000	0.42	Benign	Disease
APC-7	L1129S	n/a	T/C	0.000	0.003	0.37	0.000	0.000	n/a	Possibly damaging	Disease
APC-11	E1317Q	rs1801166	G/C	0.011	0.007	0.46	0.014	0.007	0.59	Benign	Neutral
APC-15	G2502S	rs2229995	G/A	0.010	0.021	0.19	0.009	0.012	0.83	Benign	Neutral
APC-16	R2505Q	n/a	G/A	0.000	0.002	0.43	0.000	0.000	n/a	Probably damaging	Neutral
APC-17	S2621C	rs72541816	C/G	0.003	0.011	0.25	0.000	0.012	0.12	Benign	Neutral
APC-20	8636 C/A	n/a	C/A	0.024	0.019	0.66	0.024	0.025	0.96	n/a	n/a
PMS2-1	T511A	rs2228007	A/G	0.023	0.029	0.60	0.023	0.024	0.55	Benign	Neutral
PMS2-2	T597S	rs1805318	A/T	0.034	0.020	0.23	0.043	0.012	0.19	Benign	Neutral
PMS2-3	M622I	rs1805324	G/A	0.017	0.022	0.59	0.014	0.024	0.98	Benign	Neutral
BRCA2-7	N372H	rs144848	T/G	0.320	0.286	0.23	0.329	0.298	0.60	Benign	Disease
BRCA2-8	S384F	rs41293475	C/T	0.000	0.003	0.38	0.000	0.000	n/a	Possibly damaging	Disease
BRCA2-27	R2034C	rs1799954	C/T	0.004	0.004	0.90	0.000	0.012	0.12	Benign	Disease
BRCA2-35	D2665G	rs28897745	A/G	0.003	0.000	0.22	0.005	0.000	0.53	Probably damaging	Disease
BRCA2-37	V2728I	rs28897749	G/A	0.014	0.003	**0.02**	0.019	0.000	0.21	Benign	Disease
MLH3-1	V741F	rs28756990	G/T	0.014	0.004	0.16	0.005	0.036	**0.04**	Benign	Neutral
MLH3-2	M809V	rs61752722	A/G	0.003	0.009	0.37	0.005	0.000	0.53	Benign	Neutral
MLH3-5	S845G	rs28756992	A/G	0.014	0.033	0.09	0.005	0.036	**0.04**	Benign	Neutral
AXIN1-4	D495E	n/a	C/G	0.004	0.009	0.44	0.007	0.000	0.37	Benign	Neutral
AXIN1-6	R841Q	rs34015754	G/A	0.007	0.007	0.92	0.010	0.000	0.45	Probably damaging	Disease
CDH1-1	−1128	rs13335980	A/T	0.000	0.001	0.53	0.000	0.000	n/a	S8-binding site	n/a
CDH1-2	−284	rs16260	C/A	0.246	0.308	0.07	0.218	0.291	0.12	Affects TF-binding sites	n/a
CDH1-3	IVS1+6	rs3743674	T/C	0.102	0.121	0.37	0.116	0.071	0.27	Splicing site (medium-high risk)	n/a
CDH1-5	IVS4+10	rs33963999	G/C	0.070	0.080	0.57	0.051	0.119	**0.04**	Intronic enhancer (very low-low risk)	n/a
CDH1-7	A592T	rs35187787	G/A	0.003	0.007	0.45	0.005	0.000	0.53	Benign	Neutral
CDH1-8	T599S	n/a	C/G	0.000	0.001	0.65	0.000	0.000	n/a	Benign	Neutral
CDH1-10	A634V	n/a	C/T	0.000	0.001	0.52	0.000	0.000	n/a	Possibly damaging	Disease
TP53-1	R72P	rs1042522	G/C	0.255	0.248	0.80	0.270	0.220	0.38	Benign	Disease
BRCA1-6	Q356R	rs1799950	A/G	0.042	0.060	0.22	0.043	0.038	0.53	Possibly damaging	Disease
BRCA1-8	R496H	rs28897677	G/A	0.000	0.001	0.65	0.000	0.000	n/a	Benign	Disease
BRCA1-16	T826K	rs28897683	C/A	0.000	0.001	0.65	0.000	0.000	n/a	Probably damaging	Disease
BRCA1-22	E1038G	rs16941	A/G	0.338	0.342	0.90	0.346	0.317	0.64	Probably damaging	Disease
BRCA1-28	S1512I	rs1800744	G/T	0.003	0.004	0.82	0.005	0.000	0.86	Benign	Disease
AXIN2-1	N412S	rs115931022	A/G	0.000	0.007	0.19	0.000	0.000	n/a	Benign	Neutral
SMAD4-1	A118A	n/a	G/A	0.003	0.005	0.75	0.005	0.000	0.53	Synonymous	n/a
CHEK2-1	1100 delC	n/a	C/del	0.008	0.003	0.31	0.009	0.007	0.88	p.T367fsX15	n/a

Abbreviations: dbSNP, single nucleotide polymorphism database; MAF, minor allele frequency; n/a, not available; TF, transcription factor.
Number in bold: $P \leqslant 0.05$.
Italics: typed in a subset of 227 controls.
[a]*P*-value for the comparison case vs control.
[b]*P*-value for the comparison multiple adenoma vs early-onset CRC.

Table 3 Rare variant counts in UK cases and controls

Rare variant	*Cases*[a]	*Controls*[a]	P-*value*
EPHB2-1	1/149	1/778	0.30
EPHB2-3	1/145	2/746	0.41
EPHB2-4	1/149	1/775	0.30
EPHB2-7	1/149	0/224	0.40
EXO1-2	0/150	1/751	1.00
EXO1-4	1/125	0/226	0.36
EXO1-10	1/145	4/225	0.65
EXO1-12	5/150	10/745	0.15
MLH1-1	1/147	0/740	0.17
CTNNB1-1	1/174	0/702	0.20
APC-7	0/149	4/743	1.00
APC-11	4/176	11/745	0.50
APC-16	0/150	3/729	1.00
BRCA2-8	0/147	4/748	1.00
BRCA2-27	1/142	6/740	1.00
BRCA2-35	1/149	0/224	0.40
BRCA2-37	4/148	5/744	**0.05**
MLH3-1	4/144	2/225	0.21
MLH3-2	1/148	4/227	0.65
AXIN1-4	1/117	14/749	0.71
AXIN1-6	2/146	11/746	1.00
CDH1-1	0/150	2/748	1.00
CDH1-7	1/147	11/747	0.70
CDH1-8	0/151	1/742	1.00
CDH1-10	0/150	2/731	1.00
BRCA1-8	0/151	1/735	1.00
BRCA1-16	0/149	1/744	1.00
BRCA1-28	1/149	2/225	1.00
AXIN2-1	0/128	10/737	0.37
SMAD4-1	1/148	7/739	1.00
CHEK2-1	3/178	1/179	0.37
MAF <1% (n = 31)			
Total carriers/total noncarriers 1[b]	24/146	108/739	
OR 1 (95% CI)[c]	1.13 (0.70, 1.81)		0.63
Total carriers/total noncarriers 2[b]	13/146	13/217	
OR 2 (95% CI)[c]	1.49 (0.68, 3.24)		0.33
Combined OR (95% CI)	1.21 (0.80, 1.82)		0.42
MAF< 0.5% (n = 23)			
Total carriers/total noncarriers 1[b]	11/149	41/740	
OR 1 (95% CI)[c]	1.33 (0.67, 2.65)		0.41
Total carriers/total noncarriers 2[b]	11/146	5/215	
OR 2 (95% CI)[c]	3.24 (1.10, 9.52)		**0.04**
Combined OR (95% CI)[c]	1.77 (0.97, 3.08)		**0.05**

Abbreviations: CI, confidence interval; OR, odds ratio.
Number in bold: $P \leqslant 0.05$.
Italics: typed in a subset of 227 controls.
[a]Number of individuals with variant/total number of individuals typed.
[b]Number of noncarriers corresponds to the harmonic mean of individuals without the rare variant for each variant typed.
[c]OR 1 was calculated pooling all variants typed in 866 controls, whereas OR 2 was calculated using a subset of 227 controls. Combined OR is the OR that results from combining OR 1 and OR 2 using a Mantel–Haenszel test.

seem fairly similar, there were several disagreements between programs with respect to the prediction of particular variants (Table 2). Among the rare variants with MAF <1%, there were 9 probably damaging, 5 possibly damaging and 14 benign, or 13 disease and 15 neutral variants. When only those variants with MAF <0.5% were considered, the ratio of damaging (probably + possibly)/disease to benign/neutral variants increased from ~50 to 60%. All of the low-frequency variants, in contrast, were predicted to be benign/neutral. In addition to the missense variants, there was one synonymous (SMAD4-1) and one deleterious coding (CHEK2-1) rare variants and five non-coding variants (one rare and one common variant in the promoter and two common intronic variants in *CDH1*, and one low-frequency variant in the 3′ untranslated region of *APC*). Promoter variant CDH1-1 is predicted to eliminate a S8 transcription factor-binding site, whereas the CDH1-2 promoter variant A allele has been found to decrease transcriptional efficiency by 68% with respect to the C allele, also probably by altering transcription factor-binding sites.[31] Using the programs FastSNP and F-SNP, a variant in *CDH1* intron 1 was determined to potentially affect a splicing site, whereas a variant in intron 4 of the same gene showed a low risk of being an intronic enhancer (Table 2).

DISCUSSION

We have examined 55 rare variants, 7 low-frequency variants and 8 polymorphisms in a sample of UK CRC and multiple adenoma cases and controls. Two of the four rare variants that were individually significantly associated with disease (that is, MLH1-1 and CTNNB1-1) had already been identified in the same set of individuals with multiple polyps, although not then found to be individually significant.[7] In this study, we showed that these *MLH1* and *CTNNB1* variants were not present in a different and much larger UK control population, which explains the present case–control significant difference, and were also absent from a sample of early-onset CRC UK patients. Given that these two variants have not been found in our set of French patients, and that having ~30% fewer French cases may not in itself fully explain the UK–French apparent difference, they may represent UK founder effects, as previously suggested.[4] However, replication of our findings in another UK sample of multiple adenoma patients, as well as functional studies, are necessary to establish their importance as CRC risk factors. The remaining two individually significant variants (EXO1-12 and BRCA2-37) have not been associated with CRC before. BRCA2-37 was classified as not clinically significant by the Breast Cancer Information Core database (in early 2010) and predicted to be benign by PolyPhen-2, yet it was recently found to be overrepresented among subjects with familial prostate cancer,[32] and SNP&GO considered it to be disease-associated. As mentioned above, the fact that it was not found among French patients could indicate a restricted distribution of this variant. Variant EPHB2-3, which was detected in a Finnish individual with rectal and prostate cancer in an earlier study,[33] showed a nearly significant result. All associated variants code for nonsynonymous amino-acid changes. However, CTNNB1-1, BRCA2-37 and EPHB2-3 were predicted to be benign by PolyPhen-2, whereas MLH1-1 and EXO1-12 were considered probably damaging. SNPs&GO, on the other hand, predicted CTNNB1-1 and BRCA2-37 to be disease-associated and the remaining three variants to be neutral. Recently, SNPs&GO has been found to be more accurate than PolyPhen and other similar programs.[34] However, even though it identifies CTNNB1-1 and BRCA2-37 as potentially pathogenic, it misses MLH1-1 and EXO1-12. Also, APC-11, demonstrably pathogenic,[5] was not identified as such by any of these computational methods. These discrepancies indicate that the use of *in silico* methods to evaluate the effects of nonsynonymous rare variants is not yet sufficiently reliable to be confident of their predictions. This is especially important when using them to predict which variants to focus on.

Table 4 Rare variant counts in UK multiple adenoma, early-onset CRC and control subjects for variants with MAF <0.5% in controls

Rare variant	*Multiple adenomas*[a]	*Early onset*[a]	*Controls*[a]	P-*value*[b]	P-*value*[c]
EPHB2-1	1/107	0/42	1/778	0.23	1.00
EPHB2-3	1/103	0/42	2/746	0.32	1.00
EPHB2-4	0/107	1/42	1/775	1.00	0.10
EPHB2-7	1/107	0/42	0/224	0.32	1.00
EXO1-2	0/108	0/42	1/751	1.00	1.00
EXO1-4	1/91	0/34	0/226	0.29	1.00
MLH1-1	1/105	0/42	0/740	0.12	1.00
CTNNB1-1	1/106	0/68	0/702	0.13	1.00
APC-7	0/107	0/42	4/743	1.00	1.00
APC-16	0/108	0/42	3/729	1.00	1.00
BRCA2-8	0/105	0/42	4/748	1.00	1.00
BRCA2-27	0/101	1/41	6/740	1.00	**0.05**
BRCA2-35	1/107	0/42	0/224	0.32	1.00
BRCA2-37	4/106	0/42	5/744	**0.02**	1.00
MLH3-1	1/102	3/42	2/225	1.00	**0.03**
CDH1-1	0/108	0/42	2/748	1.00	1.00
CDH1-8	0/109	0/42	1/742	1.00	1.00
CDH1-10	0/109	0/41	2/731	1.00	1.00
BRCA1-8	0/109	0/42	1/744	1.00	1.00
BRCA1-16	0/107	0/42	1/734	1.00	1.00
BRCA1-28	1/107	0/42	2/225	1.00	1.00
SMAD4-1	1/106	0/42	7/739	1.00	1.00
CHEK2-1	2/111	1/67	1/179	0.56	0.47
Adenomas vs controls					
Total carriers/total noncarriers 1[d]	9/106		41/740		
OR 1 (95% CI)[e]	1.53 (0.72, 3.24)			0.26	
Total carriers/total noncarriers 2[d]	7/103		5/215		
OR 2 (95% CI)[e]	2.92 (0.91, 9.43)			0.06	
Combined OR (95% CI)	1.87 (0.98, 3.48)			**0.05**	
Early onset vs controls					
Total carriers/total noncarriers 1[d]		2/43	41/740		
OR 1 (95% CI)[e]		0.84 (0.20, 3.59)		0.81	
Total carriers/total noncarriers 2[d]		4/42	5/215		
OR 2 (95% CI)[e]		4.10 (1.06, 15.89)		**0.03**	
Combined OR (95% CI)		1.72 (0.73, 5.27)		0.25	

Abbreviations: CI, confidence interval; OR, odds ratio.
Number in bold: $P \leqslant 0.05$.
[a]Number of individuals with variant/total number of individuals typed.
[b]*P*-value for the comparison between multiple adenoma patients and controls.
[c]*P*-value for the comparison between early-onset CRC patients and controls.
[d]Number of noncarriers corresponds to the harmonic mean of individuals without the rare variant for each variant typed.
[e]OR 1 was calculated using the larger set of controls, whereas OR 2 was calculated using a subset of 227 controls.

Table 5 Variants genotyped in French patients

Id	*Variant*	*Major/ minor allele*	*French cases*[a] *(N= 131)*	*UK cases*[a] *(N= 182)*	*Class*
EPHB2-1	R80H	G/A	0.000	0.003	Rare variant
EPHB2-4	R569W	C/T	0.000	0.003	Rare variant
EPHB2-7	M883V	A/G	0.000	0.003	Rare variant
EXO1-4	L410R	T/G	0.000	0.004	Rare variant
EXO1-10	G759E	G/A	0.015	0.003	Rare variant
MSH2-8	E808X	G/T	0.000	0.000	Rare variant
CTNNB1-1	N287S	A/G	0.000	0.003	Rare variant
APC-10	I1307K	T/A	0.000	0.000	Rare variant
APC-11	E1317Q	G/C	0.004	0.011	Rare variant
APC-17	S2621C	C/G	0.008	0.003	Low freq variant
APC-20	8636 C/A	C/A	0.051	0.024	Low freq variant
PMS2-2	T597S	A/T	0.008	0.034	Low freq variant
BRCA2-7	N372H	T/G	0.273	0.320	Polymorphism
BRCA2-35	D2665G	A/G	0.000	0.003	Rare variant
BRCA2-37	V2728I	G/A	0.000	0.014	Rare variant
BRCA2-48	P3194Q	C/A	0.000	0.000	Rare variant
MLH3-1	V741F	G/T	0.016	0.014	Rare variant
MLH3-2	M809V	A/G	0.012	0.003	Rare variant
MLH3-5	S845G	A/G	0.042	0.014	Low freq variant
CDH1-2	−284	C/A	0.315	0.246	Polymorphism
BRCA1-28	S1512I	G/A	0.004	0.003	Rare variant
CHEK2-1	1100 delC	C/del	0.000	0.008	Rare variant

[a]Minor allele frequency is shown.

The grouping of all rare variants in the association analysis (23 or 8, depending on the control set used) yielded a combined OR of ~1.2, which suggested that there was no strong evidence of an effect on CRC. However, pooling all variants with a MAF <0.5% considerably bolstered the association, taking the OR to ~1.8. Notably, even though several of the variants included in the analysis are, on the basis of the *in silico* analysis and the examination of other parameters of pathogenicity,[35] considered to be benign, neutral, not clinically significant or of unknown significance, there is nevertheless an elevated risk from their combined action. The conclusion is that these variants may well be pathologically relevant, but that the *in silico* approaches are not yet adequate to detect this. The low-frequency variants (MAF between 1 and 5%) do not appear to influence susceptibility to CRC, as we described earlier for *CCND1*.[36] It is clear that further research is needed to evaluate more fully the role of low-frequency variants in cancer.[37] Defining rare variants using a threshold based on the combined set of cases and controls, as compared with just the controls, only altered the classification of three variants in our study and so did not appreciably affect the results. This was to be expected because we had a substantially larger number of controls than cases.

Extensively studied common variants MTHFR A222V (rs1801133) and TP53 R72P (rs1042522) did not show significant frequency differences between cases and controls. Conversely, CDH1-284C/A (rs16260) exhibited a lower frequency of the A allele in patients than in controls, revealing weak statistical evidence of a protective effect of this polymorphism on colorectal disease (0.25 vs 0.31, $P = 0.07$). This is in agreement with previous findings on CRC where the C allele increases risk,[15,16] whereas in gastric, prostate and breast cancer, the

Table 6 Rare variant counts in UK and French patients

Rare variant id	*UK cases*[a]	*French cases*[a]
EPHB2-1	1/149	0/130
EPHB2-4	1/149	0/124
EPHB2-7	1/149	0/129
EXO1-4	1/125	0/130
EXO1-10	1/145	4/130
CTNNB1-1	1/174	0/130
APC-11	4/176	1/131
BRCA2-35	1/149	0/130
BRCA2-37	4/148	0/131
MLH3-2	1/148	3/129
MLH3-1	4/144	4/127
BRCA1-28	1/149	1/129
CHEK2-1	3/178	0/131
MAF <1% (n = 13)		
Total carriers/total noncarriers[b]	24/149	13/128
OR (95% CI)	1.59 (0.71, 3.02)	0.20
MAF <0.5% (n = 10)		
Total carriers/total noncarriers[b]	18/148	5/128
OR (95% CI)[e]	3.14 (1.13, 8.69)	**0.02**

Abbreviations: CI, confidence interval; OR, odds ratio.
Number in bold: $P \leqslant 0.05$.
[a]Number of individuals with variant/total number of individuals typed.
[b]Number of noncarriers corresponds to the harmonic mean of individuals without the rare variant for each variant typed.

A allele tends to be the risk allele.[38–40] However, our study is underpowered for the detection of effects from common variants.

The analysis by disease group showed that, even though the collection of rare variants in each set of patients carries a higher risk of disease, our findings are mostly driven by the effects on individuals with multiple adenomas. BRCA2-37, the rare variant with the strongest effect in this study, was, for example, found only in patients with multiple adenomas. Although the sample size for the early-onset group was limited, our results clearly suggest that the genetic influence on CRC may mostly be seen in individuals with multiple adenomas, as compared with early-onset cases. This parallels to what is found in the clear-cut familial cases of inherited CRC. The extra layer of activity needed to go from polyp to cancer leads to an additional amount of variation that may be 'less genetically determined' and so obscure the underlying genetic susceptibility due to the multiple adenomas. There were, however, no significant differences in carrier frequencies between the two groups of patients, despite the fact that over half of the variants with MAF <1% were found only in the multiple adenoma group. This, again, is probably due to the relatively smaller size of the early-onset group of patients. However, the allelic frequencies of two missense variants (one rare and one low frequency) in *MLH3* and one intronic common variant in *CDH1* differed significantly between multiple adenoma and early-onset CRC cases, with the latter exhibiting higher frequencies of these variants. This suggests that different sets of rare variants are quite likely to be involved in different pathologies, but to detect their effect would require larger numbers of patients than we were able to study.

The association *P*-values reported in this study have not been corrected for multiple hypotheses testing. Taking into account the number of variants analyzed, a Bonferroni correction would take the significance threshold to 0.001. Nonetheless, we believe that because there is an *a priori* case for each candidate variant to be potentially functional, such a correction would be unsuitably stringent. The lack of French controls precluded a similar association study from being carried out with French samples as was done for the UK samples. Using UK controls would be inappropriate because of population stratification within Europe, especially for analysis of rare variants as they are likely to be population specific. The presence of such founder effects is appreciably suggested by the fact that the variants MLH1-1 and CTNNB1-1, which are very clearly associated with multiple adenomas in UK cases, were not found in the French multiple adenoma cases. Further analysis of such differences requires larger numbers of French cases and appropriately selected French controls. Moreover, larger cohorts, such as the EPICOLON consortium experience,[41] are needed to confirm these preliminary results and meta-analysis should be performed to ensure the pathogenic effects of variants described in the present work.

In summary, because rare variants appear to be associated with higher ORs than common variants, a relatively small study like ours can uncover the effects of candidate variants with low population frequencies on complex diseases such as, in this case, CRC. We have also shown that variants with frequencies <0.5% appear to have the biggest effects regardless of the *in silico* prediction of their function. The role of the individual variant BRCA2-37 (V2728I) on the development of multiple adenomatous polyps deserves further examination. In general, the multiple adenoma phenotype seems to be more susceptible to genetic influence than early-onset CRC, but a larger early-onset patient sample would be necessary to confirm this finding. We have found some differences between UK and French patients in terms of the distribution of rare variants that justify closer inspection as population stratification within Europe can lead to spurious association results.

To conclude, we have confirmed that rare variants are important risk factors in CRC and as such, should be systematically assayed alongside common variation in the search for the genetic basis of complex diseases, taking great care to match cases with appropriate controls.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We would like to thank all individuals who contributed samples for this study. We are also grateful to Shazad Asraf, Oliver Sieber and Haitao Wang for making samples available. We thank the patients who participated in VICTOR, without whose help, the study would not have been possible. The trial was enabled by an educational study grant from Merck & Co. Inc. who also supplied the study drug. VICTOR was endorsed by the UK National Cancer Research Institute (NCRI) and Cancer Research Campaign Clinical Trials Awards and Advisory Committee (CTAAC) allowing support from the UK National Cancer Research Network. The University of Birmingham was the center for randomization and for statistics initially before these moved to the Universities of Oxford and Warwick, respectively. We also thank VICTOR trial staff: A Stanley, J Stokes, F Duchesne, K Reed, S Pendlebury, R Duvvuri, I Kennedy, G Davis, J Birtwistle, L Blair, A Bange, K Bell, L Dunham, R Deacon, A Mellor, B Mikolajczyk, M Davoudianfar and A Ferch, and VICTOR laboratory support staff: K Elliot, M Presz and H Wang, Clinical Coordinator: D Rea, The VICTOR Trial Advisory Group and the VICTOR Data and Safety Monitoring Committee and all nurses and medical staff who participated in this trial. CB was Cancer Research UK Julia Bodmer Fellow. This work was made possible by a Cancer Research UK program grant, a Wellcome Trust grant and a Serroussi Foundation grant to WFB.

URLs:

- People of the British Isles study (PoBI): www.peopleofthebritishisles.org
- Breast Cancer Information Core (BIC) dataset: http://research.nhgri.nih.gov/bic/index.shtml
- PolyPhen v.2: http://genetics.bwh.harvard.edu/pph2
- PLINK: http://pngu.mgh.harvard.edu/~purcell/plink/
- SNPs&GO: http://snps-and-go.biocomp.unibo.it/snps-and-go/
- FastSNP: http://fastsnp.ibms.sinica.edu.tw/pages/input_CandidateGeneSearch.jsp
- F-SNP: http://compbio.cs.queensu.ca/F-SNP

Author contributions: CB, BW and WB designed research; JHL, CB, BW, EJ and ST performed research; CC, TD, KH, AB, WFB, RM, DK and YP collected samples; JHL and CB analyzed data; and JHL, CB and WFB wrote the paper.

1 Fearnhead, N. S., Wilding, J. L. & Bodmer, W. F. Genetics of colorectal cancer: hereditary aspects and overview of colorectal tumorigenesis. *Br. Med. Bull.* **64,** 27–43 (2002).
2 Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J., Ellis, A., Gorman, P. *et al.* Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature* **328,** 614–616 (1987).
3 Al-Tassan, N., Chmiel, N. H., Maynard, J., Fleming, N., Livingston, A. L., Williams, G. T. *et al.* Inherited variants of MYH associated with somatic G:C–>T:A mutations in colorectal tumors. *Nat. Genet.* **30,** 227–232 (2002).
4 Bodmer, W. & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat. Genet.* **40,** 695–701 (2008).
5 Frayling, I. M., Beck, N. E., Ilyas, M., Dove-Edwin, I., Goodman, P., Pack, K. *et al.* The APC variants I1307K and E1317Q are associated with colorectal tumors, but not always with a family history. *Proc. Natl Acad. Sci. USA* **95,** 10722–10727 (1998).
6 Bodmer, W. Familial adenomatous polyposis (FAP) and its gene, APC. *Cytogenet. Cell Genet.* **86,** 99–104 (1999).
7 Fearnhead, N. S., Wilding, J. L., Winney, B., Tonks, S., Bartlett, S., Bicknell, D. C. *et al.* Multiple rare variants in different genes account for multifactorial inherited susceptibility to colorectal adenomas. *Proc. Natl Acad. Sci. USA* **101,** 15992–15997 (2004).
8 Allan, J. M., Shorto, J., Adlard, J., Bury, J., Coggins, R., George, R. *et al.* MLH1 -93G>A promoter polymorphism and risk of mismatch repair deficient colorectal cancer. *Int. J. Cancer* **123,** 2456–2459 (2008).
9 Tulupova, E., Kumar, R., Hanova, M., Slyskova, J., Pardini, B., Polakova, V. *et al.* Do polymorphisms and haplotypes of mismatch repair genes modulate risk of sporadic colorectal cancer? *Mutat. Res.* **648,** 40–45 (2008).
10 Wasielewski, M., Vasen, H., Wijnen, J., Hooning, M., Dooijes, D., Tops, C. *et al.* CHEK2 1100delC is a susceptibility allele for HNPCC-related colorectal cancer. *Clin. Cancer Res.* **14,** 4989–4994 (2008).
11 Winney, B., Boumertit, A., Day, T., Davison, D., Echeta, C., Evseeva, I. *et al.* People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *Eur. J. Hum. Genet.* **20,** 203–210 (2012).
12 Lamlum, H., Al Tassan, N., Jaeger, E., Frayling, I., Sieber, O., Reza, F. B. *et al.* Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q. *Hum. Mol. Genet.* **9,** 2215–2221 (2000).
13 Canard, G., Lefevre, J. H., Colas, C., Coulet, F., Svrcek, M., Lascols, O. *et al.* Screening for lynch syndrome in colorectal cancer: are we doing enough? *Ann. Surg. Oncol.* **19,** 809–816 (2012).
14 Lefevre, J. H., Rodrigue, C. M., Mourra, N., Bennis, M., Flejou, J. F., Parc, R. *et al.* Implication of MYH in colorectal polyposis. *Ann. Surg.* **244,** 874–880 (2006).
15 Grunhage, F., Jungck, M., Lamberti, C., Berg, C., Becker, U., Schulte-Witte, H. *et al.* Association of familial colorectal cancer with variants in the E-cadherin (CDH1) and cyclin D1 (CCND1) genes. *Int. J. Colorectal Dis.* **23,** 147–154 (2008).
16 Pittman, A. M., Twiss, P., Broderick, P., Lubbe, S., Chandler, I., Penegar, S. *et al.* The CDH1-160C>A polymorphism is a risk factor for colorectal cancer. *Int. J. Cancer* **125,** 1622–1625 (2009).
17 Levine, A. J., Figueiredo, J. C., Lee, W., Poynter, J. N., Conti, D., Duggan, D. J. *et al.* Genetic variability in the MTHFR gene and colorectal cancer risk using the colorectal cancer family registry. *Cancer Epidemiol. Biomarkers Prev.* **19,** 89–100 (2010).
18 Le Marchand, L., Wilkens, L. R., Kolonel, L. N. & Henderson, B. E. The MTHFR C677T polymorphism and colorectal cancer: the multiethnic cohort study. *Cancer Epidemiol. Biomarkers Prev.* **14,** 1198–1203 (2005).
19 Polakova, V., Pardini, B., Naccarati, A., Landi, S., Slyskova, J., Novotny, J. *et al.* Genotype and haplotype analysis of cell cycle genes in sporadic colorectal cancer in the Czech Republic. *Hum. Mutat.* **30,** 661–668 (2009).
20 Ahn, S. J., Costa, J. & Emanuel, J. R. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. *Nucleic Acids Res.* **24,** 2623–2625 (1996).
21 Sauer, S. & Gut, I. G. Genotyping single-nucleotide polymorphisms by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **782,** 73–87 (2002).
22 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
23 Svejgaard, A., Jersild, C., Nielsen, L. S. & Bodmer, W. F. HL-A antigens and disease. Statistical and genetical considerations. *Tissue Antigen* **4,** 95–105 (1974).
24 Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl Cancer Inst.* **22,** 719–748 (1959).
25 Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7,** 248–249 (2010).
26 Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L. & Casadio, R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum. Mutat.* **30,** 1237–1244 (2009).
27 Yuan, H. Y., Chiou, J. J., Tseng, W. H., Liu, C. H., Liu, C. K., Lin, Y. J. *et al.* FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.* **34,** W635–W641 (2006).
28 Lee, P. H. & Shatkay, H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res.* **36,** D820–D824 (2008).
29 Lemire, M. Defining rare variants by their frequencies in controls may increase type I error. *Nat. Genet.* **43,** 391–392 (2011).
30 Pearson, R. D. Bias due to selection of rare variants using frequency in controls. *Nat. Genet.* **43,** 392–393 (2011).
31 Li, L. C., Chui, R. M., Sasaki, M., Nakajima, K., Perinchery, G., Au, H. C. *et al.* A single nucleotide polymorphism in the E-cadherin gene promoter alters transcriptional activities. *Cancer Res.* **60,** 873–876 (2000).
32 Luedeke, M., Linnert, C. M., Hofer, M. D., Surowy, H. M., Rinckleb, A. E., Hoegel, J. *et al.* Predisposition for TMPRSS2-ERG fusion in prostate cancer by variants in DNA repair genes. *Cancer Epidemiol. Biomarkers Prev.* **18,** 3030–3035 (2009).
33 Kokko, A., Laiho, P., Lehtonen, R., Korja, S., Carvajal-Carmona, L. G., Järvinen, H. *et al.* EPHB2 germline variants in patients with colorectal cancer or hyperplastic polyposis. *BMC Cancer* **6,** 145 (2006).
34 Thusberg, J., Olatubosun, A. & Vihinen, M. Performance of mutation pathogenicity prediction methods on missense variants. *Hum. Mutat.* **32,** 358–368 (2011).
35 Spurdle, A. B. Clinical relevance of rare germline sequence variants in cancer genes: evolution and application of classification models. *Curr. Opin. Genet. Dev.* **20,** 315–323 (2010).
36 Bonilla, C., Lefevre, J. H., Winney, B., Johnstone, E., Tonks, S., Colas, C. *et al.* Cyclin D1 rare variants in UK multiple adenoma and early-onset colorectal cancer patients. *J. Hum. Genet.* **56,** 58–63 (2011).
37 Zhu, Q., Ge, D., Maia, J. M., Zhu, M., Petrovski, S., Dickson, S. P. *et al.* A genome-wide comparison of the functional properties of rare and common genetic variants in humans. *Am. J. Hum. Genet.* **88,** 458–468 (2011).
38 Humar, B., Graziano, F., Cascinu, S., Catalano, V., Ruzzo, A. M., Magnani, M. *et al.* Association of CDH1 haplotypes with susceptibility to sporadic diffuse gastric cancer. *Oncogene* **21,** 8192–8195 (2002).
39 Qiu, L. X., Li, R. T., Zhang, J. B., Zhong, W. Z., Bai, J. L., Liu, B. R. *et al.* The E-cadherin (CDH1)–160 C/A polymorphism and prostate cancer risk: a meta-analysis. *Eur. J. Hum. Genet.* **17,** 244–249 (2009).
40 Beeghly-Fadiel, A., Shu, X. O., Long, J., Li, C., Cai, Q., Cai, H. *et al.* Genetic polymorphisms in the MMP-7 gene and breast cancer survival. *Int. J. Cancer* **124,** 208–214 (2009).
41 Castellvi-Bel, S., Ruiz-Ponte, C., Fernandez-Rozadilla, C., Abuli, A., Munoz, J., Bessa, X. *et al.* Seeking genetic susceptibility variants for colorectal cancer: the EPICOLON consortium experience. *Mutagenesis* **27,** 153–159 (2012).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

ARTICLES CITÉS DANS LA THÈSE

Author's personal copy

Ann Surg Oncol (2012) 19:809–816
DOI 10.1245/s10434-011-2014-7

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – COLORECTAL CANCER

Screening for Lynch Syndrome in Colorectal Cancer: Are We Doing Enough?

Guillaume Canard, MD[1], Jeremie H. Lefevre, MD[1], Chrystelle Colas, MD[2], Florence Coulet, PharmD, PhD[2], Magali Svrcek, MD, PhD[3], Olivier Lascols, MD[4], Richard Hamelin, PhD[5], Conor Shields, MD[1], Alex Duval, MD, PhD[5], Jean-Francois Fléjou, MD, PhD[3], Florent Soubrier, MD, PhD[2], Emmanuel Tiret, MD[1], and Yann Parc, MD, PhD[1]

[1]Department of Digestive Surgery, Hôpital Saint-Antoine, Université Pierre et Marie Curie, Paris, France; [2]Department of Oncogenetics, Hôpital Pitié-Salpêtrière, Université Pierre et Marie Curie, Paris, France; [3]Department of Pathology, Hôpital Saint-Antoine, Université Pierre et Marie Curie, Paris, France; [4]Department of Biology and Molecular Genetics, Hôpital Saint-Antoine, Université Pierre et Marie Curie, Paris, France; [5]INSERM UMRS 938, Equipe Instabilité des microsatellites et Cancer, Université Pierre et Marie Curie, Paris, France

ABSTRACT

Purpose. The purpose of this study was to assess the efficacy of screening for the detection of Lynch syndrome (LS) in an unselected population undergoing surgery for a colorectal cancer.

Methods. A total of 1,040 patients were prospectively included between 2005 and 2009. LS screening modalities included the Bethesda criteria, immunochemistry (IHC) for MLH1, MSH2, and MSH6, and microsatellite instability (MSI) by using pentaplex markers. Promoter methylation was assessed in tumors with a loss of MLH1 expression. Gene sequencing was offered to patients with abnormal IHC or MSI status without promoter methylation.

Results. A total of 105 patients had an abnormal result: 102 (9.8%) exhibited a loss of protein on IHC and 98 (9.4%) had MSI. A discordant result was observed in 10 patients with eventual proven LS in 6 patients. Loss of MLH1 ($n = 64$) was due to promoter methylation in 43 patients (67.2%). Overall, of 62 patients with an abnormal result, 38 had genetic sequencing leading to 25 (65.8%) identified with a germ-line mutation. Loss of *MSH2* on IHC was associated with a mutation in 78.3% (18 of 23) of cases. Among the 62 patients with abnormal results, 23 (37.1%) did not meet the Bethesda criteria.

Conclusions. Strict application of the Bethesda criteria does not lead to identification of all patients with LS. IHC and MSI testing are complementary methods and should be used in association to identify potential LS patients.

© Society of Surgical Oncology 2011

First Received: 9 October 2010;
Published Online: 31 August 2011

J. H. Lefevre, MD
e-mail: jeremie.lefevre@sat.aphp.fr

Lynch syndrome (LS) is the most common type of hereditary colorectal cancer (CRC), accounting for 2–4% of all CRC.[1–4] The pathogenesis of this autosomal dominant predisposition is attributed to inactivation of the mismatch repair (MMR) system, and is caused by germline mutations in MMR genes (mostly *hMLH1* and *hMSH2*, which account for approximately 90% of families with LS and to a lesser extent, *hMSH6* and *hPMS2*).[5,6] The MMR defect leads to microsatellite instability (MSI), characterised by a change in the length of nucleotide repeat sequences within tumor DNA. Diagnosis of LS is based on the identification of germ-line MMR mutations, and has important implications for long-term follow-up and management of at-risk relatives.[7]

Because of practical and financial constraints, appropriate selection of patients with CRC for genetic testing is necessary. Both the Amsterdam criteria and the Bethesda guidelines have been used to select CRC patients for either MSI evaluation or immunohistochemical (IHC) analysis of MMR proteins.[8,9] If MSI is detected, a germ-line MMR gene mutation is sought.[7,10] By means of this screening strategy, the proportion of cases with an inherited MMR gene mutation that remained undetected is estimated to be between 10 and 25%.[1,7] The evidence on which to base a choice between IHC and MSI testing for primary screening

of LS is conflicting: the sensitivity of MSI testing is slightly better than IHC for the diagnosis of LS but the advantage of IHC is that it directs mutation analysis.[1,11,12]

Neither IHC nor MSI testing are specific for LS: about 15% of sporadic CRC exhibit an MSI phenotype, with a loss of MLH1 expression on IHC due to methylation of the promoter of *MLH1*.[13]

The aim of this study was to assess the performance in daily practice of a screening strategy for the identification of patients with LS, based on systematic IHC and MSI testing in a consecutive series of unselected patients with CRC.

METHODS

Screening Strategy of LS

Between January 2005 and August 2009, 1,040 consecutive and unselected patients underwent surgery for CRC in our department, excluding patients with a dominant polyposis syndrome or a history of inflammatory bowel disease. IHC staining for MMR proteins and MSI testing were performed on all specimens. Methylation of the promoter of *hMLH1* was sought if IHC showed a loss of MLH1. If MSI in the absence of methylation of the MLH1 promoter, or a defect of the MMR protein was detected, the patient was referred for genetic counseling to screen for mutations of the MMR genes.

Immunohistochemical Staining

IHC staining for MSH2 and MLH1 was systematically performed as described previously.[14,15] The primary antibody used for MSH2 was FE11 (Calbiochem, Cambridge, MA), that for MSH2 was G168-728 (Pharmingen, San Diego, CA). Lymphocytes and normal epithelium exhibit strong nuclear staining for MMR proteins and served as positive internal controls for staining of these proteins. IHC staining for MSH6 and PMS2 was performed in case of MLH1/MSH2 normal expression in patients younger than 60 years, or with personal and/or family history of LS, or with MSI. The primary antibody used for MSH6 was clone 44 (Becton Dickinson, Lexington, NC) and that for PMS2 was clone A16-4 (BD Pharmingen, Le Pont de Claix, France).

Analysis of MSI

For each patient, tumour DNA was extracted from formalin-fixed, paraffin-embedded sections with the QiampKit (Qiagen, Santa Clarita, CA) and was subjected to polymerase chain reaction (PCR) assay with the use of the MSI Analysis System (Promega France, Charbonnières-les-Bains, France), that is derived from a pentaplex assay (NR21, NR22, NR24, BAT25, BAT26) developed in our laboratory.[16] Because of the quasimonomorphic nature of these markers in white populations, MSI status of tumors was established without reference to matching normal DNA.[17] Tumors were scored as MSI-H if at least 3 of the 5 markers showed instability, as compared to the normal quasimonomorphic variation range of each marker.[17]

Analysis of MLH1 Promoter Methylation

Methylation status of the proximal part of the promoter region of *MLH1* was assessed by methylation-specific PCR (MSP). All DNA samples were treated chemically with sodium bisulphite according to the Zymo Research DNA-methylation Kit Subsequent methylation-specific PCRs were carried out as described.[18] Two PCRs were performed for each chemically modified DNA sample, one with primers annealing specifically to the methylated gene and the other to unmethylated DNA, with a fluorescent primer in each case. PCR products were run on an ABI 3730 DNA sequencer (Applied Biosystems, Foster City, CA). The resulting data were analyzed with the GeneMapper software, version 4.0 (Applied Biosystems). All PCRs were performed with positive controls of both methylated and unmethylated DNA. If methylation of the promoter could not be assessed because of technical failure, it was considered as a failure of methylation.

Germ-Line Mutation Analysis

DNA Isolation DNA was isolated from peripheral blood after obtaining informed consent for MMR genetic analysis. Human genomic DNA was prepared from whole blood samples with a semiautomated Extragene extractor (Genomic Industry, Archamps, France) and a DNA extraction kit (Promega France).

PCR Amplification and Sequencing of the Coding Regions of MLH1 and MSH2 All *hMLH1*, *hMSH2*, and *hMSH6* exons were amplified by PCR with primers designed with primer3 software (http://Frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Amplification was performed on a GeneAmp PCR system 9700 (Applied Biosystems). PCR products were cleaned up on a MultiScreen PCR 96-well plate (Millipore, Billerica, MA) and sequencing reactions were carried out with the BigDye Terminator v1.1 cycle sequencing kit (Applied Biosystems). The products of the sequencing reactions are cleaned up with the Sephadex G-50 (GE Healthcare, Life Sciences) in a MultiScreen-HV 96-well filter plate (Millipore), then run up on an ABI 3730 DNA sequencer (Applied Biosystems). The resulting sequence data were analyzed by SeqScape software, version 2.5 (Applied Biosystems) in comparison with the

reference sequences of *MLH1* and *MSH2* genes (accession numbers, respectively: NM_000249 and NM_000251).

Ligation-Dependent Probe Amplification (MLPA) The SALSA MLPA P003 MLH1/MSH2 probe mix and SALSA MLPA P008 kits (MRC-Holland, Holland, The Netherlands) were used according to the manufacturer's instructions to screen rearrangements of one or more exons of the *hMLH1*, *hMSH2*, and *hMSH6* genes. Fragment analysis of multiplex PCR was carried out on the ABI 3730 DNA analyser (Applied Biosystems) and results were analyzed by GeneMapper software, version 4.0 (Applied Biosystems). Normalization of samples was performed against the mean control ratios. DNA samples from unaffected individuals were used as controls in each experiment. Patient samples were compared to unaffected individuals samples.

Statistical Analysis

Clinicopathologic data were compared with *t*-test and ANOVA for quantitative data and with contingency table and a Pearson χ^2 test for proportions. Statistical analysis of the data was carried out with XLSTAT, version 2007.7. A *P* value of 0.05 was considered the threshold of statistical significance.

RESULTS

Among 1,040 patients operated on for CRC in the department of Surgery, 105 (10.1%) had an abnormal test. Results of screening in the 1,040 patients are summarized in Fig. 1. IHC was abnormal in 102 (9.8%) of the 1,040 tumors: 66 patients had a loss of MLH1 expression (64.7%), 33 a loss of MSH2 (32.3%), and 3 a loss of MSH6 only (3%). PCR showed MSI in 98 (9.4%) of the 1,040 tumors.

Methylation Analysis

Methylation of the *hMLH1* promoter was assessed in the 66 CRC tumors with loss of MLH1 expression. Overall, 43 (65%) of them had hypermethylation of the *hMLH1* promoter and were considered as sporadic cases. Eighteen (27%) of them had no methylation of the MLH1 promoter, whereas 5 patients (8%) with a methylation testing failure were eligible for MMR mutation analysis.

Genetic Counseling and Mutation Analysis

Sixty-two patients were eligible for MMR mutation analysis because of abnormal results on IHC and/or MSI testing, without methylation of the *hMLH1* promoter, and

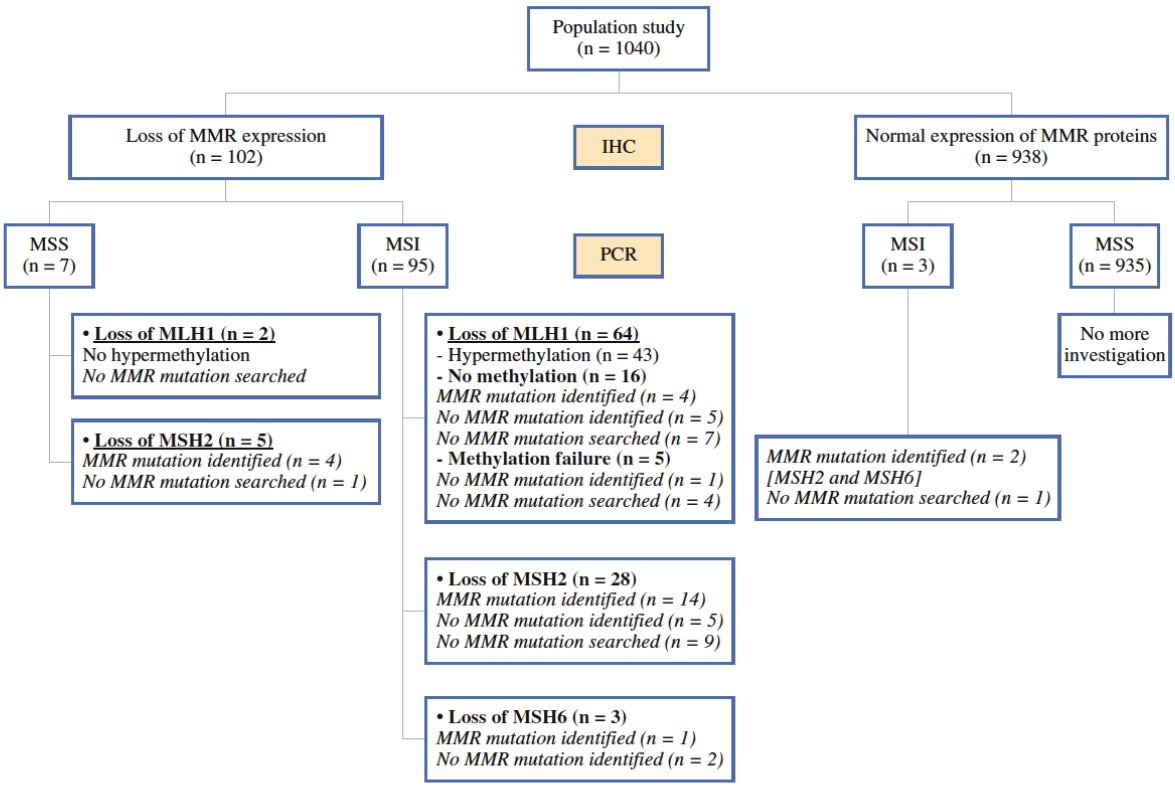


FIG. 1 Diagram of the screening strategy and main results of the study

were referred for genetic counseling. Among them, 38 (61.3%) had a genetic sequencing leading to 25 (65.8%) identified germ-line mutations (*hMLH1* = 4; *hMSH2* = 19; *hMSH6* = 2). For 13 patients, an MMR mutation was sought but none identified (MLH1 loss = 6; MSH2 loss = 5; MSH6-only loss = 2).

Twenty-four patients (38.7%) declined gene sequencing. These patients were significantly older than the other patients undergoing exploration for MMR mutations (63.2 ± 12.9 years vs. 49.4 ± 12.8 years; $P = 0.0001$) and had a tendency toward a less frequent history of CRC (11 of 24, 45.8%, vs. 25 of 38, 65.8%; $P = 0.1$).

Efficiency of Screening Test

Sensitivity of IHC and MSI testing for detecting identified LS patients were, respectively, 92% (23 of 25) and 84% (21 of 25). After exclusion of patients not tested for an MMR mutation, the positive predictive value of IHC and MSI testing were 29.1% (23 of 79) and 27.3% (21 of 77), respectively.

Concordance of Screening Tests

Ten patients had a discordant result of IHC and MSI testing (10 of 105, 9.5%) (Table 1). These patients represented 0.96% of the entire cohort. Seven patients with abnormal IHC staining (7 of 102, 6.9%) were found to be MSS. In most cases, this was due to the very low tumor content in the sample because none of the samples were tumor enriched by microdissection before MSI analysis. Three patients with MSI-H phenotype (3 of 98, 3.1%) had a normal expression of the four MMR proteins on IHC.

Among these 10 patients with discordant results, LS was ultimately proven in 6 cases (60%): 2 patients with MSI phenotype and normal expression of MMR proteins on IHC had mutation in MSH2 and MSH6. Surprisingly, none was missense mutation: one was a frameshift and the other skip exon 5 and is usually associated with loss of expression in IHC; 4 patients with MSS phenotype and with loss of MSH2 on IHC. The remaining 4 patients did not have genetic testing.

Clinical Details of Patients with Abnormal Results

Among the 105 patients with abnormal results on IHC and/or PCR, 43 were finally considered as sporadic, because of methylation of the MLH1 promoter, whereas 62 were suspicious for LS: 23 because of the loss of MLH1 on IHC with no methylation, or with methylation failure; 36 because of loss of MSH2 or MSH6 detected on IHC; and 3 because of MSI positivity with a normal expression of MMR proteins on IHC. Clinical characteristics of sporadic cases and possible LS are shown in Table 2. Patients with possible LS were significantly younger than sporadic cases (54.8 ± 14.5 vs. 77.3 ± 13.5; $P < 0.0001$). Right-sided CRC were predominant in sporadic cases group, unlike possible LS group (36 of 43, 83.7%, vs. 29 of 62, 46.8%, $P < 0.0001$). Moreover, all patients with rectal cancer were classified in the possible LS group (0 of 43, 0%, vs. 17 of 62, 27.4%, $P = 0.0002$). Among the 15 patients with a rectal tumor, only 3 had a loss of expression of MLH1 (20%) and all of them had an methylation failure test. The remaining patients with a rectal tumor had a loss of expression of MSH2 ($n = 9$; 60%) or MSH6 ($n = 2$; 13%). The last patient had no loss of protein expression, but a MSI phenotype. Finally, 3 patients had a MSS phenotype and a loss of protein expression and 2 of them had a neoadjuvant radiochemotherapy treatment. Tumors from sporadic cases were significantly more invasive than those

TABLE 1 Characteristics of the 10 patients with discordant results

Patient no.	Sex	Age (years)	CRC site	Amsterdam[a]	Bethesda[b]	Radiotherapy[c]	IHC	MSI testing	Mutation gene
1	M	51	Rectum	N	Y	Y	MLH1	MSS	NS
2	F	76	Ascending	N	N	N	MLH1	MSS	NS
3	M	66	Rectum	N	Y	Y	MSH2	MSS	NS
4	M	16	Rectum	N	Y	Y	MSH2	MSS	MSH2
5	M	53	Ascending	N	Y	N	MSH2	MSS	MSH2
6	M	46	Ascending	Y	Y	N	MSH2	MSS	MSH2
7	F	49	Rectum	Y	Y	Y	MSH2	MSS	MSH2
8	M	60	Rectum	N	N	N	Normal	MSI	MSH6
9	M	53	Ascending	N	N	N	Normal	MSI	MSH2
10	M	62	Ascending	N	N	N	Normal	MSI	NS

NS not searched

[a] Yes if patient meets Amsterdam II criteria

[b] Yes if patient meets revised Bethesda guidelines

[c] Yes if patients had preoperative radiotherapy

TABLE 2 Characteristics of the 105 patients with abnormal results on IHC and/or PCR

Characteristic	Sporadic cases ($n = 43$)	Possible LS ($n = 62$)
Mean age, y (min–max)	77 (21–98)	55 (16–86)
Sex (M/F)	14/29	41/21
CRC site		
Ascending	36 (83.7%)	29 (46.8%)
Transverse	5 (11.6%)	6 (9.7%)
Descending	1 (2.3%)	10 (16.7%)
Rectum	0 (0)	15 (24.2%)
Multifocal	1 (2.3%)	2 (3.2%)
pT		
1–2	3 (6.9%)	18 (29.0%)
3	25 (58.1%)	36 (58.1%)
4	15 (34.9%)	8 (12.9%)
pN		
N0	30 (69.8%)	48 (77.4%)
N+	13 (30.2%)	14 (22.6%)
pM		
M0	42 (97.7%)	61 (98.4%)
M+	1 (2.3%)	1 (1.6%)
Fulfilment of Amsterdam criteria	1 (2.3%)	17 (27.4%)
Fulfilment of revised Bethesda criteria	10 (23.2%)	39 (62.9%)

from patients with possible LS (pT3–4: 40 of 43, 93%, vs. 44 of 62, 71%, $P = 0.005$) but no difference was observed regarding the N score ($P = 0.512$). Patients with possible LS fulfilled more frequently the Amsterdam criteria (27.4 vs. 2.3%; $P = 0.002$) and the revised Bethesda criteria (62.9 vs. 23.2%; $P < 0.0001$) than patients with MSI sporadic tumors.

Clinical Details of Patients with Proven LS

The overall prevalence of LS identified in this series was 2.4% (25 of 1040). This represents the minimal prevalence as 24 patients with possible LS were not tested for germline mutation. The clinical and molecular characteristics of these 25 patients are shown in Table 3. The mean age at diagnosis was 48.6 ± 11.6 years (16–69), and 40% ($n = 10$) were older than 50 at the time of diagnosis. Five patients (20%) had no family history of LS spectrum tumors. Eleven patients (44%) fulfilled the Amsterdam II criteria and 22 patients (88%) fulfilled at least one of the revised Bethesda criteria. The three patients who did not fulfil the revised Bethesda criteria were diagnosed with a MSI-H phenotype despite a normal IHC staining ($n = 2$).

DISCUSSION

Guidelines for the identification of LS in patients with CRC recommend selection for prescreening tests, such as IHC and MSI testing, based on the revised Bethesda criteria.[7,9] These include parameters such as age at diagnosis, personal and family history and histological features of MSI. Some studies have already demonstrated that using this strategy, between 10 and 25% of the mutation carriers remain undetected.[1,7,10] Since 2005, IHC and MSI testing are performed systematically on all CRCs in our institution, independent of the revised Bethesda criteria. The aim of this study was to evaluate the efficacy of this strategy, especially in reducing the number of MMR mutation carriers missed by the current recommended strategy.

Overall, 10.1% of patients in our series had MSI-H and/or MMR defective tumors, below the percentage of 15–20% that is generally found in the literature.[1,19] It should be noted that our group is accredited for both IHC and MSI analysis.[14,15,20] This discordance is probably due to the fact that our surgical department specializes in rectal cancer. As a consequence, rectal cancers were over-represented in our series, tumors which are known to exhibit MSI uncommonly (about 2% in the literature). 2.4% were identified as presenting with LS, which is concordant with the literature, but it is obvious that this percentage is underestimated because the number of possible LS patients in which germ-line MMR mutations were not found (13 cases) or even looked for (24 cases).[1–4] We showed that IHC and MSI-testing were quite similar with a high sensitivity for LS (92 and 84%, respectively), as has been previously demonstrated.[1,11,21,22] The pentaplex markers used in our study for MSI testing are an alternative to the reference panel recommended by the Bethesda guidelines, and have the advantage of a 100% sensitivity and specificity without the need to match normal DNA. Interestingly, 10 patients had discordant results between IHC and MSI testing, leading to a diagnosis of LS in 6 patients. This rate is lower than reported by Hampel et al., who found 21 discordant results between MSI and IHC in a cohort of 500 patients.[23] In most cases, MSI negative results with an MMR deficiency are found in rectal cancers after radiotherapy in which tumor content is below the MSI detection level. Most authors recommend the use of IHC as the first step for screening as this technique is more widely available than MSI testing and allows the identification of the probable mutated gene. In a recent report, screening for LS by MSI-testing in all tumors, and then by IHC for MSI-H tumors was recommended.[7,10] The adoption of this strategy would have resulted in the nondetection of 16% of patients with proven LS in this series. Therefore, we suggest that both modalities be used used to avoid this scenario.

TABLE 3 Characteristics of 24 patients with LS identified

Patient no.	Sex	Age	CRC site	Amsterdam[a]	Bethesda[b]	IHC	MSI testing	Gene	Nucleotide change	Protein change
1	F	45	Descending	N	Y	MLH1 negative	MSI	*MLH1*	c.753 C>G	p.Tyr251X
2	M	33	Rectum	Y	Y	MLH1 negative	MSI	*MLH1*	c.790+5G>T	splice defect
3	M	63	Ascending	Y	Y	MLH1 negative	MSI	*MLH1*	c.676C>T	p.Arg226X
4	M	46	Ascending	N	Y	MLH1 negative	MSI	*MLH1*	c.2T>C	p.Met1?
4	M	63	Descending	Y	Y	MSH2 negative	MSI	*MSH2*	c.2541del	p.Ala848ProfsX44
6	F	69	Transverse	N	Y	MSH2 negative	MSI	*MSH2*	c.942+3A>T	splice defect
7	F	56	Rectum	Y	Y	MSH2 negative	MSI	*MSH2*	c.942+3A>T	splice defect
8	M	36	Descending	N	Y	MSH2 negative	MSI	*MSH2*	c.1077−1G>C	splice defect
9	M	47	Ascending	Y	Y	MSH2 negative	MSI	*MSH2*	c.2075 G>T	p.Gly692Val
10	M	56	Transverse	N	Y	MSH2 negative	MSI	*MSH2*	c.749del	p.Gly250GlufsX4
11	F	50	Rectum	N	Y	MSH2 negative	MSI	*MSH2*	c.1351 C>T	p.Gln451X
12	M	42	Descending	N	Y	MSH2 negative	MSI	*MSH2*	c.1-?_1386+? (delex1–8)	-
13	M	47	Transverse	N	Y	MSH2 negative	MSI	*MSH2*	c.2472–2473delins GAT	p.Ser825MetfsX11
14	M	37	Ascending	Y	Y	MSH2 negative	MSI	*MSH2*	c.725dup	p.Asn242LysfsX14
15	M	56	Rectum	N	N	MSH2 negative	MSI	*MSH2*	c.2662del	p.Leu888CysfsX4
16	M	47	Ascending	Y	Y	MSH2 negative	MSI	*MSH2*	c.1786_1788del	p.Asn596del
17	M	42	Ascending	Y	Y	MSH2 negative	MSI	*MSH2*	c.942+3A>T	splice defect
18	M	34	Descending	Y	Y	MSH2 negative	MSI	*MSH2*	TACSTD1 deletion	promoter methylation
19	M	16	Rectum	N	Y	MSH2 negative	MSS	*MSH2*	c.2038 C>T	p.Arg680X
20	M	53	Ascending	N	Y	MSH2 negative	MSS	*MSH2*	c.367-?_1386+? (delex3–8)	-
21	F	49	Rectum	Y	Y	MSH2 negative	MSS	*MSH2*	c.942+3A>T	splice defect
22	M	46	Ascending	Y	Y	MSH2 negative	MSS	*MSH2*	c.2089T>C	p.Cys697Arg
23	M	53	Ascending	N	N	Normal	MSI	*MSH2*	c.942+3A>T	splice defect
24	M	60	Rectum	N	N	Normal	MSI	*MSH6*	c.3264–3270del	p.Glu1090LysfsX23
25	F	62	Ascending	N	Y	MSH6 negative	MSI	*MSH6*	c.562dup	p.Ile188AsnfsX2

[a] Yes if patients meet Amsterdam II criteria

[b] Yes if patients meet revised Bethesda guidelines

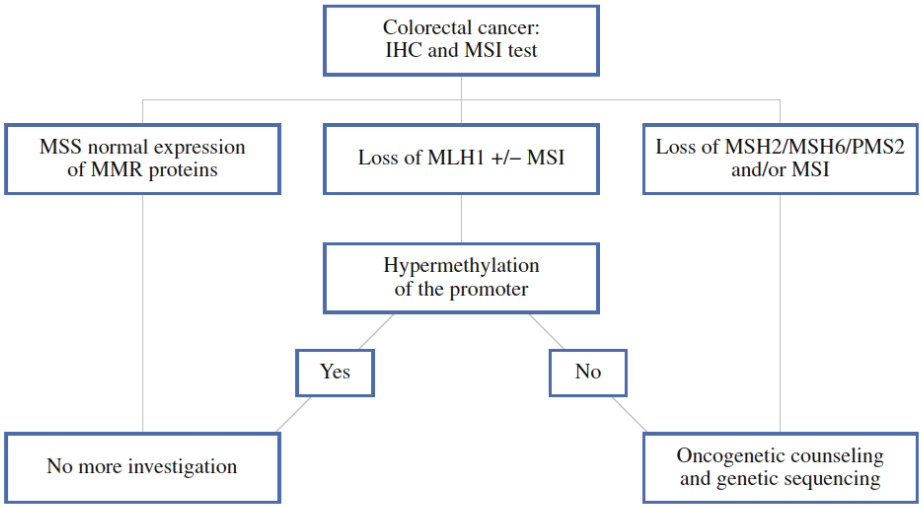


FIG. 2 Systematic strategy for LS identification in sporadic CRC

In our study, 12% of the patients with proven LS and 37.1% of the patients with possible LS did not fulfil the revised Bethesda criteria; moreover, a restriction of screening to patients younger than 50 years would have missed about half of patients with proven LS. Family history of LS-associated cancers is often underreported by patients, and recent recommendations underline this problem.[24,25] Even if LS represents only a small cohort, when diagnosed it allows the early identification of at-risk relatives with appropriate intensive surveillance. We believe that selection of patients with CRC for MMR mutation identification must be based on prescreening tests, such as IHC or MSI-testing, performed systematically on all CRCs. The cost-effectiveness of this strategy has to be evaluated in routine practice but it seems to not increase the number of genetic analyses: in our series, about 6% of patients met the criteria. IHC for MLH1/MSH2 loss of expression costs €14.05 and PCR for MSI testing about €80.

In our series, we found two possible predictive factors associated with LS. Firstly, loss of MSH2 on IHC is strongly associated with LS (78.3%), as we have previously described.[15] Secondly, no patient with a rectal cancer and a loss of MLH1 ($n = 4$) had methylation of the promoter, and therefore, they were all considered as possible LS: among them, an MLH1 mutation was sought and found in one patient. Moreover, among the 15 patients with a rectal cancer and an abnormal result (IHC and/or MSI-H), 8 were tested and all were found to have a germ-line mutation of a MMR gene.

The prevalence of LS in this series is low. Limitations of this study include the fact that genetic sequencing was not performed in 24 patients with possible LS. The fact that about 39% of the patients were not subjected to genetic counseling for MMR mutations may be explained by the age at diagnosis and the absence of a personal or family history of CRC: these patients were older and were frequently without a family history of CRC. As a reduction of up to 60% in cancer incidence and related death is observed in known at-risk relatives, patient education about the merits of genetic testing may improve uptake.[26] Sequencing and/or deletion detection techniques also need improvement, as MMR germ-line mutations were not detected in 13 cases in which MSI testing and IHC analysis were strongly indicative of LS. *MSH2* promoter methylation due to *TACSTD1* germ-line deletion, mosaicism, or somatic mutation may explain some of the unsuccessful searches.[27] Finally, we did not search for *PMS2* mutation.

In conclusion, we recommend screening for LS in all patients with CRC, independent of the revised Bethesda criteria (Fig. 2). In the present study, 6% of cases are probably LS, although this was genetically proven in only 2.4%. Strict application of the Bethesda criteria would have resulted in the non-identification of 37% of possible cases. A further refinement of the clinical criteria may reduce this number. IHC and MSI testing are complementary with a discordant result rate of up to 10%, and should both be used: IHC alone misses 8%, and PCR misses 16% of LS. Patients must be advised on the importance of genetic counseling, especially elderly patients or patients with no history of CRC.

REFERENCES

1. Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *N Engl J Med.* 2005;352:1851–60.
2. Cunningham JM, Kim CY, Christensen ER, Tester DJ, Parc Y, Burgart LJ, et al. The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am J Hum Genet.* 2001;69:780–90.
3. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med.* 2003;348:919–32.

4. 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;89: 1758–62.
5. de la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev Cancer.* 2004;4:769–80.
6. Renkonen E, Zhang Y, Lohi H, Salovaara R, Abdel-Rahman WM, Nilbert M, et al. Altered expression of *MLH1, MSH2,* and *MSH6* in predisposition to hereditary nonpolyposis colorectal cancer. *J Clin Oncol.* 2003;21:3629–37.
7. Vasen HF, Moslein G, Alonso A, Bernstein I, Bertario L, Blanco I, et al. Guidelines for the clinical management of Lynch syndrome (hereditary non-polyposis cancer). *J Med Genet.* 2007; 44:353–62.
8. 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;58:5248–57.
9. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst.* 2004;96:261–8.
10. Julie C, Tresallet C, Brouquet A, Vallot C, Zimmermann U, Mitry E, et al. Identification in daily practice of patients with Lynch syndrome (hereditary nonpolyposis colorectal cancer): revised Bethesda guidelines–based approach versus molecular screening. *Am J Gastroenterol.* 2008;103:2825–35.
11. Pinol V, Castells A, Andreu M, Castellvi-Bel S, Alenda C, Llor X, et al. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA.* 2005;293:1986–94.
12. Barnetson RA, Tenesa A, Farrington SM, Nicholl ID, Cetnarskyj R, Porteous ME, et al. Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer. *N Engl J Med.* 2006;354:2751–63.
13. Kane MF, Loda M, Gaida GM, Lipman J, Mishra R, Goldman H, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res.* 1997;57: 808–11.
14. Rigau V, Sebbagh N, Olschwang S, Paraf F, Mourra N, Parc Y, et al. 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.
15. Ewald J, Rodrigue CM, Mourra N, Lefevre JH, Flejou JF, Tiret E, et al. Immunohistochemical staining for mismatch repair proteins, and its relevance in the diagnosis of hereditary nonpolyposis colorectal cancer. *Br J Surg.* 2007;94:1020–7.
16. Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology.* 2002;123:1804–11.
17. Buhard O, Cattaneo F, Wong YF, Yim SF, Friedman E, Flejou JF, et al. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J Clin Oncol.* 2006;24:241–51.
18. Capel E, Flejou JF, Hamelin R. Assessment of *MLH1* promoter methylation in relation to gene expression requires specific analysis. Oncogene. 2007;26:7596–600.
19. de la Chapelle A, Hampel H. Clinical relevance of microsatellite instability in colorectal cancer. *J Clin Oncol.* 2010;28:3380–7.
20. Parc Y, Gueroult S, Mourra N, Serfaty L, Flejou JF, Tiret E, et al. Prognostic significance of microsatellite instability determined by immunohistochemical staining of *MSH2* and *MLH1* in sporadic T3N0M0 colon cancer. *Gut.* 2004;53:371–5.
21. Baudhuin LM, Burgart LJ, Leontovich O, Thibodeau SN. Use of microsatellite instability and immunohistochemistry testing for the identification of individuals at risk for Lynch syndrome. *Fam Cancer.* 2005;4:255–65.
22. Palomaki GE, McClain MR, Melillo S, Hampel HL, Thibodeau SN. EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. *Genet Med.* 2009;11:42–65.
23. Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, et al. Feasibility of screening for Lynch syndrome among patients with colorectal cancer. *J Clin Oncol.* 2008;26: 5783–8.
24. Mitchell RJ, Brewster D, Campbell H, Porteous ME, Wyllie AH, Bird CC, et al. Accuracy of reporting of family history of colorectal cancer. *Gut.* 2004;53:291–5.
25. Vasen HF, Moslein G, Alonso A, Aretz S, Bernstein I, Bertario L, et al. Recommendations to improve identification of hereditary and familial colorectal cancer in Europe. *Fam Cancer.* 2010;9: 109–15.
26. Jarvinen HJ, Aarnio M, Mustonen H, Aktan-Collan K, Aaltonen LA, Peltomaki P, et al. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. *Gastroenterology.* 2000;118:829–34.
27. 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;41:112–7.

Original article

Risk factors for development of desmoid tumours in familial adenomatous polyposis

J. H. Lefevre[1], Y. Parc[1], S. Kernéis[2], N. Goasguen[1], M. Benis[1], R. Parc[1] and E. Tiret[1]

[1]Department of Digestive Surgery, Hospital Saint-Antoine, Assistance Publique-Hôpitaux de Paris (AP-HP), Pierre and Marie Curie University, and [2]Department of Biostatistics and Medical Informatics, Hospital Saint-Louis, AP-HP, University of Paris VII, Paris, France

Correspondence to: Professor Y. Parc, Department of Digestive Surgery, Hôpital Saint-Antoine, Université Pierre et Marie Curie, 184 Rue du Faubourg Saint-Antoine, F-75571 Paris, France (e-mail: yann.parc@sat.ap-hop-paris.fr)

Background: **Desmoid tumours (DTs) are the primary cause of death of patients with familial adenomatous polyposis (FAP) following restorative proctocolectomy. The aim of this study was to identify risk factors for DT in a French population.**

Methods: **Clinical data for 442 patients with FAP from 1983 to 2004 were reviewed retrospectively.**

Results: **A total of 124 DTs were documented in 50 patients (25 female). DT sites were mesenteric (73 tumours), abdominal wall (44) and extra-abdominal (seven). Female patients developed DT earlier than males. Although DTs appeared after colectomy in 34 patients, the type of surgery did not influence the risk of DT. An identified point mutation in the adenomatous polyposis coli (*APC*) gene after codon 1444 was a significant risk factor (hazard ratio 3·3 (95 per cent confidence interval 1·5 to 7·3)). Belonging to a family affected by DT did not increase the individual's risk in this population.**

Conclusion: **No risk factor for life-threatening mesenteric DT could meaningfully modify the management of patients with FAP.**

Presented to the 15th United European Gastroenterology Week, Paris, France, October 2007, and published in abstract form as *Gut* 2007; **56**(Suppl III): A323

Paper accepted 26 March 2008

Published online 25 June 2008 in Wiley InterScience (www.bjs.co.uk). **DOI:** 10.1002/bjs.6241

Introduction

Since the introduction of prophylactic colectomy, extracolonic manifestations have become the new therapeutic challenge in familial adenomatous polyposis (FAP). Desmoid tumours (DTs) are the leading cause of death after restorative proctocolectomy[1]. The term aggressive fibromatosis is sometimes employed to describe better the marked cellularity and aggressive local behaviour of these lesions[2]. Between 7 and 15 per cent of patients with FAP are estimated to be affected by DT[2,3]. Surgery is safe when the tumour lies in the abdominal wall but the perioperative mortality rate can reach 36 per cent for mesenteric desmoids[2,3].

Four predictive factors for developing DT have been reported: antecedent trauma, oestrogen exposure, family history, and presence and location of adenomatous polyposis coli (*APC*) mutations[2,4–6]. The aim of this retrospective study was to identify risk factors for DT in a French population.

Methods

Charts of all patients with FAP treated between January 1983 and September 2004 were reviewed retrospectively. The diagnosis of DT was ascertained by a combination of clinical examination, surgical exploration and radiological investigation.

Statistical analysis

Continuous data are presented as median (range). Fisher's exact test was used to compare categorical variables and Mann–Whitney or Wilcoxon test for continuous variables. All tests were two sided at the 0·050 significance level. Risk factors for DT were studied by entering six prognostic

The Editors are satisfied that all authors have contributed significantly to this publication

Copyright © 2008 British Journal of Surgery Society Ltd
Published by John Wiley & Sons Ltd

co-variates in the model (sex, site of frameshift mutation, total number of polyps, age at first surgery, type of surgical intervention, number of surgical procedures). To study the impact of being in a family affected by DT, a Cox proportional hazards regression model was fitted, with the family effect regarded as a random effect, the occurrence of a DT as the response variable, and the location of the mutation as a co-variate[7,8]. Analyses were performed using the R statistical package (R Foundation for Statistical Computing, Vienna, Austria).

Results

Patient characteristics are detailed in *Table 1*. An APC germline mutation was found in 323 patients. Fifty patients (11·2 per cent) had a DT (*Table 2*). Nineteen patients (38 per cent) had DTs at several sites. DTs were discovered before any surgery in five patients (median 5·7 (0·2–23) years before colorectal surgery), during the surgical procedure in 11 patients, and a median of 5·3 (0·2–31) years after operation in the remaining 34 patients. DTs were asymptomatic in 31 patients (62 per cent) and discovered on clinical examination during either follow-up or surgery. They were found during the investigation of abdominal pain (three patients), small bowel obstruction (four), altered gastrointestinal function (six) and ureteric obstruction with pyeloureteral dilatation (six).

Two patients died from mesenteric desmoids at the ages of 46 and 48 years. One suffered fistulation of the DT into the rectum leading to massive rectal bleeding. The other developed multiple fistulas, with necrosis of a large irresectable DT causing multiple organ failure.

Table 1 Characteristics of 442 patients with familial adenomatous polyposis

	n
Female sex	222 (50·2)
Genetics	
Identified mutations	323 (73·1)
Point mutations	266 (60·2)
Genomic deletions	37 (8·4)
Splicing	14 (3·2)
Other	6 (1·4)
No identified mutation	32 (7·2)
Mutation not searched	87 (19·7)
Median (range) age at first surgery (years)*	31 (9–68)
Type of surgical intervention†	
Ileal pouch–anal anastomosis	334 (78·6)
Ileorectal anastomosis	57 (13·4)
Abdominoperineal resection	29 (6·8)
Other	5 (1·1)
Patients with ≥ 2 surgical procedures	68 (16·0)
Desmoid tumours	50 (11·3)

Values in parentheses are percentages unless indicated otherwise. Data based on *428 or †425 patients.

Desmoid tumour risk factors

Characteristics of patients with and without desmoid tumours are compared in *Table 3*. Development of a DT did not show sex predilection, but the median age at diagnosis was significantly lower in women than in men (27·5 (5–38) *versus* 38·9 (9–68) years; $P = 0·001$). Type of prophylactic surgery did not appear to influence the risk of developing a DT (*Table 3*).

Among 266 patients with a point mutation in the *APC* gene, the proportion of patients with DT was 12·9 per cent (31 of 241) when the mutation was 5′ to codon 1444

Table 2 Characteristics of 50 patients with desmoid tumours

	n
Female sex	25 (50)
Median (range) age at diagnosis of DT	32 (5–68)
Site of DTs	124
Mesenteric	73 (58·9)
Abdominal wall	44 (35·5)
Extra-abdominal	7 (5·6)
Genetics	
Identified mutations	42
Point mutations	41 (82)
Genomic deletions	1 (2)
Mutation not searched	8 (16)

Values in parentheses are percentages unless indicated otherwise. DT, desmoid tumour.

Table 3 Characteristics of patients with or without desmoid tumours

	No desmoid ($n = 392$)	Desmoid ($n = 50$)
Site of point mutation before 1444	16 of 225 (7·1)	31 of 41 (76)
Total no. of polyps ≥ 100	98 of 163 (60·1)	15 of 19 (79)
Female sex	197 (50·3)	25 (50)
Median (range) age at first surgery (years)	30·6 (9–68)	28·8 (14–68)
Type of surgical intervention*		
Ileal pouch–anal anastomosis	294 (78·4)	40 (80)
Ileorectal anastomosis	52 (13·9)	5 (10)
Abdominoperineal resection	26 (6·9)	3 (6)
Other	3 (0·8)	2 (4)
Patients having ≥ 2 surgical procedures	64 (17·1)	4 (8)

Values in parentheses are percentages unless indicated otherwise. *Data based on 425 patients.

Copyright © 2008 British Journal of Surgery Society Ltd
Published by John Wiley & Sons Ltd

Table 4 Cox proportional hazards regression analysis of risk factors for desmoid tumours

	n	Hazard ratio	*P*
Site of point mutation	266		
≤ 1444		1	
> 1444		3·3 (1·5, 7·3)	0·032
Total no. of polyps	182		
< 100		1	
≥ 100		2·5 (0·8, 8·0)	0·130
Sex	442		
M		1	
F		1·2 (0·7, 2·2)	0·520

Values in parentheses are 95 per cent confidence intervals.

compared with 40 per cent (ten of 25) when the mutation was 3′ to codon 1444 ($P < 0·001$). An identified point mutation after codon 1444 was a significant risk factor (*Table 4*). Intra-abdominal tumours tended to be more frequent if the mutation occurred after codon 1444 (27 of 30 *versus* seven of ten), but the difference was not statistically significant ($P = 0·125$). For abdominal wall DTs (11 of 30 *versus* eight of ten; $P = 0·018$) and extra-abdominal DTs (one of 30 *versus* five of ten; $P < 0·001$), the location of the mutation was significant. However, interpretation of these findings is limited by the very small sample size.

Among the 71 families of patients with FAP, representing 189 cases, 18 patients with DTs belonged to a FAP family. The test of random effect, when adjusting for the site of frameshift mutation, was not significant ($P = 0·240$).

Discussion

The prevalence of DT (11·3 per cent) reported here is comparable to published values of 9–15 per cent[1–3,9–12]. No increased risk for females was found, although DTs appeared earlier in women, in keeping with a possible hormonal influence on tumour growth[2,9,13].

The search for a phenotype–genotype correlation revealed an increased risk for DT at the 3′ end of codon 1444, similar to previous findings[1,10]. The difference was not significant when only life-threatening mesenteric DTs were evaluated, but the study was not powered for this subanalysis. A recent report[12] on a national Dutch cohort did not find a correlation between DT and location of the mutation. A more limited area of the *APC* gene is linked to extra-abdominal DT development (codons 1444–1580)[10]. Thus the impact of surgery on patients with this mutation may be less catastrophic (mesenteric DTs are more life threatening)[3].

A strong family history is considered a significant risk, regardless of the mutation site[6,12]. In this series, however, belonging to a family affected by DT did not increase the risk of desmoids, when adjusted for the site of the mutation and taking into account the clustered structure of the population. This may be the first time that the present methodology has been used to investigate the effect of family on the occurrence of DT. Contrasting results with existing studies may be partly explained by the statistical approach and further investigation is needed.

The only modifiable factor is the type of surgical procedure, but this did not have a significant influence on the risk of DT. The findings of this cohort analysis should be considered with caution as the sample size is limited.

References

1 Speake D, Evans DG, Lalloo F, Scott NA, Hill J. Desmoid tumours in patients with familial adenomatous polyposis and desmoid region adenomatous polyposis coli mutations. *Br J Surg* 2007; **94**: 1009–1013.

2 Shields CJ, Winter DC, Kirwan WO, Redmond HP. Desmoid tumours. *Eur J Surg Oncol* 2001; **27**: 701–706.

3 Clark SK, Neale KF, Landgrebe JC, Phillips RK. Desmoid tumours complicating familial adenomatous polyposis. *Br J Surg* 1999; **86**: 1185–1189.

4 Penna C, Tiret E, Parc R, Sfaii A, Kartheuser A, Hannoun L *et al*. Operation and abdominal desmoid tumors in familial adenomatous polyposis. *Surg Gynecol Obstet* 1993; **177**: 263–268.

5 Caspari R, Olschwang S, Friedl W, Mandl M, Boisson C, Böker T *et al*. Familial adenomatous polyposis: desmoid tumours and lack of ophthalmic lesions (CHRPE) associated with APC mutations beyond codon 1444. *Hum Mol Genet* 1995; **4**: 337–340.

6 Sturt NJH, Gallagher MC, Bassett P, Philp CR, Neale KF, Tomlinson IP *et al*. Evidence for genetic predisposition to desmoid tumours in familial adenomatous polyposis independent of the germline APC mutation. *Gut* 2004; **53**: 1832–1836.

7 Heo M, Leon AC. Comparison of statistical methods for analysis of clustered binary observations. *Stat Med* 2005; **24**: 911–923.

8 Zeger SL, Karim MR. Generalized linear models with random effects, a Gibbs sampling approach. *J Am Stat Assoc* 1991; **86**: 78–87.

9 Clark SK, Phillips RKS. Desmoids in familial adenomatous polyposis. *Br J Surg* 1996; **83**: 1494–1504.

10 Friedl W, Caspari R, Sengteller M, Uhlhaas S, Lamberti C, Jungsk M *et al*. Can APC mutation analysis contribute to therapeutic decisions in familial adenomatous polyposis? Experience from 680 FAP families. *Gut* 2001; **48**: 515–521.

Copyright © 2008 British Journal of Surgery Society Ltd
Published by John Wiley & Sons Ltd

Article 3. Family history, surgery and APC mutation are risk factor for desmoid tumors in FAP: An international cohort study (Dis Colon Rectum)

ORIGINAL CONTRIBUTION

Family History, Surgery, and *APC* Mutation Are Risk Factors for Desmoid Tumors in Familial Adenomatous Polyposis: An International Cohort Study

Marry H. Nieuwenhuis, M.D., M.Sc.[1] • Jérémie H. Lefevre, M.D.[2]
Steffen Bülow, M.D., Ph.D.[3] • Heikki Järvinen, M.D., Ph.D.[4]
Lucio Bertario, M.D., Ph.D.[5] • Solen Kernéis, M.D.[6] • Yann Parc, M.D., Ph.D.[2]
Hans F. A. Vasen, M.D., Ph.D.[1,7]

1 The Netherlands Foundation for the Detection of Hereditary Tumors, Leiden, The Netherlands
2 Department of Digestive Surgery, Hospital Saint-Antoine, AP-HP, University Pierre et Marie Curie Paris VI, Paris, France
3 The Danish Polyposis Register, Department of Surgery, Hvidovre University Hospital, Hvidovre, Denmark
4 Department of Surgery, University of Helsinki, Helsinki, Finland
5 Department of Predictive and Preventive Medicine, National Cancer Institute, Milan, Italy
6 Cochin-Pasteur Vaccinology Center (CIC BT505), Université Paris Descartes, INSERM, Assistance Publique Hôpitaux de Paris (AP-HP), Paris, France
7 Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, The Netherlands

BACKGROUND: Ability to identify patients with familial adenomatous polyposis who have a high risk of developing desmoid tumors may affect decisions in clinical practice.

OBJECTIVES: Our aim was to assess several risk factors for desmoid tumor development in an international cohort of patients with familial adenomatous polyposis and to evaluate the clinical relevance of risk factors.

DESIGN: This was a retrospective cohort study.

SETTING AND PATIENTS: Polyposis registries in The Netherlands, France, Denmark, Finland, and Italy provided information on familial adenomatous polyposis patients with desmoid tumors.

MAIN OUTCOME MEASURES: We used univariate and multivariable analyses of data from registries in The Netherlands, France, Denmark, and Finland to test whether gender, *APC* mutation site, previous colorectal surgery, colorectal cancer, and family history for desmoid tumors contribute to risk of developing desmoid tumors at any location, or specifically at an intra-abdominal location. The effect of family history was tested with a generalized linear mixed model.

RESULTS: Of 2260 patients with familial adenomatous polyposis from 912 families in The Netherlands, France, Denmark, and Finland, 220 patients (10%) had desmoid tumors (101 men). In 387 patients with desmoid tumors (including 167 patients from the Italian registry), the median age at diagnosis of the first desmoid tumor was 31 years (range, 4 months–74 years). Desmoid locations were intra-abdominal (53%), abdominal wall (24%), extremities (9%), and unknown sites or combinations of sites (14%). Multivariable analysis of risk factors for desmoids at any location showed surgery (OR, 2.58; $P = .0004$), an *APC* mutation 3′ of codon 1444 (OR, 3.0; $P < .0001$), and a positive family history ($P < .0001$) to be independently associated with desmoid development. When only intra-abdominal location was analyzed, *APC* mutation site was not associated with desmoid development.

LIMITATIONS: Selection bias may have occurred.

CONCLUSIONS: A positive family history for desmoid tumors, abdominal surgery, and *APC* mutation site are significant risk factors for development of desmoid tumors. The results may have implications for

Financial Disclosure: None reported.

Presented at the meeting of the International Society for GI Hereditary Tumours, San Antonio, TX, March 30 to April 2, 2011.

Correspondence: Marry H. Nieuwenhuis, M.D., M.Sc., The Netherlands Foundation for the Detection of Hereditary Tumors, Rijnsburgerweg 10, Poortgebouw Zuid, 2333 AA Leiden, The Netherlands. E-mail: m.nieuwenhuis@stoet.nl

Dis Colon Rectum 2011; 54: 1229–1234
DOI: 10.1097/DCR.0b013e318227e4e8
©The ASCRS 2011

Copyright © The American Society of Colon & Rectal Surgeons, Inc. Unauthorized reproduction of this article is prohibited.

determining the optimal management of FAP patients and guide future studies.

KEY WORDS: Desmoid tumor; Familial adenomatous polyposis; Risk factors.

Desmoid tumors are a serious manifestation of familial adenomatous polyposis (FAP), occurring in about 12% of patients.[1,2] A germline mutation in the adenomatous polyposis coli gene (*APC*) in patients with FAP leads to development of hundreds of colorectal adenomas. Without prophylactic colectomy, colorectal cancer inevitably results. In addition to the development of colorectal adenomas, patients with FAP may display various other manifestations, with desmoid tumors as one of the most threatening.

Although desmoid tumors are histologically benign fibroblastic neoplasms, intra-abdominal desmoids in particular cause substantial morbidity and even mortality in patients with FAP. The etiology of desmoid tumors has not yet been elucidated. Several potential risk factors have been suggested, including surgical trauma, estrogens, pregnancy and female gender, a strong family history of desmoids, and mutations at the 3′ end of the *APC* gene.[1–11] However, some potential risk factors, for example female gender, a positive desmoid family history, and a 3′ *APC* mutation site, were not consistently found in previous studies.[5,7,8,10] Moreover, the clinical relevance of some risk factors is questionable.

The aim of the present study was to assess risk factors for desmoid tumor development in a large European database of FAP patients. A secondary aim was to evaluate the potential clinical implications of such risk factors.

METHODS

Data were retrieved from 5 European polyposis registries located in The Netherlands, France, Denmark, Finland, and Italy. Patients with either a genetically (germline *APC* mutation) or clinically (more than 100 colorectal adenomas) confirmed diagnosis of FAP were included in the analysis. The Italian, Dutch, and French data had been used in previous studies on desmoid risk factors.[2,7,8] Because the Italian registry included only data on patients with desmoid tumors, these data were not included in the risk factor analysis. All patients in the registries had given informed consent for use of their data for research, and all patient data were handled anonymously.

For all patients, information was obtained on gender, *APC* mutation site (upstream of codon 1444 in the 5′ direction vs downstream of codon 1444 in the 3′ direction), age at first colorectal surgical procedure, type of colorectal surgery, and colorectal cancer at any time in life. For the patients with desmoid tumors, additional data were collected including age at diagnosis of the desmoid tumor and desmoid location.

Results are reported as median (range) for continuous variables, and number (percentage) for categorical variables. The associations between 4 potential risk factors (gender, mutation site, surgery, colorectal cancer) and development of a desmoid tumor were first assessed by univariate analysis, using Student *T* test for continuous variables and the χ^2 test for categorical variables. All tests were 2-sided at a .05 significance level. Variables with a *P* value lower than .2 were included in a multivariable logistic regression model. Analyses were performed using StatView computer software (version 5, 1992–1998, SAS Institute, Cary, NC).

To assess the impact of belonging to a family affected by desmoid tumors, a generalized linear mixed model was fitted, with the occurrence of a desmoid tumor as the response variable, the family effect regarded as a random effect, and the location of the mutation as a covariate. These analyses were restricted to individuals belonging to families where data were available for at least 2 family members. This mixed-effect model was compared to a fixed logistic regression model with the same response variable and covariate (likelihood ratio test). Testing this random effect allowed us to determine whether belonging to a family affected by desmoid tumors would modify an individual's risk of displaying a desmoid tumor, while simultaneously allowing us to adjust for the site of a frameshift mutation. Analyses were performed using the R statistical package (R Development Core Team; R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org).

RESULTS

Characteristics of Patients With FAP

Data available for patients with FAP registered in databases in Denmark, Finland, France, and The Netherlands (Table 1) included 912 families comprising 2260 patients (52% men). Most patients (61%) had undergone surgery once, at a median age of 30 years. In the majority of cases, a subtotal colectomy with ileorectal anastomosis (IRA, n = 740, 39%) or total colectomy with ileal pouch–anal anastomosis (IPAA, n = 806, 43%) was performed. Of all patients, 508 (22%) had had a diagnosis of colorectal cancer. *APC* mutations 5′ of codon 1444 were the most common, occurring in 62% of patients, whereas mutations 3′ of codon 1444 occurred in 9%. A total of 220 patients (10%) developed at least 1 desmoid tumor.

Characteristics of Patients With Desmoid Tumors

The total cohort of FAP patients with desmoid tumors consisted of 387 patients, including the 220 patients from

TABLE 1. Characteristics of FAP patients from 4 European registries

	Netherlands (n = 992)	*France (n = 434)*	*Denmark (n = 498)*	*Finland (n = 336)*	*Total (n = 2260)*
No. of families	289	315	187	121	912
Sex, n (%)					
Men	517 (52)	214 (49)	276 (55)	157 (47)	1164 (52)
Women	475 (48)	220 (51)	222 (45)	179 (53)	1096 (48)
APC mutations, n (%)					
5′ of codon 1444	674 (68)	266 (61)	252 (50)	214 (60)	1406 (62)
3′ of codon 1444	82 (8)	22 (5)	46 (9)	49 (15)	199 (9)
Large deletion	53 (5)	36 (8)	0 (0)	13 (4)	69 (3)
Whole gene deletion	2 (0.2)	1 (0.2)	25 (5)	8 (2)	36 (2)
No/unknown mutation	181 (12)	109 (31)	175 (36)	52 (15)	517 (23)
Age at first surgery, y					
Median (range)	28 (5–91)	30 (6–67)	31 (5–87)	31 (1–71)	30 (1–91)
Number of operations, n (%)					
No operation	243 (24)	1 (0.2)	76 (15)	58 (17)	378 (17)
One operation	572 (58)	378 (87)	227 (46)	199 (59)	1376 (61)
More than one	177 (18)	55 (13)	195 (39)	79 (24)	506 (22)
Type of first surgery, n (%)					
IRA	347 (46)	63 (15)	167 (40)	127 (45)	740 (39)
IPAA	272 (36)	340 (79)	156 (37)	74 (27)	806 (43)
Proctocolectomy + ileostomy	57 (8)	24 (6)	20 (5)	42 (15)	143 (8)
Other procedure	73 (10)	6 (1)	79 (19)	35 (13)	193 (10)
Colorectal cancer, n (%)					
Yes	141 (14)	90 (21)	183 (37)	94 (28)	508 (22)
No	851 (86)	344 (79)	315 (63)	242 (72)	1752 (78)
At least 1 desmoid tumor	83 (8)	59 (14)	33 (7)	45 (13)	220 (10)

FAP = familial adenomatous polyposis; IRA = subtotal colectomy with ileorectal anastomosis; IPAA = total colectomy with ileal pouch–anal anastomosis.

Denmark, Finland, France, and The Netherlands, plus 167 patients from the Italian registry. Characteristics of this cohort are shown in Table 2. The median age at diagnosis was 31 years (range, 4 months–74 years). Approximately half (53%) of the initial desmoids had an intra-abdominal location. Intra-abdominal desmoid tumors were diagnosed at a significantly later median age than tumors at other locations (35 vs 28 years, $P < .0001$). Of all 387 patients with desmoid tumors, 324 (84%) had an *APC* mutation, with 5′ mutations more common than 3′ mutations (65% vs 19%). In addition, a total of 63 patients (16%) had large deletions or unknown mutations.

TABLE 2. Characteristics of FAP patients with desmoid tumors from 5 European registries

	Netherlands (n = 83)	*France (n = 59)*	*Denmark (n = 33)*	*Finland (n = 45)*	*Italy (n = 167)*	*Total (n = 387)*
Age at 1st desmoid diagnosis, median (range), y	31 (8–58)	39 (16–68)	30 (17–74)	28 ([a]–69)	31 ([b]–67)	31 (0–74)
Location of 1st desmoid, n (%)						
Intra-abdominal (mesentery)	51 (61)	30 (51)	16 (48)	24 (53)	84 (50)	205 (53)
Abdominal wall	13 (16)	8 (13)	13 (39)	15 (33)	46 (28)	95 (24)
Extremities	5 (6)	1 (2)	4 (12)	6 (13)	18 (11)	34 (9)
Unknown/combination of sites	14 (17)	20 (34)	0	0	19 (11)	53 (14)
Desmoid without colorectal surgery, n (%)	22 (27)	9 (15)	7 (21)	15 (33)	56 (33)	109 (28)
Desmoid after colorectal surgery, n (%)	61 (73)	50 (85)	26 (79)	30 (67)	111 (66)	278 (72)
Time between surgery and desmoid, median (range), months	49 (1–474)	60 (1–261)	30 (13–318)	30 (5–262)	32 (1–366)	36 (1–474)
APC mutations, n (%)						
5′ of codon 1444	62 (75)	36 (61)	17 (52)	16 (35)	120 (72)	251 (65)
3′ of codon 1444	8 (10)	11 (19)	4 (12)	18 (40)	32 (19)	73 (19)
Large deletion	1 (1)	2 (3)	0 (0)	0 (0)	0 (0)	3 (1)
Whole gene deletion	0 (0)	0 (0)	2 (6)	4 (9)	1 (1)	7 (2)
No/unknown mutation	12 (14)	10 (17)	10 (30)	7 (16)	14 (8)	53 (13)

FAP = familial adenomatous polyposis; NA = not available.
[a]Minimum, 8 months.
[b]Minimum, 4 months.

TABLE 3. Univariate and multivariable analysis of risk factors for desmoid tumors at any location

	Univariate		*Multivariable*	
	Desmoid tumor, n/N (%)	*P*	*OR (95% CI)*	*P*
Sex		.08		NS
Women	119/1096 (10.9)			
Men	101/1164 (8.7)			
APC mutation site		<.0001		<.0001
5′ of codon 1444	133/1406 (9.5)		1	
3′ of codon 1444	41/199 (20.6)		3.0 (1.99–4.46)	
Age at first surgery		.003		NS
≤31 years	71/859 (8.3)			
>31 years	126/1007 (12.5)			
Previous abdominal surgery		.003		.0004
No operation	21/378 (5.6)		1	
Operation	199/1882 (10.6)		2.58 (1.53–4.35)	
Type of surgery		.53		
IRA	78/704 (11.1)			
IPAA	85/842 (10.1)			
Colorectal cancer		.45		
Yes	45/508 (20.4)			
No	175/1752 (9.9)			

N = number of patients with available data from registries in The Netherlands, France, Denmark, and Finland; OR = odds ratio; NS = not significant; IRA = subtotal colectomy with ileorectal anastomosis; IPAA = total colectomy with ileal pouch–anal anastomosis.

Desmoid tumors occurred without a history of abdominal surgery in 109 patients (28%) and after surgery in 278 patients (72%). Mutations 3′ of codon 1444 occurred more frequently in patients without a history of abdominal surgery than in those with a postoperative desmoid tumor (40% vs 15%, $P < .0001$). Analysis of available data on tumor location showed an intra-abdominal location in 49% (47 of 96) for preoperative desmoids and in 66% (158 of 238) for desmoids developing after surgery, $P = .0031$. Postoperative desmoids developed at a median of 36 months (range, 1–474 months) after surgery.

Risk Factor Analysis

Our analysis of data of patients from Denmark, Finland, France, and The Netherlands showed that location of an *APC* mutation 3′ of codon 1444 and previous abdominal surgery were significant risk factors for developing a desmoid tumor, regardless of tumor location (Table 3). Sub-

TABLE 4. Univariate and multivariable analysis of risk factors for intra-abdominal desmoid tumors

	Univariate		*Multivariable*	
	Intra-abdominal desmoid tumor, n/N (%)	*P*	*OR (95% CI)*	*P*
Sex		.88		
Women	62/1096 (5.7)			
Men	60/1164 (5.2)			
APC mutation site		.19		NS
5′ of codon 1444	80/1406 (5.7)			
3′ of codon 1444	16/199 (8.0)			
Age at first surgery		.05		NS
≤31 years	41/859 (4.8)			
>31 years	70/1007 (7)			
Previous abdominal surgery		.009		.0016
No operation	10/378 (2.6)		1	
Operation	112/1882 (6)		2.33 (1.21–4.49)	
Type of surgery		.50		
IRA	45/704 (6.4)			
IPAA	47/842 (5.6)			
Colorectal cancer		.32		
Yes	23/508 (4.5)			
No	99/1752 (5.7)			

N = number of patients with available data from registries in The Netherlands, France, Denmark, and Finland; OR = odds ratio; NS = not significant; IRA = subtotal colectomy with ileorectal anastomosis; IPAA = total colectomy with ileal pouch–anal anastomosis.

analysis specifically for intra-abdominal desmoid tumors (Table 4) showed abdominal surgery to be the only significant risk factor for desmoid development. Gender and previous colorectal cancer were not associated with desmoid development. Although age was a significant factor in univariate analysis, it was not an independent risk factor in the multivariable analysis.

Based on the likelihood ratio test, the mixed-effect model, with family regarded as a random effect variable, differed significantly from the fixed logistic regression model (difference of deviances = 69.8, $P < 10^{-5}$). This suggests that belonging to a family affected by desmoid tumors significantly increases the individual risk of displaying a desmoid tumor and that this effect is independent of the frame-shift mutation site.

DISCUSSION

In the present study in a large international cohort of FAP patients, an *APC* mutation 3′ of codon 1444, previous abdominal surgery, and a positive family history of desmoid tumors were independent risk factors for desmoid tumor development. As far as we know, this study used the largest desmoid database currently in existence. However, because of the retrospective design of the study, we must take into account the possibility of selection bias for desmoid cases. Moreover, there may be differences in registration methods in the different countries. Despite these caveats, we believe that our cohort provides a reliable representation of European FAP patients with clinically relevant desmoid disease.

Several previous studies of risk factors for desmoid tumor development have reported contradictory results. Only previous abdominal surgery has been consistently found to influence desmoid development.[1,9–12] Our current study confirms previous abdominal surgery as an evident risk factor for intra-abdominal desmoid development.

Familial clustering of desmoid tumors was reported as a risk factor in at least 3 studies,[2,5,10] but was not found in a French study.[8] Because FAP is a hereditary disease, analysis of the impact of familial clustering is complicated. We used robust statistical analysis, taking into account information on *APC* mutations, and our calculations provided strong evidence for familial clustering of desmoid tumors.

Since the first reports on genotype-phenotype correlations in patients with FAP, several investigators have shown a high risk of desmoid tumors in patients with 3′ *APC* mutations, particularly mutations 3′ of codon 1444.[2,4,8,9] However, the cutoff at this codon was arbitrary and other studies could not confirm this association.[8,10] Remarkably, in our study, a mutation site 3′of codon 1444 was a statistically significant risk factor for the development of desmoid tumors when we analyzed all tumor locations together, but not when we analyzed only intra-abdominal desmoids, which are clinically the most threatening ones.[8]

There is controversy about the influence of female gender on the risk of developing FAP-related desmoid tumors.[2,5,7–10] In a recent study, we showed that sporadic desmoid tumors were more common in female patients than in male patients, whereas FAP related desmoids were evenly distributed over both sexes.[13] Our current results confirm equal gender distribution for desmoids associated with FAP. We also investigated whether previous colorectal cancer was associated with desmoid development, but could not confirm such an association. Although this variable was not previously shown to be a risk factor, we wanted to exclude a possible relation between desmoids and colorectal cancer.

The increased risk of desmoid development after surgery can be explained by the function of fibroblasts, which are involved in repairing tissue damage. In a subset of FAP patients, the surgical procedure triggers uncontrolled growth of fibroblasts, resulting in a desmoid tumor. However, as this does not occur in all FAP patients, the question is which additional factors induce desmoid development. The role of the *APC* mutation site is unclear, because it does not influence the risk of development of intra-abdominal desmoids. However, other genetic modifiers may play an important role in desmoid development. The fact that intra-abdominal desmoids were found later than those at other locations may also be attributable to a need for additional somatic mutations/inactivation in order for such tumors to develop. Possibly, a combination of genetic changes and environmental factors causes desmoid development. Genetic modifiers would also explain the finding of a strong family effect, as family members share genetic variations.

For clinical practice, it is important to identify those FAP patients at high risk of desmoid tumors. Deferring colorectal surgery may be considered in cooperative patients in whom frequent surveillance is guaranteed, particularly if patients have a positive family history of desmoids.[14] If surgery is needed, there is a choice between IRA and IPAA. Elayi et al[11] proposed performance of minimally invasive surgery (i.e., IRA) in patients with a high risk of desmoid tumors. However, to avoid a situation in which the development of intra-abdominal desmoids precludes secondary proctectomy,[7] we advise consideration of primary IPAA in such patients. In addition to surgery, prophylactic treatment with antiestrogens and nonsteroidal anti-inflammatory drugs (NSAIDs) may be considered[15] in patients at high risk for desmoid development. Currently, information on the mutation site seems not to be very useful in clinical practice. Particularly for intra-abdominal desmoid tumors, the relevance of the mutation site is not evident. Further efforts should be made to find genetic modifiers.

CONCLUSION

A positive family history for desmoid tumors, abdominal surgery, and *APC* mutation site are significant risk factors for development of desmoid tumors. The results may have implications for determining the optimal management of FAP patients and guide future studies.

ACKNOWLEDGMENTS

The authors thank Paola Sala for support with the Polyposis Registry and Paolo Radice for laboratory work.

REFERENCES

1. Sinha A, Tekkis PP, Gibbons DC, Phillips RK, Clark SK. Risk factors predicting desmoid occurrence in patients with familial adenomatous polyposis: a meta-analysis [published online ahead of print June 10, 2010]. *Colorectal Dis.* doi:10.1111/j.1463-1318.2010.02345.x.
2. Bertario L, Russo A, Sala P, et al. Genotype and phenotype factors as determinants of desmoid tumors in patients with familial adenomatous polyposis. *Int J Cancer.* 2001;95:102–107.
3. Lips DJ, Barker N, Clevers H, Hennipman A. The role of APC and beta-catenin in the aetiology of aggressive fibromatosis (desmoid tumors). *Eur J Surg Oncol.* 2009;35:3–10.
4. Caspari R, Olschwang S, Friedl W, et al. Familial adenomatous polyposis: desmoid tumours and lack of ophthalmic lesions (CHRPE) associated with APC mutations beyond codon 1444. *Hum Mol Genet.* 1995;4:337–340.
5. Sturt NJ, Gallagher MC, Bassett P, et al. Evidence for genetic predisposition to desmoid tumours in familial adenomatous polyposis independent of the germline APC mutation. *Gut.* 2004;53:1832–1836.
6. Sturt NJ, Clark SK. Current ideas in desmoid tumours. *Fam Cancer.* 2006;5:275–288.
7. Nieuwenhuis MH, De Vos Tot Nederveen Cappel W, Botma A, et al. Desmoid tumors in a dutch cohort of patients with familial adenomatous polyposis. *Clin Gastroenterol Hepatol.* 2008;6:215–219.
8. Lefevre JH, Parc Y, Kerneis S, et al. Risk factors for development of desmoid tumours in familial adenomatous polyposis. *Br J Surg.* 2008;95:1136–1139.
9. Durno C, Monga N, Bapat B, et al. Does early colectomy increase desmoid risk in familial adenomatous polyposis? *Clin Gastroenterol Hepatol.* 2007;5:1190–1194.
10. Gurbuz AK, Giardiello FM, Petersen GM, et al. Desmoid tumours in familial adenomatous polyposis. *Gut.* 1994;35:377–381.
11. Elayi E, Manilich E, Church J. Polishing the crystal ball: knowing genotype improves ability to predict desmoid disease in patients with familial adenomatous polyposis. *Dis Colon Rectum.* 2009;52:1762–1766.
12. Clark SK, Phillips RK. Desmoids in familial adenomatous polyposis. *Br J Surg.* 1996;83:1494–1504.
13. Nieuwenhuis MH, Casparie M, Mathus-Vliegen LM, et al. A nation-wide study comparing sporadic and familial adenomatous polyposis-related desmoid-type fibromatoses. *Int J Cancer.* 2011;129:256–261.
14. Vasen HF, Moslein G, Alonso A, et al. Guidelines for the clinical management of familial adenomatous polyposis (FAP). *Gut.* 2008;57:704–713.
15. Janinis J, Patriki M, Vini L, Aravantinos G, Whelan JS. The pharmacological treatment of aggressive fibromatosis: a systematic review. *Ann Oncol.* 2003;14:181–190.

Article 4. Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists. (Clin Genet)

CLINICAL GENETICS An International Journal of Genetics, Molecular and Personalized Medicine

Clin Genet 2011
Printed in Singapore. All rights reserved

© 2011 John Wiley & Sons A/S
CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2011.01676.x

Original Article

Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists

Mongin C, Coulet F, Lefevre JH, Colas C, Svrcek M, Eyries M, Lahely Y, Fléjou J-F, Soubrier F, Parc Y. Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists.
Clin Genet 2011. © John Wiley & Sons A/S, 2011

Two main colorectal polyposis syndromes have been described, familial adenomatous polyposis and *MUTYH*-associated polyposis syndromes. Some polyposis remains unexplained: 20% of adenomatous polyposis and serrated polyposis. The aim of this study was to evaluate in a cohort of patients with unexplained polyposis whether a genetic defect could be detected. Individuals presenting polyposis with more than 40 adenomas or more than 20 serrated polyps (hyperplastic, sessile serrated and mixed), without causative mutation identified, were included. Complementary explorations on *APC* or *MUTYH* were performed: search for *APC* mosaicism, splicing-affecting mutations, large genomic rearrangement of *MUTYH*. Four genes of Wnt pathway (*AXIN2*, *PPP2R1B*, *WIF1*, *SFRP1*) and two genes of transforming growth factor-β (TGF-β) pathway (*SMAD4*, *BMPR1A*) were screened for germline mutation. Twenty-five patients had an unexplained adenomatous polyposis (familial or sporadic). Five pathogenic mutations were found: four in *APC* gene (with one case of mosaicism) and one in *BMPR1A* gene. The exploration of *APC* mosaicism was better performed from adenoma DNA with high-resolution melting. The screening of the candidate genes did not find any causative mutation. Thirteen individuals had an unexplained serrated polyposis and a frameshift on *SMAD4* gene was identified. All mutations were identified in familial cases of polyposis. After new pathological examination, both BMPR1A and SMAD4 cases were found to be associated with a juvenile polyposis while the polyposis was initially described as adenomatous or undetermined. In 17% (6/38) of the patients the causative mutation of the polyposis was identified. Genetic causes were heterogeneous. Sporadic polyposis patients must be considered as potential *APC* mosaicism. The histological classification of polyposis is strongly important in direct genetic exploration.

Conflict of interest

No conflict of interest to report.

C Mongin[a,b], F Coulet[b], JH Lefevre[a], C Colas[b], M Svrcek[c], M Eyries[b], Y Lahely[b], J-F Fléjou[c], F Soubrier[a] and Y Parc[a]

[a]Oncogenetic and Molecular Angiogenetic Laboratory, Hospital Pitié-Salpêtrière AP-HP and [b]Department of Digestive and General Surgery, and [c]Department of Pathology, Hospital Saint-Antoine AP-HP, University Pierre et Marie Curie, Paris VI, Paris, France

Key words: adenomatous polyposis – colorectal cancer – hereditary mixed polyposis syndrome – hyperplastic polyposis – serrated polyposis – Wnt signalling pathway

Corresponding author: Prof Florent Soubrier, Oncogenetic and Molecular Angiogenetic Laboratory, Hospital Pitié-Salpêtrière, 83 bld De l'Hopital, 75013 Paris, France.
Tel.: +33 01 42 16 13 46;
fax: +33 1 42 16 13 64;
e-mail: florent.soubrier@psl.aphp.fr

Received 24 February 2011, revised and accepted for publication 4 April 2011

Two main hereditary adenomatous colorectal polyposis syndromes have been described, namely familial adenomatous polyposis (FAP) and *MUTYH*-associated polyposis (MAP) syndromes. FAP has an incidence at birth of about 1/8300, and accounts for less than 1% of colorectal cancer (CRC) cases. FAP is characterized by an autosomal dominant inheritance where the affected subjects

develop hundreds to thousands of adenomatous polyps throughout the whole colon, usually during their teenage years (1). The occurrence of colon cancer is a frequent event after the second decade of life. This syndrome is also characterized by the development of a variable range of life-threatening extracolonic manifestations (upper gastrointestinal tumours, desmoid tumours and more benign manifestations like osteoma) (2). Germline mutations in the tumour suppressor gene *APC* are responsible for FAP and the mutation detection rate ranges above 75% (1). If not inherited, germline mutations causing genetic disorders may occur *de novo* in early embryogenesis and account for 10–25% of FAP (3, 4). *De novo* mutations may occur during cell division, leading to a mosaicism of wild-type and mutated cells in some cases. Few cases of mosaic *APC* mutations have been previously reported (5–7). Currently, mosaicism is estimated to be present in one fifth of *de novo* cases of FAP (8, 9). Bi-allelic mutations of *MUTYH*, responsible for adenomatous polyposis, with an increased risk of CRC, were described in 2002, and this recessively inherited polyposis is known as MAP (10). Subsequent studies have shown that patients with bi-allelic mutations of *MUTYH* have between 20 and 100 colorectal polyps, some duodenal and gastric polyps but exceptionally extracolonic features of FAP such as demoid and hepatoblastoma (11). MAP is responsible for about 20% of adenomatous polyposis without *APC* mutation identified (12).

Despite progress in understanding adenomatous polyposis, around 20% of cases remain *APC* and *MUTYH* mutation negative by conventional methods, including numerous sporadic cases. Patients without genetic explanation are considered as FAP and they as their families will undergo endoscopic screening and follow-up similar to polyposis. Different causes of unexplained polyposis are suggested such as unclassified variants identified in genomic DNA of known genes (*APC*, *MUTYH*) whose deleterious consequences remain to be determined (splicing effect, co-segregation analysis); unrecognized mutations in known genes (*APC*, *MUTYH*) by routine molecular methods; mutations in genes involved in non-adenomatous polyposis (*SMAD4*, *BMPR1A*) (13) if polyposis is misclassified. By screening different genes for germline mutations in 47 'multiple' adenoma patients, Lipton et al. (14) identified putative pathogenic change in *BMPR1A*. Mutations of other genes might be involved such as other regulatory genes of the Wnt signalling pathway. Indeed, *APC* regulates Wnt signalling by controlling the levels of β-catenin reaching the cell nucleus (2).

Aside from polyposis with adenomas, serrated polyposis has been recently described and associated with a certain risk of CRC (15). The serrated polyposis includes hyperplastic polyposis syndrome (HPS), sessile serrated polyposis and mixed polyposis. HPS is a rare disease including patients with one of the following symptom: more than 20–30 colorectal hyperplastic polyps (HP), at least one HP with a size greater than 1 cm, at least one HP in the proximal colon with a relative with HP (16). It remains poorly defined, and little is known about its genetic basis. Hereditary mixed polyposis syndromes comprise a collection of polyposis syndromes showing a mixture of various types of polyps. In some families with this syndrome, mutations in the *BMPR1A* gene were shown (17–19). The nosological status of the remaining cases of serrated polyposis remains unclear.

In all these mutation-negative adenomatous or serrated polyposis, a constitutive genetic or epigenetic cause is likely because the occurrence of numerous polyps cannot be satisfactorily explained by exogenous factors. The knowledge of the underlying factors has major significance for our understanding of colorectal tumourigenesis and for counselling affected families. Therefore, the aim of this study was to re-investigate the genetic origin of the disease in a series of patients with unexplained adenomatous polyposis (familial or sporadic) after routine molecular diagnosis. Additional investigations were performed in known polyposis genes (*APC*, *MUTYH*) to increase the mutation detection rate. We also completed analysis by screening genes of transforming growth factor-β (TGF-β) pathway (*BMPR1A*, *SMAD4*). Finally, we selected and screened four genes of the Wnt signalling pathway: *AXIN2*, *PPP2R1B*, *WIF1*, *SFRP1* both of which have been implicated in colorectal tumourigenesis and have already been found mutated in somatic colorectal cells (19–22). We also searched for a genetic explanation in serrated polyposis: we screened genes of TGF-β pathway (*SMAD4*, *BMPR1A*) for germline mutations in patients with unexplained serrated polyposis.

Materials and methods

Patient selection and characteristics

Our polyposis registry included 210 patients with multiple colorectal polyposis (with more than five polyps) who came for counselling at the oncogenetic consultation. Consent for the genetic study was obtained from patients, the pedigree of the family was drawn and a blood sample was collected for DNA extraction. All patients

were screened for *APC* germline mutations by common diagnostic methods (direct sequencing of exons and large rearrangement screening) and for MUTYH germline mutations by direct sequencing.

DNA and RNA extraction

Patients' DNA was isolated from peripheral lymphocytes using an automatized Extragene® extractor (Genomic Industry, Archamps, France) with a DNA extraction kit Wizard Genomic DNA® (Promega France, Charbonnières-les-Bains, France) according to the manufacturer's standard protocol. Extraction from buccal swab used FTA® (WHATMAN, Maidstone Kent, UK). DNA extraction from frozen tissues used DNA mini Kit Qiagen® (QIAGEN, Courtaboeuf, France). Total cellular RNA was extracted with PAXgene Blood RNA purification system (QIAGEN). Quality of DNA and RNA was assessed with the Nanodrop® technology (Coleman Technologies, Orlando, FL).

Mutation analysis

Mutation analysis of the *APC*, *MUTYH*, *AXIN2*, *PPP2R1B*, *WIF1*, *SFRP1*, *SMAD4*, *BMPR1A* genes was performed by direct sequencing. The complete coding sequence of these genes, including exons and intronic junctions, was amplified by polymerase chain reaction (PCR). Each primers were designed with PRIMER3 software (Frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primer sequences and PCR conditions are available on request. The PCR products were cleaned up on a Multi-Screen PCR 96-well plate (Millipore, Billerica, MA). Sequencing reactions were performed in forward and reverse orientations using the ABI BigDye Terminator v1.1 cycle sequencing kit (QIAGEN). The products of the sequencing reactions were cleaned up using Sephadex G-50 (GE Healthcare Life Sciences, Piscataway, NJ) in a Multi-Screen-HV 96-well filter plate then run up on an ABI 3730 DNA sequencer (Applied Biosystems, Courtaboeuf, France). The resulting sequence data were analysed with ABI SEQSCAPE software, version 2.5 in comparison with the reference sequences of human *APC*, *MUTYH*, *BMPR1A*, *SMAD4*, *SFRP1*, *WIF1*, *PPP2R1B*, *AXIN2* genes (NM_00003814, NM_0011048171, NM_004329.2, NM_005359, NM_003012, NM_007191, NM_181699).

Search for *APC* mosaicism

All patients included in our polyposis registry as *de novo* cases with no known family history of FAP in the siblings or the parents were selected for potential *APC* mosaicism ($n = 17$). A second reading of *APC* sequences was carried out for these patients to look for weak signals suggesting mosaicism. In case of a suspected signal or a suggestive family history (affected offspring), we analysed tumour-derived cells extracted from frozen tissue by DNA sequencing (if frozen tissue available). The identified mosaic was analysed by high-resolution melting (HRM) analysis. The level of mutation detection was calculated by performing mutated DNA dilutions. HRM analysis of the genomic DNA samples was carried out on a LightCycler 480 System (Roche Diagnostics, Meylan, France). Reaction mixture for HRM consisted in 0.3 µM of each primer, 1X LC480 HRM Master Mix (containing fluorescent intercalant agent ResoLight dye) (Roche Diagnostics) and 20 ng of genomic DNA samples. Assays were carried out in a 96-well format in 20 µl reaction volume and were performed using the touchdown PCR cycling and HRM conditions as follows. PCRs were initiated with a 10-min hold at 95°C, followed by 42 cycles of 95°C for 10 s, a touchdown annealing step (decreasing 1°C/cycle) ranging from 65°C to 55°C for 10 s and 72°C for 20 s. Each PCR run contained one negative (no DNA template) control.

Analysis of large genomic rearrangement for *APC*, *SMAD4* and *BMPR1A*

The screening of large deletions was performed by multiplex ligation probe amplification (MLPA) using the MLPA kit SALSA P043 for the *APC* gene and SALSA MLPA kit P158-B1 JPS for the *SMAD4* and *BMPR1A* gene (MRC Holland, Amsterdam, the Netherlands) according to the manufacturer's instructions. Fragment analysis of multiplex PCR was carried out on the ABI 3730 DNA analyser (Applied Biosystems), using GENEMAPPER software, version 4.0 (Applied Biosystems).

cDNA analysis of *APC* and *MUTYH* genes

cDNA was generated using the oligodT protocol of SuperScript™ II Invitrogen® (Invitrogen, Carlsbad, CA). PCR products were generated from the *APC* cDNA and MUTYH cDNA (primers available on request). The screening of large rearrangement in MUTYH gene was performed through cDNA analysis. Because of the small length of the coding sequence, we performed the amplification of the total cDNA (1770 bp) and looked

Table 1. Clinical features of unexplained polyposis patients

Characteristics	Adenomatous polyposis	Unclassified polyposis
Patients	25	13
Age at the time of diagnosis	44 years ± 16	51 years ± 20
More than 100 polyps	4 (16%)	3 (23%)
Familial history of polyposis	7 (28%)	5 (38%)
Familial history of CRC	12 (48%)	6 (46%)
Upper gastrointestinal polyposis	10 (40%)	0 (0%)
Histological types of polyps		
Adenomatous	25 (100%)	9 (69%)
Serrated	0 (0%)	6 (46%)
Hyperplastic	0 (0%)	11 (85%)

CRC, colorectal cancer.

for a difference in the length of the amplicon performing the migration of the PCR products on a 1% ethidium bromide-stained agarose gel. cDNA analysis was performed on the *APC* mRNA to study the effect of an unclassified variant (c.1549-8A>C) on splicing.

Results

Patients

Thirty-eight patients were included. We included all individuals with more than 40 adenomas ($n = 25$) or more than 20 polyps (i.e. serrated, hyperplastic or association of several types of polyps) ($n = 13$) and with no identified mutation in *APC* or *MUTYH* gene. Table 1 presents clinical features of the patients. The mean age at diagnosis was 46 (±18). Seven patients (18%) had more than 100 adenomas. A family polyposis history was present in 32% and a family history of CRC was present in 47% of the patients.

Complement analysis for *APC* and *MUTYH* genes for point mutation and large rearrangement

For three patients with few polyps and a later age at diagnosis, screening of only *MUTYH* was previously performed. For these patients, we completed the analysis with point mutation and large rearrangement of the *APC* gene. For one patient with FAP-marked phenotype, MLPA was performed for the second time and identified an exon 14 deletion that had been missed in previous experiments. The new MLPA experiment was performed using a new set of probes provided by the manufacturer.

Large rearrangement analysis was performed on cDNA for *MUTYH* gene ($n = 6$). Electrophoresis migration of cDNA *MUTYH* was performed for six patients with available RNA. PCR amplification from cDNA *MUTYH* resulted in a predicted product of 1770 bp in all samples. We did not find any large rearrangement of the *MUTYH* gene.

Search for *APC* mosaicism

The second reading of *APC* sequences of 17 patients with no family history, selected as potential mosaic, revealed one frameshift mutation c.3202_3205del (p.Ser1068GlyfsX57). This mutation was not a mosaic, but had been missed because of a too short overlap of the previous set of primers designed for this fragment of exon 15. Second, among the 17 patients in whom a mosaicism was searched, one case was found: a 36-year old male with mandible osteomas and multiple colorectal adenomas had a 5-year old daughter with mandible osteomas. Direct sequencing of *APC* in the daughter found the c.4666del (p.Thr1556LeufsX9) mutation. This deletion was not identified in the first sequencing of the proband *APC* gene, but we found a weak signal in the father's sequence corresponding to mutation (Fig. 1). The low level of the mutated allele was confirmed for the proband by analysis of other tissues. Colon adenomas, normal colonic cells and cell of the oral epithelium were available. The mutation was hardly detectable, in the DNA extracted from normal colonic cells and cells of the oral epithelium but the mutation was found at a higher level in DNA extracted from adenomas compared to leucocyte DNA (Fig. 2).

The HRM method was also tested on this patient and his daughter. Several dilutions of mutated DNA in wild-type DNA showed that the HRM technique allows a 2.5% proportion of the mutated allele to be detected (Fig. 3) and was better than sequencing. HRM mutated DNA dilution analysis determined that cells with the mutated *APC* allele represent 2.5% in father's leucocytes DNA.

Analysis of *APC* splice mutation

An *APC* intronic variant c.1549-8A>C had been previously identified. *In silico* predictions [Splice Site Finder (http://violin.genet.sickkids.on.ca/~ali/splicesitefinder.html), MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html), NNsplice (http://www.fruitfly.org/seq_tools/splice.html), Gene Splicer (http://www.tigr.org/tdb/GeneSplicer/gene_spl.html)]of splicing did not support a deleterious splicing, but to address the possible molecular consequence of this intronic variant on mRNA splicing, PCR was performed

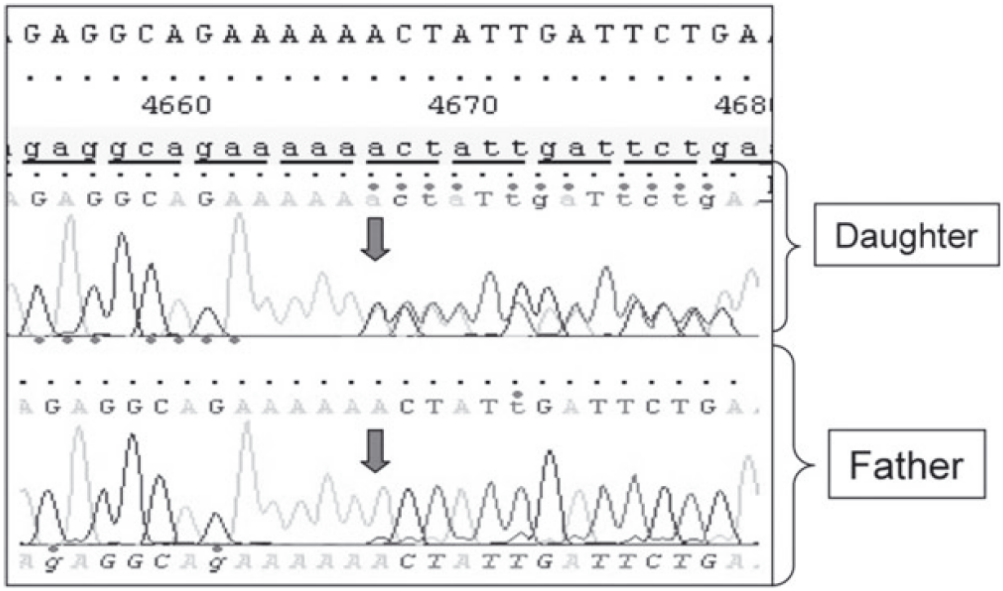


Fig. 1. Sequence analysis showing the APC mutation c.4666delA in genomic DNA from blood of the father (mosaicism) and his daughter.

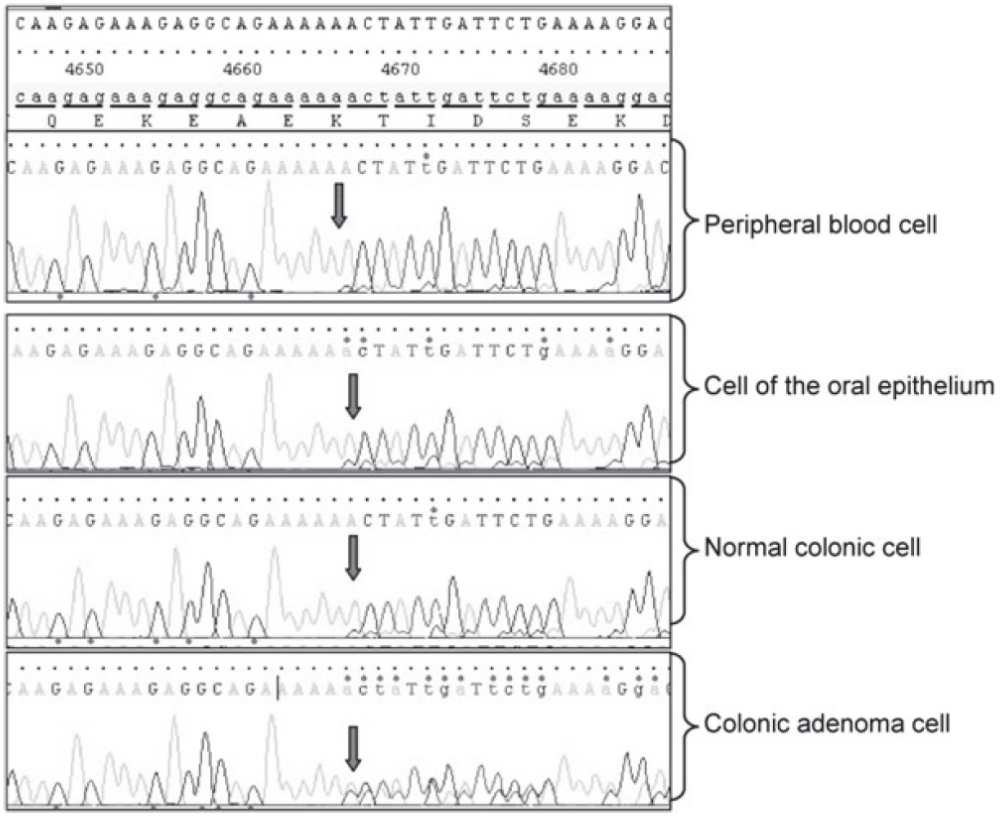


Fig. 2. Comparison of sequencing results obtained from blood, normal colonic cells, oral epithelium cells and colonic adenomas.

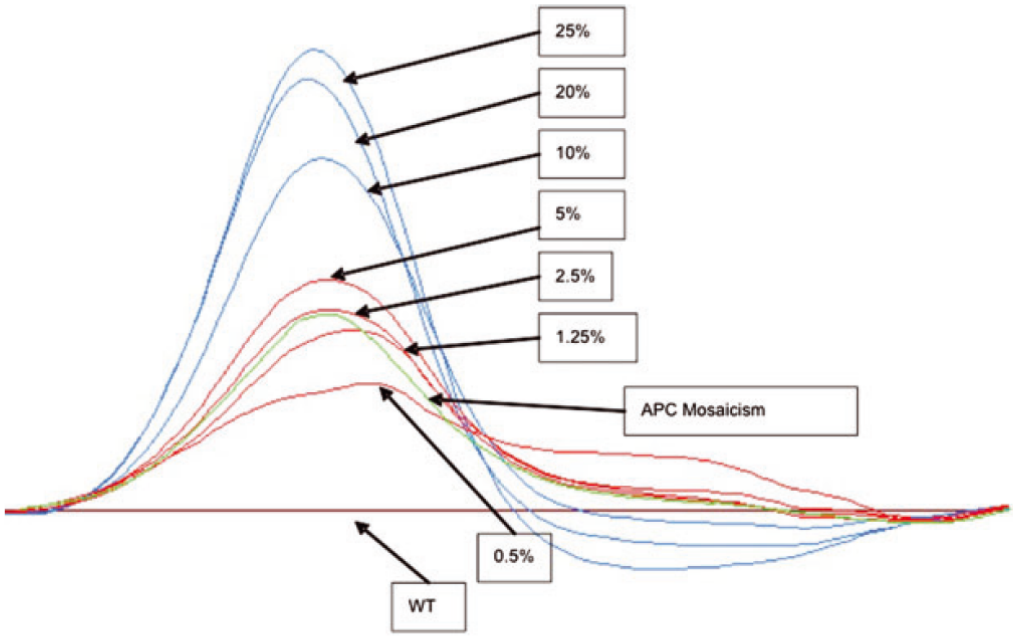


Fig. 3. Sensitivity test of the high-resolution melting (HRM) technique with dilutions of wild-type and mutated DNA (APC mutation c.4666delA) and quantification of the mosaic level in peripheral blood cell of the father.

on the *APC* cDNA. Characterization of the cDNA from this mutant *APC* allele revealed a longer mRNA in which seven nucleotides was added to exon 12 through a cryptic acceptor splice site. This insertion created a frameshift and the introduction of a premature stop codon (Fig. 4).

Mutational analysis of two TGF-β pathway genes

Screening of *SMAD4* and *BMPR1A* genes in adenomatous polyposis ($n = 25$) showed a deleterious duplication c.218dupA in *BMPR1A* (p.Asn73Lysfs X2). Initial histology of this case reported some inflammatory polyps associated to the adenomas. A second analysis by another pathologist revealed that these polyps are compatible with juvenile polyposis. We did not find any large genomic rearrangement of *SMAD4* and *BMPR1A* gene.

Mutational analysis of *SMAD4* and *BMPR1A* on 13 patients with unexplained serrated polyposis showed a pathogenic mutation c.1264_1265delinsA (p.Pro422MetfsX14) in the *SMAD4* gene. The first histological analysis found some unclassified polyps. A second analysis by another pathologist revealed that these polyps were compatible with juvenile polyposis.

Mutational analysis of four Wnt pathway genes

Mutation analysis of four genes involved in Wnt signalling pathway, *SFRP1*, *WIF1*, *AXIN2* and *PPP2R1B*, showed no deleterious mutation in the

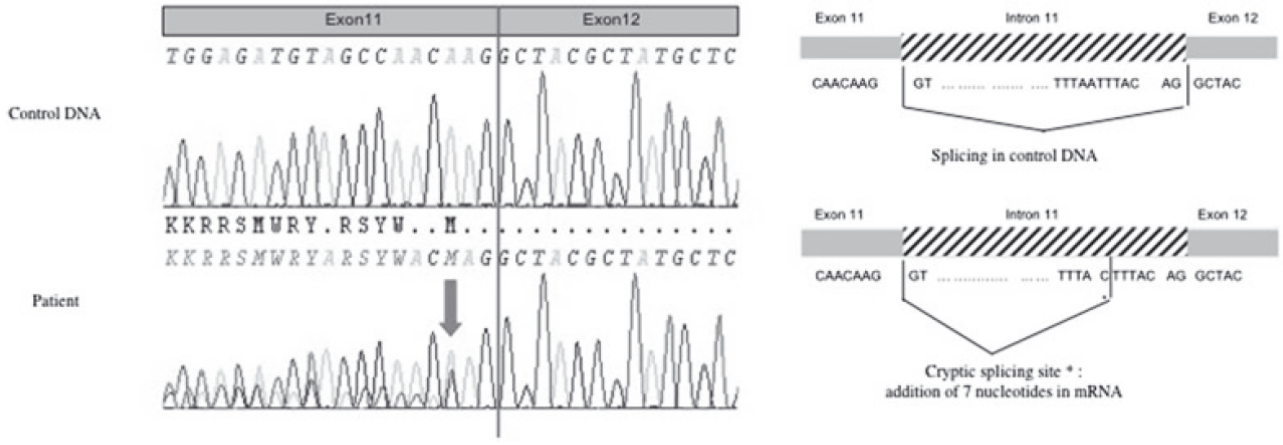


Fig. 4. cDNA sequence analysis (left) and schematic representation of aberrant splicing (right) due to the APC variant c.1549-8A>C.

set of 25 patients. A rare variant c.1235A>G (p.Asn412Ser) was found on the *AXIN2* gene. We tested this variant on 275 healthy control subjects with the HRM method and sequencing of variant profiles and it was identified in four healthy subjects (allele frequency = 0.72%) showing that this variant is a polymorphism.

Discussion

FAP and MAP syndromes are the two main causes of hereditary colorectal polyposis. About 20% of polyposis remains genetically unexplained with no mutation identified in the *APC* or *MUTYH* gene. We have defined two groups of unexplained polyposis: unexplained adenomatous polyposis and serrated unexplained polyposis (hyperplastic, sessile serrated and mixed polyposis). A combination of extensive mutation screening in *APC* and *MUTYH* genes and screening of other candidate genes for polyposis predisposition was performed in 38 unexplained polyposis patients. Six pathogenic mutations were found: four mutations in the *APC* gene: a frameshift, a large deletion, a frameshift mosaic mutation, and a splice mutation in *APC* gene with the demonstration of the pathogenic effect of the c.1549-8A>C variant, two mutations in genes of TGF-β pathway: a frameshift mutation in the *BMPR1A* gene and a frameshift mutation in the *SMAD4* gene. Overall, an additional causal mutation of polyposis was identified in 17% (6/38) of the cases selected in our study. All mutations were identified in familial cases of polyposis (even in the mosaic case with associated signs in the family) whereas 32% of included patients had familial cases of polyposis.

These results highlight the importance of complementary explorations for patients with typical polyposis. The results emphasize the need to upgrade molecular analysis of genes when the phenotype is evocative, and if significant improvements have been made on molecular method.

cDNA analysis is important for the molecular diagnosis of adenomatous polyposis, in particularly for assessing the functionality of rare variant identified in genomic DNA. Several cases of splice variants have already been described for *APC* and *MUTYH* (20–23). We showed, in our study, that *APC* intronic variant c.1549-8A>C is pathogenic even if *in silico* splice sites are unable to predict deleterious effect.

Mosaicism might underlie a fraction of sporadic FAP cases. It is estimated that 10–20% of *de novo* cases of FAP are mosaic (8, 9). Mosaicism might also underlie false-negative results in *APC* mutation screening because of the mutation detection level by usual techniques of screening (mainly sequencing). Mutations that appear *de novo* could represent germinal mosaicism in one of the parents. It is mandatory to test the offspring if there are affected individuals as the mutation is easier to detect. For this reason, testing a younger affected generation, if available, is preferable to testing the first affected generation, even if the polyposis phenotype is not yet completely expressed in the second generation, but a mutation carrier status is likely due to the presence of some symptoms as in our case (mandible osteomas). We determined that the proband had 2.5% of mutated cells in tested tissues (lymphocytes, normal colonic cells, and oral epithelium). Because the mutation was easily detectable with DNA extracted from colon adenomas as previously described in other case of *APC* mosaicism (9), frozen adenomas shall be conserved by the pathologist in all polyposis in order to perform DNA analysis if necessary.

Direct sequencing is not the best technique for mosaicism detection as it has a sensitivity of 10–20% for detecting a mutated allele (8). Prescreening techniques seem to be more sensitive than direct sequencing in the mosaicism mutation detection (8). We confirmed that the HRM method was more efficient in this application (detection threshold at 2.5–5% for the mutated allele). Another study evaluated HRM sensitivity to 5% (24). The mutation analysis of genes with high proportion of *de novo* mutation should be performed by pre-screening technique like HRM.

We did not find any large rearrangement of *MUTYH* gene in cases with available RNA. In the literature, no case of large genomic rearrangement of *MUTYH* has been described yet (25, 26). The structure of the gene and the small introns length do not favour the occurrence of large rearrangement. The search for new candidate tumour suppressor genes by exploration of the Wnt pathway was not successful. No mutation was identified on *SFRP1*, *WIF1*, *PPP2R1B* and *AXIN2*. Sequencing of *AXIN2* found a rare variant c.1235A>G (p.Asn412Ser) that we classed as polymorphism by studying a control healthy population. This variant has already been identified in a patient with unknown adenomatous polyposis and also classified as non-pathogenic by the authors on conservation species and physicochemical arguments (27). Moreover, the patient did not have dental abnormalities as they have been described associated to the polyposis due to this gene (27, 28). However, the exploration of 82 CRC and 27 polyposes, did not find any mutation on *AXIN2* gene (29, 30). The germline *AXIN2* mutation is probably a

rare event responsible of polyposis associated with severe tooth agenesis.

Screening of the TGF-β superfamily members, identified a pathogenic mutation in *BMPR1A* on a previously identified patient with an adenomatous polyposis and an *SMAD4* mutation on a patient with a first histological diagnosis of serrated polyposis. Mutations in the *BMPR1A* or *SMAD4* gene are already known to predispose to juvenile polyposis. These results are consistent with second histological examination of the specimen of both cases. Some inflammatory or unclassified polyps are compatible with juvenile polyps. The histological classification of polyposis is important because it points genetic investigations. Referral for molecular genetic testing should be performed in a targeted manner after appropriate endoscopic and histological workup has been carried out. These results confirm that pathological examination is critical as already observed by Sweet et al. (31). If nosological doubt remains, a second look on the polyps from a skilled pathologist is recommended.

After these investigations, some polyposis still remains genetically unexplained. Other genes (of the Wnt signalling pathway or TGF-β/bone morphogenic protein (BMP) pathway) may be involved and several other candidates can be proposed. For example, *MCC* (mutated in colorectal cancer) had already been described in colorectal tumourigenesis (32) or *PTEN*. Whole exome next generation sequencing could be performed in familial cases of unexplained polyposis. Mechanisms other than coding region mutation can be involved: germline epigenetic inactivation, miRNA, low frequency of susceptibility variants (33–36).

We have to report some bias in this work. The number of patients is relatively low even if such unexplained polyposis patients are very rare. We did not screen the *MUTYH* gene in the group of patients with an unexplained polyposis but mutations in this gene have not been found in a recent work on serrated polyposis (31).

A suspected diagnosis based on clinical and histological features is a requirement for rational, targeted testing for mutation. Misinterpretation of histological findings can misled genetic testing. Demonstration of a causal mutation is essential for the differential diagnosis, assessment of the risk of recurrence, monitoring, and genetic counselling (pre-symptomatic diagnosis). Close collaboration between human geneticists, pathologists and gastroenterologists is necessary for genetic predisposition to be identified.

References

1. Half E, Bercovich D, Rozen P. Familial adenomatous polyposis. Orphanet J Rare Dis 2009: 4: 22.
2. Fearnhead NS, Britton MP, Bodmer WF. The ABC of APC. Hum Mol Genet 2001: 10: 721–733.
3. Bisgaard ML, Fenger K, Bulow S, Niebuhr E, Mohr J. Familial adenomatous polyposis (FAP): frequency, penetrance, and mutation rate. Hum Mutat 1994: 3: 121–125.
4. Aretz S, Uhlhaas S, Caspari R et al. Frequency and parental origin of *de novo* APC mutations in familial adenomatous polyposis. Eur J Hum Genet 2004: 12: 52–58.
5. Mandl M, Kadmon M, Sengteller M et al. A somatic mutation in the adenomatous polyposis coli (APC) gene in peripheral blood cells – implications for predictive diagnosis. Hum Mol Genet 1994: 3: 1009–1011.
6. Farrington SM, Dunlop MG. Mosaicism and sporadic familial adenomatous polyposis. Am J Hum Genet 1999: 64: 653–658.
7. Davidson S, Leshanski L, Rennert G, Eidelman S, Amikam D. Maternal mosaicism for a second mutational event – a novel deletion – in a familial adenomatous polyposis family harboring a new germ-line mutation in the alternatively spliced-exon 9 region of APC. Hum Mutat 2002: 19: 83–84.
8. Hes FJ, Nielsen M, Bik EC et al. Somatic APC mosaicism: an underestimated cause of polyposis coli. Gut 2008: 57: 71–76.
9. Aretz S, Stienen D, Friedrichs N et al. Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). Hum Mutat 2007: 28: 985–992.
10. Al-Tassan N, Chmiel NH, Maynard J et al. Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectal tumors. Nat Genet 2002: 30: 227–232.
11. Lefevre JH, Parc Y, Svrcek M et al. APC, MYH, and the correlation genotype-phenotype in colorectal polyposis. Ann Surg Oncol 2009: 16: 871–877.
12. Lefevre JH, Rodrigue CM, Mourra N et al. Implication of MYH in colorectal polyposis. Ann Surg 2006: 244: 874–879; discussion 879–880.
13. Zbuk KM, Eng C. Hamartomatous polyposis syndromes. Nat Clin Pract Gastroenterol Hepatol 2007: 4: 492–502.
14. Lipton L, Sieber OM, Thomas HJW et al. Germline mutations in the TGF-beta and Wnt signalling pathways are a rare cause of the "multiple" adenoma phenotype. J Med Genet 2003: 40: e35.
15. Boparai KS, Mathus-Vliegen EM, Koornstra JJ et al. Increased colorectal cancer risk during follow-up in patients with hyperplastic polyposis syndrome: a multicentre cohort study. Gut 2010: 59: 1094–1100.
16. Aretz S. The differential diagnosis and surveillance of hereditary gastrointestinal polyposis syndromes. Dtsch Arztebl Int 2010: 107: 163–173.
17. Cao X, Eu KW, Kumarasinghe MP et al. Mapping of hereditary mixed polyposis syndrome (HMPS) to chromosome 10q23 by genomewide high-density single nucleotide polymorphism (SNP) scan and identification of BMPR1A loss of function. J Med Genet 2006: 43: e13.
18. O'Riordan JM, O'Donoghue D, Green A et al. Hereditary mixed polyposis syndrome due to a BMPR1A mutation. Colorectal Dis 2010: 12: 570–573.
19. Cheah PY, Wong YH, Chau YP et al. Germline bone morphogenesis protein receptor 1A mutation causes colorectal tumorigenesis in hereditary mixed polyposis syndrome. Am J Gastroenterol 2009: 104: 3027–3033.
20. Tuohy TM, Done MW, Lewandowski MS et al. Large intron 14 rearrangement in APC results in splice defect and attenuated FAP. Hum Genet 2010: 127: 359–369.
21. Neklason DW, Solomon CH, Dalton AL, Kuwada SK, Burt RW. Intron 4 mutation in APC gene results in splice

defect and attenuated FAP phenotype. Fam Cancer 2004: 3: 35–40.
22. Varesco L, Gismondi V, Presciuttini S et al. Mutation in a splice-donor site of the APC gene in a family with polyposis and late age of colonic cancer death. Hum Genet 1994: 93: 281–286.
23. Fostira F, Papademitriou C, Efremidis A, Yannoukakos D. An in-frame exon-skipping MUTYH mutation is associated with early-onset colorectal cancer. Dis Colon Rectum 2010: 53: 1197–1201.
24. Vossen RH, Aten E, Roos A, den Dunnen JT. High-resolution melting analysis (HRMA): more than just sequence variant screening. Hum Mutat 2009: 30: 860–866.
25. Peterlongo P, Mitra N, Sanchez de Abajo A et al. Increased frequency of disease-causing MYH mutations in colon cancer families. Carcinogenesis 2006: 27: 2243–2249.
26. Kanter-Smoler G, Bjork J, Fritzell K et al. Novel findings in Swedish patients with MYH-associated polyposis: mutation detection and clinical characterization. Clin Gastroenterol Hepatol 2006: 4: 499–506.
27. Lejeune S, Guillemot F, Triboulet JP et al. Low frequency of AXIN2 mutations and high frequency of MUTYH mutations in patients with multiple polyposis. Hum Mutat 2006: 27: 1064.
28. Lammi L, Arte S, Somer M et al. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. Am J Hum Genet 2004: 74: 1043–1050.
29. Renkonen ET, Nieminen P, Abdel-Rahman WM et al. Adenomatous polyposis families that screen APC mutation-negative by conventional methods are genetically heterogeneous. J Clin Oncol 2005: 23: 5651–5659.
30. Peterlongo P, Howe LR, Radice P et al. Germline mutations of AXIN2 are not associated with nonsyndromic colorectal cancer. Hum Mutat 2005: 25: 498–500.
31. Sweet K, Willis J, Zhou XP et al. Molecular classification of patients with unexplained hamartomatous and hyperplastic polyposis. JAMA 2005: 294: 2465–2473.
32. Fukuyama R, Niculaita R, Ng KP et al. Mutated in colorectal cancer, a putative tumor suppressor for serrated colorectal cancer, selectively represses beta-catenin-dependent transcription. Oncogene 2008: 27: 6044–6055.
33. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 2004: 36: 497–501.
34. Wang X, Lam EK, Zhang J, Jin H, Sung JJ. MicroRNA-122a functions as a novel tumor suppressor downstream of adenomatous polyposis coli in gastrointestinal cancers. Biochem Biophys Res Commun 2009: 387: 376–380.
35. Fearnhead NS, Wilding JL, Winney B et al. Multiple rare variants in different genes account for multifactorial inherited susceptibility to colorectal adenomas. Proc Natl Acad Sci U S A 2004: 101: 15992–15997.
36. Wong HL, Peters U, Hayes RB et al. Polymorphisms in the adenomatous polyposis coli (APC) gene and advanced colorectal adenoma risk. Eur J Cancer 2010: 46: 2457–2466.

Comprehensive assessment of variation at the transforming growth factor β type 1 receptor locus and colorectal cancer predisposition

Luis G. Carvajal-Carmona[a,1], Mike Churchman[a], Carolina Bonilla[b,c], Axel Walther[a], Jeremie H. Lefèvre[b,c], David Kerr[c], Malcolm Dunlop[d], Richard Houlston[e], Walter F. Bodmer[b,c,1], and Ian Tomlinson[a]

[a]Molecular and Population Genetics Laboratory, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom; [b]Cancer and Immunogenetics Laboratory, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DZ, United Kingdom; [c]Department of Clinical Pharmacology, University of Oxford, Oxford OX3 7DQ, United Kingdom; [d]Colon Cancer Genetics Group, Medical Research Council, Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, United Kingdom; and [e]Section of Cancer Genetics, Institute of Cancer Research, Sutton SM2 5NG, United Kingdom

Contributed by Walter F. Bodmer, March 15, 2010 (sent for review February 1, 2010)

The role of transforming growth factor β receptor type 1 (*TGFBR1*) polymorphisms, particularly a coding CGC insertion (rs11466445, *TGFBR1*6A/9A*) in exon 1, has been extensively investigated in regard to colorectal cancer (CRC) risk. These investigations have generated conflicting results. More recently, allele-specific expression (ASE) of *TGFBR1* mRNA has been suggested as predisposing to CRC, with a relative risk of nearly 10-fold and a population attributable risk of ∼10%. Owing to the potential importance of *TGFBR1* variants in CRC, we performed a comprehensive examination of tagging SNPs at and around the gene in 3,101 CRC cases and 3,334 controls of northern European ancestry. To test whether rare or subpolymorphic *TGFBR1* variants were associated with CRC risk, we sequenced the gene's exons in a subset of patients. We also evaluated *TGFBR1* ASE in a panel of CRC cases and controls. Overall, we found no association between *TGFBR1* polymorphisms and CRC risk. The rare variant screen did not identify any changes of potentially pathogenic effects. No evidence of greater ASE in cases than controls was detected, and no haplotype around *TGFBR1* could account for the ASE reported in other studies. We conclude that neither genetic variation nor ASE at *TGFBR1* is likely to be a major CRC risk factor.

genetic susceptibility | transforming growth factor β receptor type 1 *6A/9A | candidate gene | low penetrance | TGF signaling

Transforming growth factor β receptor type 1 (TGFBR1) is a serine-threonine kinase that mediates growth-inhibiting signals from TGFB1 through a complex with TGFBR2. The TGFB pathway and the related bone morphogenetic protein (BMP) pathway play important roles in the pathogenesis of colorectal cancer (CRC) and other intestinal tumors: Inactivating *TGFBR2* mutations occur in CRCs with microsatellite instability (MSI) (1), germline mutations in *SMAD4* and *BMPR1A* predispose to juvenile polyposis (2, 3), and *SMAD4* is mutated in some MSI-negative CRCs (4). In addition, common germline variants in the BMP pathway are associated with increased risk for CRC (5). It is therefore highly plausible that germline variation at *TGFBR1* itself has effects on the risk for CRC in the general population.

Considerable attention has been focused on a common *TGFBR1* polymorphism (*TGFBR1**6A/9A, rs11466445) in exon 1 that results in the deletion of three alanines from a stretch of nine alanines. This 9-bp deletion is located within the predicted signal sequence cleavage region. Functional studies have suggested that *TGFBR1**6A responds less well than the *TGFBR1**9A allele to growth inhibitory signals of TGFB1. Some studies have suggested that individuals who carry a *TGFBR1**6A allele are at increased risk for cancers of several types, including CRC (6). Overall, however, evidence for the association with CRC is mixed and relatively small cohorts have been analyzed. The effects on cancer risk of other common polymorphisms and rare variants at *TGFBR1* have not specifically been assessed.

A second line of evidence for the importance of *TGFBR1* in predisposition to CRC has recently emerged. Valle et al. (7) reported that CRC cases show imbalanced expression of *TGFBR1* alleles compared with controls. This so-called "germline" allele-specific expression (ASE) was detected using the SNaPshot (Applied Biosystems) method to screen *TGFBR1* mRNA from peripheral blood in a panel of 138 cases and 105 controls. Although no causal variation was found, this tendency was said to be present in about 20% of CRC cases, to be dominantly inherited, and to confer a nearly 10-fold increased risk for CRC. A further report from Guda et al. (8) extended the study of Valle et al. (7) using pyrosequencing. Guda et al. (8) found ASE in only about 5% of familial CRC cases and in no sporadic CRC patients or controls.

We wished to undertake a thorough assessment of *TGFBR1* in relation to CRC risk in a large set of cases and controls. Using a total of over 3,000 CRC cases and over 3,300 controls, we genotyped haplotype-tagging SNPs from throughout the *TGFBR1* region, imputed other SNPs in the region, genotyped the *TGFBR1*6A/9A* polymorphism, and screened the coding region of the gene for rare variants with potential effects on disease risk. In parallel, we undertook ASE analysis in a subset of patients using two complementary techniques.

Results

Association Between TGFBR1*6A/9A and CRC Risk. We typed *TGFBR1*6A/9A* in 828 familial colorectal tumor cases and 913 controls from the COloRectal Gene Identification (CORGI) study. There was no association with CRC susceptibility ($P_{per\ allele}$ = 0.71, β = 0.04, SE = 0.13; Table 1). To investigate the association between this allele and CRC in two other case–control cohorts, we used IMPUTEv2 (9) to generate *in silico* genotypes, with the CORGI samples as a diploid reference (https://mathgen.stats.ox.ac.uk/impute/impute.html). The imputation quality was very good, with proper_info scores >80% (Table 1) in all cases and >98% correspondence between imputed and typed genotypes in the CORGI study. In the additional cohorts, we found no evidence for an association between *TGFBR1*6A/9A* and CRC risk ($P_{per\ allele}$ in the Scotland series = 0.07, β = 0.21; $P_{per\ allele}$ in the VQ58 series = 0.58, β = 0.05;

Author contributions: L.G.C.-C., D.K., M.D., R.H., W.F.B., and I.T. designed research; L.G.C.-C., M.C., C.B., A.W., J.H.L., and I.T. performed research; L.G.C.-C., W.F.B., and I.T. analyzed data; and L.G.C.-C., W.F.B., and I.T. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: luis@well.ox.ac.uk or walter.bodmer@hertford.ox.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/1002816107/DCSupplemental.

Table 1. Genotypes at rs11466445 in each of the three CRC case–control series

Series	Genotype counts* (9A9A/9A6A/6A6A)		*P* value and associated regression coefficient β (SE)†				Proper_info score‡
	Cases	Controls	Additive	Dominant	Recessive	Genotypic	
CORGI	746/159/8	673/145/10	0.70, 0.04 (0.13)	0.81, 0.02 (0.12)	0.49, 0.32 (0.47)	0.47, 0.01 (0.12)	—
Scotland	772/152/9	843/140/7	0.07, 0.21 (0.15)	0.07, 0.23 (0.13)	0.54, 0.31 (0.50)	0.50, 0.22 (0.13)	0.84
VQ58	938/201/13	1119/200/14	0.58, 0.05 (0.09)	0.55, 0.06 (0.10)	0.96, 0.01 (0.38)	0.38, 0.01 (0.38)	0.85
All			0.13, 0.09 (0.06)	0.13, 0.10 (0.07)	0.17, 0.18 (0.25)	0.17, 0.09 (0.07)	

Tests of association (additive, dominant, recessive, and genotypic) are shown.
*Genotypes in the CORGI study were obtained by direct genotyping. Genotypes in the Scotland and VQ58 series were imputed using IMPUTEv2 software. Counts in the Scotland and VQ58 series were derived from individual genotypes with a probability of 90% or higher.
†*P* values were derived from the additive, dominant, recessive, and genotypic disease models as implemented using logistic regression in the program SNPTEST. For each series, we present the *P* value, followed by the regression coefficient and the regression coefficient's SE (in parentheses).
‡Proper_info scores relate to the quality of the imputation. These scores contain a measure of the relative statistical information about the parameters of interest, β in this case. This measure lies in the range [0,1], with 1 indicating perfect information and 0 indicating no information. For more details on this score, see the article by Marchini et al. (12).

Table 1). A meta-analysis of the three studies found no association between *TGFBR1*6A/9A* alleles and CRC ($P_{per\ allele} = 0.12$; Fig. 1).

We then investigated whether particular haplotypes around *TGBFR1*6A/9A* could mediate CRC risk. The 6A allele was principally present on two haplotypes with frequencies of ∼1% and ∼8%. However, neither haplotype was associated with CRC risk (*SI Appendix*).

Although *TGFBR1*6A/9A* was not a major detectable CRC risk factor on its own, it might increase risk by interacting with polymorphisms in other genes. Prime candidates are four CRC SNPs that map close to the TGFB/BMP pathway genes *GREM1*, *BMP2*, *BMP4*, and *SMAD7* (5). We carried out simple multiplicative genotypic interaction tests in cases only between *TGFBR1*6A/9A* and the SNPs rs4779584, rs961253, rs4444235, and rs4939827, corresponding to each of the four risk loci above. No interactions were found ($P = 0.05$, $P = 0.49$, $P = 0.45$, and $P = 0.56$, respectively; *SI Appendix*).

Association Between Other SNPs Near TGFBR1 and CRC. We genotyped up to 15 tagging SNPs spanning the ∼75-kb haplotype block containing the *TGFBR1* locus (chr9: 100,897,875–100,973,680) in 3,125 CRC cases and 3,372 controls, all of northern European ancestry. To generate additional genotypes in this block, we used the genotype data from the Centre d'Etude du Polymorphisme Humain (CEPH) samples ($n = 63$) obtained by the 1000genomes project (http://browser.1000genomes.org/index.html). This dataset contained nearly five times more SNPs and lower frequency variants than those present in release 22 of the Hapmap project (260 SNPs vs. 56 SNPs), including a much larger fraction of variants with MAF < 0.01 (97 vs. 1), with 0.01 < MAF < 0.05 (67 vs. 1), and with MAF > 0.05 (96 vs. 54). Large proportions of these variants (91%, 74%, and 27%, respectively) were not reported in database of single nucleotide polymorphisms (dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/). The use of this reference population for imputation permitted a comprehensive evaluation of common variants and some rarer variants close to *TGFBR1* for CRC susceptibility. Genotype imputation was successful for 89–97 variants depending on the study cohort (differences in SNP array genotyping are discussed in *Materials and Methods*); most variants had MAF > 0.05 (*SI Appendix*), bringing the total variants evaluated in this haplotype block to 102.

We found no evidence of association between SNPs in the *TGFBR1* haplotype block and CRC risk in the VQ58 and Scottish cohorts. Two SNPs had nominally significant associations in the CORGI study (9-100973331, $P = 0.018$; 9-100973680, $P = 0.019$; *SI Appendix*), but none of these remained significant after correction for multiple testing. An association meta-analysis of the three cohorts found no evidence that any SNP was significantly associated with CRC (*SI Appendix*). Therefore, common variation at the *TGFBR1* locus is unlikely to be associated with CRC risk in populations of northern European ancestry.

To investigate whether long-range regulation could mediate the association between *TGFBR1* and CRC risk and explain mechanisms such as ASE, we extended our analysis to ∼500 kb on each side of the *TGFBR1* haplotype block (chr9: 100,385,599–101,513,263). We found very little evidence for an association between SNPs in this region and CRC risk after multiple testing correction (*SI Appendix*). However, one marker (rs410180), located ∼100 kb downstream of *TGFBR1*, had a suggestive association ($P_{per\ allele} = 0.00008$, $\beta = -0.24555$, $SE = 0.06$) in all three sample sets combined. We were intrigued by this possible association between rs410180 and CRC risk and decided to obtain genotypes from two additional cohorts that included ∼3,000 sporadic cases and ∼3,000 controls from England and ∼2,000 sporadic cases and ∼2,000 sporadic controls from the Scotland series. We found no support for the association between rs410180 and disease in these latter cohorts ($P_{per\ allele} = 0.15$ and $P_{per\ allele} = 0.59$, respectively; meta-analysis for all five cohorts: $P = 0.003$, odds ratio = 1.13, 95% confidence interval: 1.04–1.22), suggesting that rs410180 is unlikely to be a CRC susceptibility variant.

A further possibility of long-range regulation of *TGFBR1* expression could involve multimarker haplotypes rather than alleles. To examine this possibility, we used the Haploview program (10) to investigate associations between haplotypes in this region and disease risk. The distal and proximal flanking regions contained 19 and 12 haplotype blocks defining 82 and 45 haplotypes, respectively. We found no haplotypes with significant associations after correction for multiple testing (lowest nominal *P* value = 0.007). Therefore, neither alleles nor haplotypes at or around *TGBFR1* are likely to be associated with CRC risk.

Association Between Other TGFB Pathway Genes and CRC Risk. In previous studies, we have shown that alleles at a number of TGFB pathway genes are associated with CRC risk. These genes include *SMAD7, GREM1, BMP2*, and *BMP4* (5,11). It is thus possible that other genes in this pathway interact with *TGFBR1* to increase risk for cancer. To investigate this possibility, we examined the association between CRC and alleles at and around *TGFB1*, *TGFB2*, *TGFB3*, *TGFBR2*, and *TGFBR3* in our three cohorts (*SI Appendix*). Overall, we found very little evidence for an association between these genes and CRC, with the lowest *P* values being at two *TGFBR2* intragenic SNPs: 3-30632242 ($P = 0.00083$, $\beta = 0.035$) and 3-30644255 ($P = 0.00099$, $\beta = 0.43$).

Although no SNPs in TGFB pathway genes examined in the present study were individually associated with CRC, we tested epistasis between genotyped SNPs in the *TGFBR1*-containing

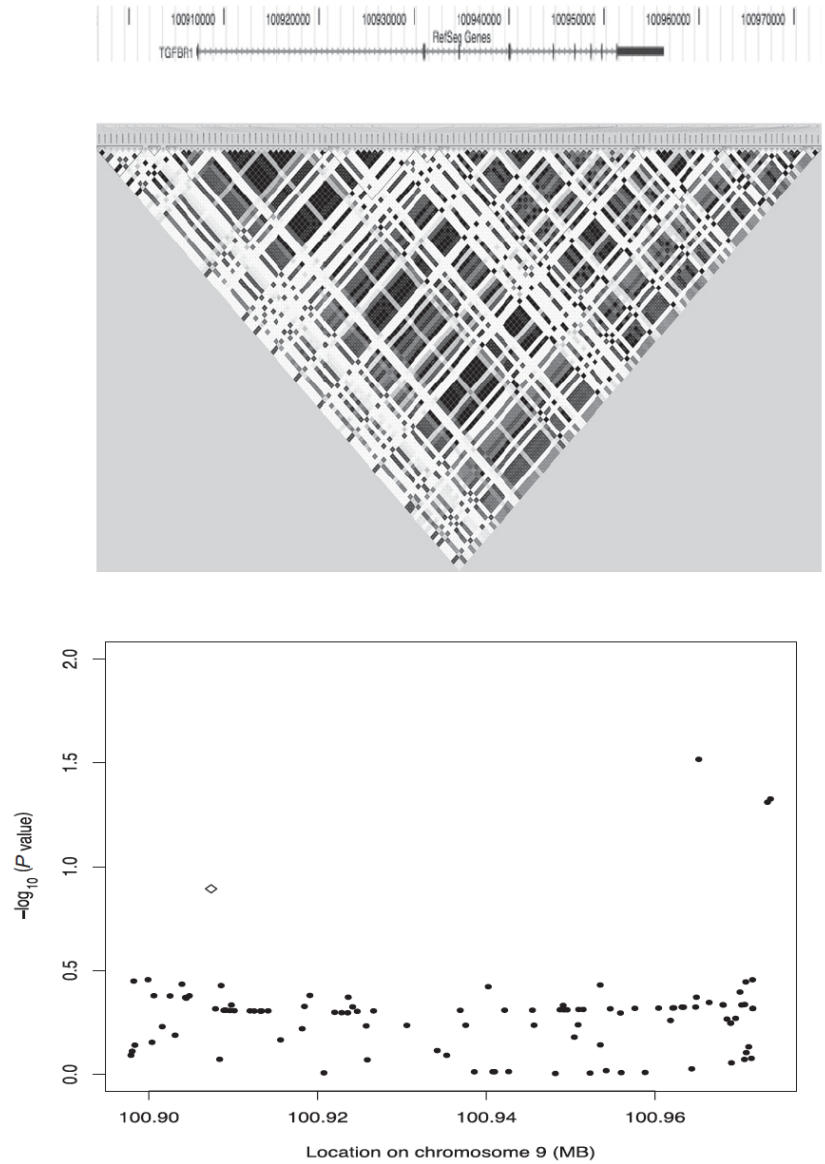


Fig. 1. Association between SNPs in the TGFBR1 locus haplotypes and CRC risk. Approximate location of *TGFBR1* (*Top*), linkage disequilibrium patterns (r^2) in the CEPH 1000genome data (*Middle*), and $-\log_{10}$ (*P* values) for the allelic test across the region (*Bottom*) are shown.

haplotype block and SNPs genotyped close to *TGFB1*, *TGFB2*, *TGFB3*, *TGFBR2*, and *TGFBR3*. We failed to detect any significant interaction at $P = 0.001$, suggesting that epistasis between *TGFBR1* and these five closely-related genes is unlikely to mediate CRC risk.

Coding Variation at TGFBR1. We screened the entire coding sequence of *TGFBR1* in 96 CRC cases from the CORGI cohort. All these patients had at least one first-degree relative with CRC and did not have mutations in known highly penetrant CRC genes. We found only one change in the *TGFBR1* coding region, a synonymous serine-to-serine change at residue 39 (exon 2). We failed to identify the Tyr401Asn change reported by Valle et al. (7). An examination of the 1000genomes CEPH genotype data did not reveal the existence of this change or any other *TGFBR1* nonsynonymous variant. Thus, we found no evidence of uncommon disease-associated variants in the coding region of *TGFBR1*, suggesting that these do not represent a major risk factor for CRC.

Germline ASE. We examined ASE in lymphoblastoid cell line-derived cDNA at *TGFBR1*6A/9A* (located in the 3′ UTR of the gene) and in a 5′-UTR marker (rs1590) previously examined by Valle et al. (7). Twenty-five individuals (8 patients and 17 controls) were informative (heterozygous) at *TGFBR1*6A/9A*, and 44 individuals (16 patients and 28 controls) were informative at rs1590.

We initially searched for any evidence of ASE as a rare trait by looking for outliers in terms of their relative allelic expression of *TGFBR1* alleles (Fig. 2). One control showed highly biased expression at *TGFBR1*6A/9A*, and a single case showed similarly biased expression at rs1590. However, apparently biased allelic dosages were also seen at rs1590 in genomic DNA (gDNA) from one control and two cases. In the absence of constitutional copy number variation at *TGFBR1* (http://genome.ucsc.edu/ and http://projects.tcag.ca/variation/), it seemed most likely that the outlying individuals resulted from inherent occasional problems with quantitative genotyping.

Despite our lack of clear evidence for a discrete category of ASE, we assessed whether ASE, in the sense of the degree of

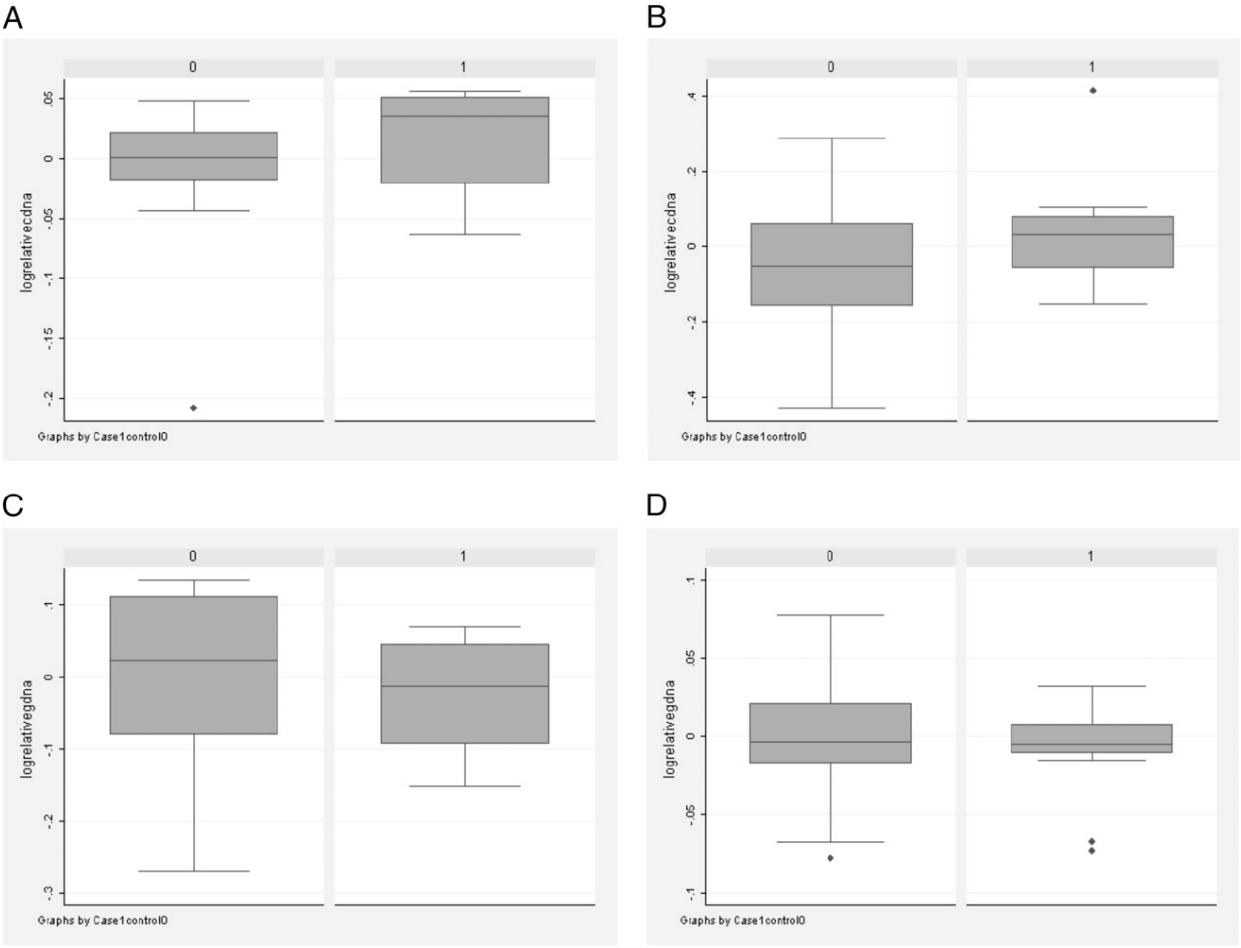


Fig. 2. Box-and-whiskers plots showing log (allelic dosage) at rs11466445 and rs1590 relative to the geometrical mean for that polymorphism in all samples. Dosage is shown for cases and controls at TGFBR1*6A/9A in cDNA (*A*), rs1590 in cDNA (*B*), TGFBR1*6A/9A in gDNA (*C*), and rs1590 in gDNA (*D*). Note the outlying samples shown by diamonds.

biased allelic expression, was more prevalent in our cases than in controls. There was a nominally significant difference between the cases and controls ($P = 0.028$, Wilcoxon test; $t = -2.09$ and $P = 0.040$, t test), but the trend to ASE was actually more pronounced in the controls (mean = +0.027 deviation from mean, SD = 0.12) than in the cases (mean = −0.041, SD = 0.13).

In the above analysis, we reasoned that ASE should be assessed relative to the hypothetical 1:1 ratio of allelic dosage rather than by comparing cDNA and gDNA dosage from a single sample, essentially so that experimental noise was minimized. However, we also replicated the tests performed by Valle et al. (7) and Guda et al. (8) by calculating the dosage ratio, $(A_{cDNA}/B_{cDNA})/(A_{gDNA}/B_{gDNA})$, for alleles A and B in each individual. Nineteen individuals showed ASE according to the criteria of Valle et al. (7), that is, ratios >1.5, or <0.67. Of these, 7 were patients and 12 were controls, demonstrating a nonsignificant difference ($\chi^2_1 = 0.049$, $P = 0.83$). Quantitative analyses also showed no significant difference in ratios between cases and controls ($P = 0.09$, Wilcoxon test; $P = 0.13$, t test).

Discussion

Despite its excellent candidate gene status and previous reports, we found no evidence to show that common or rare genetic variants at or near *TGFBR1* are associated with the risk for CRC. There was, moreover, no evidence of a disease-associated haplotype in this region or that variants near *TGFBR1* had any effect on risk in association with other variants in the TGFB pathway. Although we did not specifically test for common copy number variants, there is very little evidence of such variation in public databases. Our study was empowered to detect common alleles with moderate or greater effects (e.g., >80% power to find an allele with a frequency of 0.3 with an additive 1.1-fold effect on disease risk at $P = 0.1$).

We failed to confirm the reported finding of higher levels of ASE at *TGFBR1* in CRC cases than controls, and we did not identify any disease-associated haplotype that could account for the previously reported ASE at this site. The question remains as to whether ASE occurs at all at *TGFBR1*. In our data, we observed no greater variation in relative allelic dosage in cDNA than in gDNA, suggesting that it is difficult to find rare ASE events using the techniques employed.

To date, hypothesis-free genome-wide association (GWA) studies in colorectal and other cancers have identified tens of variants associated with differential risk for disease. The genes involved act in multiple different pathways, although there is accumulating evidence that the TGFB/BMP pathway plays a central role in CRC predisposition. The prior focus of research groups on *TGFBR1* as a candidate gene has therefore been vindicated by GWA approaches. The absence of detectable CRC-associated genetic variation at *TGFBR1* might result from

variants of very weak effects or low frequency but may also result from chance or may reflect relatively strong natural selective constraints owing to the central importance of the type 1 TGFB receptor in development and tissue maintenance.

Materials and Methods

Study Samples. We used three CRC case–control series based on samples of northern European ancestry. The first series was ascertained in England through the CORGI consortium and included 920 familial CRC or significant adenoma cases and 929 cancer-free controls. Familial patients in the CORGI cohort were individuals with two first-degree relatives with colorectal tumors and from whom known Mendelian syndromes had been excluded. The second case–control sample comprised 1,003 early-onset Scottish CRC patients (<55 years of age) who had no mutations in the known highly penetrant genes and 979 population controls of Scottish origin. The third series comprised 1,216 samples from the post-treatment stage of a Phase III, randomised, double blind, placebo-controlled study of rofecoxib (VIOXX) in colorectal cancer patients following potentially curative therapy (VICTOR) (n = 920; http://www.octo-oxford.org.uk/alltrials/infollowup/vic.html) and from the multicentre international study of capecitabine ± bevacizumab as adjuvant treatment of colorectal cancer (QUASAR2) (n = 356; http://www.octo-oxford.org.uk/alltrials/trials/q2.html). In the latter study, we used publicly available control genotype data from 1,437 individuals belonging to the UK 1958 Birth Cohort. We refer to these three series as the CORGI, Scotland, and VQ58 series (5). Full informed consent was obtained from all individuals under the auspices of UK Research Ethics Committees. Previous analysis had shown no detectable evidence of gross population stratification or other sources of systematic bias within each of these sample sets.

Nucleic Acid Isolation and Genotyping. DNA samples were isolated from peripheral blood using standard methods and quantified with picogreen. Tagging SNPs around *TGFBR1* were typed with Illumina Hap300 (VQ58) or Hap550 (CORGI and Scotland1 series) SNP arrays; these contain 9 and 15 tagging SNPs, respectively, in the haplotype block that comprises the *TGFBR1* coding region (chr9: 100,897,000–100,973,999). Duplicate samples were used to check genotyping quality. General quality control assessment was as previously described, and all SNPs and samples described herein passed the required thresholds (5). The *TGFBR1*6A/9A* SNP, rs11466445, was genotyped using standard PCR conditions and the following primers: 5′-GAGGTTTGCTGGGGTGAG and 5′-AGCAGGAGCGAGCCAGAG. PCR products were run on an ABI3730XL sequencer (Applied Biosystems), and genotypes were read using GeneMapper (Applied Biosystems). For a subset of CORGI cases and controls, lymphoblastoid cell lines were made. cDNA was extracted from these samples using standard methods.

Exon Sequencing. We used conventional PCR sequencing to screen all protein coding exons of *TGFBR1* for mutations in 96 CORGI patients. Primers used to amplify and sequence these exons are listed in *SI Appendix*. Sequences were visualized using the 4Peaks program. Sequence changes were confirmed by independent amplification and sequence reactions.

Statistical Analyses. Genotype frequencies at each SNP were tested for deviations from Hardy–Weinberg equilibrium (HWE) and rejected at $P < 10^{-6}$. Logistic regression was used to test additive, genotypic, dominant, recessive, and genotypic models of association between genetic variants and disease. The program SNPTEST (www.stats.ox.ac.uk/~marchini/software/gwas/snptest.html) was used to obtain association P values and to estimate regression coefficients (β) in these models along with their standard errors (SEs). Epistasis and HWE test statistics were calculated using the PLINK package (12), STATA software (Stata Corporation), and SNPTEST. We used genotype data from the CEPH 1000genome samples (ftp://ftp.1000genomes.ebi.ac.uk) and IMPUTEv1 software (13) to generate *in silico* genotypes at additional common polymorphisms in and around (within 50 kb of) *TGFBR1* and selected other loci in the TGFB pathway. *In silico* genotypes at rs11466445 were generated in the Scotland and VQ58 series using reference genotype data from the CORGI cohort and the program IMPUTEv2 (8). Meta-analysis of association data was carried out with the program Meta (http://www.stats.ox.ac.uk/~marchini/software/gwas/gwas.html). Linkage disequilibrium analyses, including estimation of haplotype frequencies and haplotype association χ^2 tests, were carried out with Haploview (9). Haplotype blocks were defined with the solid spine method incorporated into Haploview. The program VCFtools (kindly provided by Adam Auton) was used to summarize genotypes and linkage disequilibrium (LD) patterns at the *TGFBR1* locus in the 1000genome CEPH samples.

ASE. *TGFBR1* ASE was examined at rs1590 and rs11466445 in 43 CRC cases and 55 controls in at least two replicates of each sample in all cases. ASE at rs1590 was examined in DNase-treated lymphoblastoid mRNA from cases and controls using the SNaPshot protocol, as also employed by Valle et al. (7). Owing to some inherent noise in the genotyping signal using this method, we initially compared dosage of the minor allele (C) with that of the major allele (A) in cDNA from informative (heterozygote) samples, based on an expectation that the underlying allelic dosage was 1:1, using nonparametric analysis. We then normalized each cDNA dosage to the geometrical mean of all the dosages derived from the full set of cDNAs. Finally, we normalized each cDNA ratio to the gDNA ratio on a per sample basis. For rs11466445, relative allelic dosages were examined using a fluorescent PCR and Genescan/Gennemmapper (Applied Biosystems) analysis of peak areas based on allelic separation by size. Other analyses were performed as for rs1590.

ACKNOWLEDGMENTS. We thank Dr. Boris Pasche for sharing the experimental conditions used in the ASE experiments and Sarah Louise West-Spain for sharing her expertise with IMPUTE. All research was approved by UK Ethics Committees and performed according to the Declaration of Helsinki. Funding was provided by Cancer Research UK, the Wellcome Trust, the Oxford Biomedical Research Centre, and the European Union FP7 Genetic study of Common HeredItary Bowel Cancers in Hispania and the Americas (CHIBCHA) project.

1. Markowitz S, et al. (1995) Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science* 268:1336–1338.
2. Howe JR, et al. (1998) Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science* 280:1086–1088.
3. Howe JR, et al. (2001) Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. *Nat Genet* 28:105–107.
4. Hahn SA, et al. (1996) DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* 271:350–353.
5. Houlston RS, et al.; Colorectal Cancer Association Study Consortium; CoRGI Consortium; International Colorectal Cancer Genetic Association Consortium (2008) Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet* 40:1426–1435.
6. Daley D, et al. (2007) Is TGFBR1*6A a susceptibility allele for nonsyndromic familial colorectal neoplasia? *Cancer Epidemiol Biomarkers Prev* 16:892–894.
7. Valle L, et al. (2008) Germline allele-specific expression of TGFBR1 confers an increased risk of colorectal cancer. *Science* 321:1361–1365.
8. Guda K, et al. (2009) Infrequent detection of germline allele-specific expression of TGFBR1 in lymphoblasts and tissues of colon cancer patients. *Cancer Res* 69: 4959–4961.
9. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 5: e1000529.
10. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265.
11. Jaeger E, et al.; CORGI Consortium (2008) Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet* 40: 26–28.
12. Purcell S, et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575.
13. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39: 906–913.