



Étude fonctionnelle de SMAP1 : un nouveau gène à la croisée du trafic vésiculaire et de l'oncogenèse

Fatiha Sangar

► To cite this version:

Fatiha Sangar. Étude fonctionnelle de SMAP1 : un nouveau gène à la croisée du trafic vésiculaire et de l'oncogenèse. Cancer. Université Pierre et Marie Curie - Paris VI, 2012. Français. NNT : 2012PA066121 . tel-00836162

HAL Id: tel-00836162

<https://theses.hal.science/tel-00836162>

Submitted on 20 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.

UNIVERSITÉ PARIS VI - PIERRE ET MARIE CURIE

THÈSE DE SCIENCES

Pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ PARIS VI - PIERRE ET MARIE CURIE

École doctorale : Complexité du vivant

Spécialité : Génétique des tumeurs

Présentée par

Fatima SANGAR

**ÉTUDE FONCTIONNELLE DE SMAP1 :
UN NOUVEAU GÈNE À LA CROISÉE DU TRAFIC
VÉSICULAIRE ET DE L'ONCOGÉNÈSE**

Dirigée par **Françoise PRAZ**

Soutenue le **Jeudi 12 Avril 2012**

Devant le Jury composé de :

Pierre NETTER

Philippe CHAVRIER

Carla OLIVEIRA

Philippe JAY

Olivier DE WEVER

Françoise PRAZ

Président

Rapporteur

Rapporteur

Examineur

Examineur

Directrice de Thèse

RESUME

La déficience du système de réparation des mésappariements de bases aboutit à une instabilité des séquences répétées ou microsatellites (MSI) qui engendre des mutations au niveau de gènes cibles de l'oncogenèse MSI.

L'objectif de ma Thèse consistait à définir les conséquences fonctionnelles des mutations d'un nouveau gène cible de la tumorigenèse colorectale MSI : le gène SMAP1 (Small ArfGAP1) qui code une protéine de la famille ArfGAP (ADP ribosylation factor GTPase Activating Protein) spécifique d'Arf6, protéine impliquée dans de nombreux mécanismes cellulaires.

Les mutations de SMAP1 sont spécifiques des tumeurs MSI de différentes origines tissulaires et n'apparaissent qu'au niveau de la répétition (A10). Dans les tumeurs colorectales primaires, la fréquence de mutations de SMAP1 observées diminue au cours de la progression tumorale suggérant que les tumeurs dépourvues de mutations de SMAP1 sont plus invasives. D'un point de vue fonctionnel, les mutations de SMAP1 ont pour conséquences un défaut dans le recyclage rapide du récepteur à la transferrine, une augmentation de la prolifération cellulaire et une diminution du pouvoir invasif en maintenant les jonctions adhérentes.

Ainsi, nos observations montrant que les mutations de SMAP1 augmentent le pouvoir prolifératif mais diminuent le pouvoir invasif des lignées cellulaires issues de CCR MSI pourraient expliquer certaines caractéristiques cliniques des CCR MSI, les tumeurs MSI étant en effet des tumeurs volumineuses, ayant un faible pouvoir métastatique.

Mots clés : cancer colorectal, système MMR, instabilité des microsatellites, gène cible, SMAP1, Arf6, endocytose clathrine-dépendante, prolifération cellulaire, invasion/migration.

ABSTRACT

FUNCTIONAL STUDY OF SMAP1: A NEW GENE LINKING VESICULAR TRAFFIC AND ONCOGENESIS

Tumours that are deficient in mismatch repair system accumulate numerous mutations throughout the genome in repeated sequences, called microsatellites. Genes that contain microsatellite in their sequence are target genes of microsatellite instability (MSI).

The aim of my thesis was to determine the functional consequences of mutations occurring in a newly identified target gene of the MSI oncogenesis : SMAP1 (Small ArfGAP1), which encodes a protein of the ArfGAP (ADP ribosylation factor GTPase Activating Protein) family specific for Arf6, a protein involved in clathrin-mediated endocytosis and actin remodelling.

SMAP1 mutations are specific for MSI tumours of various tissue origins and occur in its coding poly (A)₁₀ microsatellite. SMAP1 mutations are especially frequent in MSI colorectal cancers (CRC), in which its mutation frequency decreases with tumour progression suggesting that tumours devoid of SMAP1 mutations are more invasive. Functional experiments allowed us to show that SMAP1 mutations lead to a defect in the fast recycling step of the transferrin receptor trafficking, to increased cell proliferation and to decreased invasiveness by maintaining the adherent junctions.

Interestingly, our observations showing that SMAP1 mutations increase the proliferative potential but reduce the invasiveness of MSI CRC-derived cell lines may explain some of the MSI CRCs clinical features, notably their large size and low metastatic potential.

Key words: colorectal cancer, mismatch repair, microsatellite instability, target gene, SMAP1, Arf6, clathrin-mediated endocytosis, cell proliferation, invasion/migration.

Adresse du laboratoire :

Centre de Recherche Saint-Antoine

Hôpital Saint-Antoine

Bâtiment Kourilsky – 4^{ème} étage

184 rue du faubourg Saint-Antoine

75571 PARIS Cedex 12

TABLE DES MATIERES

ABREVIATIONS.....	9
AVANT-PROPOS.....	12
PARTIE I. INTRODUCTION	15
I. ONCOGENESE MSI (MICROSATELLITE INSTABILITY).....	17
1. DECOUVERTE DE L'ONCOGENESE MSI	17
2. LE SYSTEME DE REPARATION DES MESAPPARIEMENTS DE BASES (MMR)	17
3. LE SYNDROME HNPCC (HEREDITARY NON POLYPOSIS COLORECTAL CANCER) ET L'INSTABILITE DES MICROSATELLITES	21
4. LES CANCERS SPORADIQUES	24
5. CARACTERISTIQUES CLINIQUES DES TUMEURS MSI	25
6. LE TRAITEMENT ADJUVANT DES TUMEURS STADE III PAR CHIMIOTHERAPIE	25
7. LA NOTION DE GENE CIBLE	27
II. LE GENE SMAP1 (SMALL ARFGAP1) : UN NOUVEAU GENE CIBLE DE L'ONCOGENESE MSI ...	32
1. LE GENE SMAP1	32
2. LA PROTEINE SMAP1	34
3. LES FONCTIONS CELLULAIRES DE SMAP1	36
III. LA PROTEINE ARF6 : ROLES PHYSIOPATHOLOGIQUES	38
1. LES PROTEINES ARF	38
2. LA PROTEINE ARF6.....	44
2.1. Arf6 et endocytose clathrine-dépendante	44
2.1.1. Endocytose clathrine-dépendante.....	45
2.1.2. Le recyclage des récepteurs membranaires	53
2.2. Rôle d'Arf6 dans la migration et l'invasion cellulaires.....	62
2.2.1. L'endocytose de la E-cadhérine.....	63
2.2.2. L'endocytose de l'intégrine $\beta 1$	68
IV. OBJECTIFS DE L'ETUDE	71
PARTIE II. RESULTATS	74
I. ARTICLE	74
II. RESULTATS COMPLEMENTAIRES.....	120
1. MATERIELS ET METHODES.....	120
1.1. Immunofluorescence et cytométrie en flux	120

1.2. Migration en deux dimensions	121
1.3. Invasion à travers une matrice de Matrigel	121
2. RESULTATS.....	122
2.1. Les mutations de SMAP1 stabilisent les jonctions adhérentes	122
2.2. Les mutations de SMAP1 augmentent la migration cellulaire	124
2.3. Les mutations de SMAP1 diminuent le pouvoir invasif des cellules	125
PARTIE III. DISCUSSION & PERSPECTIVES	129
I. DISCUSSION	129
1. LA CARACTERISATION DES MUTATIONS DE SMAP1 DANS LES TUMEURS	130
2. LE ROLE DE SMAP1 DANS L'ENDOCYTOSE CLATHRINE-DEPENDANTE DU RECEPTEUR DE LA TRANSFERRINE	132
3. LE ROLE DE SMAP1 DANS LA CYTOCINESE	135
4. LE ROLE DE SMAP1 DANS LA MIGRATION ET L'INVASION CELLULAIRES	138
II. PERSPECTIVES	142
PARTIE IV. CONCLUSION GENERALE	147
ANNEXES.....	150
BIBLIOGRAPHIE.....	184

TABLE DES ILLUSTRATIONS

FIGURES

FIGURE 1 : DISTRIBUTION DES DIFFERENTS TYPES DE CANCERS COLORECTAUX.....	15
FIGURE 2 : LE SYSTEME DE REPARATION DES MESAPPARIEMENTS DE BASES (MMR) CHEZ LA BACTERIE (A) ET CHEZ L'HOMME (B)	19
FIGURE 3 : REPARTITIONS DES MUTATIONS GERMINALES DES GENES DU SYSTEME MMR CHEZ LES PATIENTS ATTEINTS DU SYNDROME HNPCC.....	24
FIGURE 4: L'ONCOGENESE MSI.....	28
FIGURE 5 : FREQUENCES DE MUTATIONS DES GENES CIBLES DE L'ONCOGENESE COLORECTALE MSI.....	29
FIGURE 6 : LES PROTEINES ARFGAP.....	33
FIGURE 7: CYCLE GDP/GTP D'ARF6.....	34
FIGURE 8 : ALIGNEMENT DES PROTEINES SMAP1 ET SMAP2.....	35
FIGURE 9: LES CINQ FAMILLES DES PROTEINES ARFGEF.....	41
FIGURE 10: LE CYCLE GDP/GTP DES PROTEINES ARF.....	42
FIGURE 11: L'ENDOCYTOSE CLATHRINE-DEPENDANTE.....	46
FIGURE 12: REPRESENTATION SCHEMATIQUE DE LA CLATHRINE.....	50
FIGURE 13: STRUCTURE TRIDIMENSIONNELLE D'UN PANIER DE CLATHRINE.....	51
FIGURE 14: LE RECYCLAGE DES PROTEINES MEMBRANAIRES INTERNALISEES PAR LA VOIE CLATHRINE-DEPENDANTE.....	54
FIGURE 15: TRAFIC VESICULAIRE AU PONT INTERCELLULAIRE LORS DE LA CYTOKINESE.....	61
FIGURE 16: LES JONCTIONS ADHERENTES.....	64
FIGURE 17: ENDOCYTOSE CLATHRINE-DEPENDANTE DE LA E-CADHERINE LORS DE LA TRANSITION EPHITELIO-MESENCHYMATEUSE.....	65
FIGURE 18: EFFET DE NM23-H1 DANS L'INVASION CELLULAIRE.....	68
FIGURE 19: ACTIVATION DES INTEGRINES.....	69
FIGURE 20: ENDOCYTOSE DE L'INTEGRINE B1 AU COURS DE MIGRATION CELLULAIRE.....	70

FIGURE 21: LOCALISATION SUBCELLULAIRE DE LA E-CADHERINE.....	123
FIGURE 22: LOCALISATION SUBCELLULAIRE DE LA B-CATENINE.....	124
FIGURE 23: MIGRATION DEUX DIMENSIONS.....	125
FIGURE 24: INVASION A TRAVERS UN GEL DE MATRIGEL.....	127
FIGURE 25 : MODELE DU ROLE DE SMAP1 DANS LE RECYCLAGE RAPIDE DU TFR.....	134
FIGURE 26: MODELE DE L'ETAPE DE CYTOKINESE APRES REEXPRESSION DE SMAP1.....	136
FIGURE 27: MODELE DE LA CYTOKINESE DANS LES CELLULES HCT116.....	137
FIGURE 28: MODELE DU ROLE DE SMAP1 DANS LA STABILISATION DES JONCTIONS.....	141

TABLEAUX

TABLEAU 1: CLASSIFICATION TNM DES TUMEURS.....	27
TABLEAU 2: LES CLASSES DE PROTEINES ARF.....	38
TABLEAU 3: LES PRINCIPALES PROTEINES ADAPTATRICES QUI ONT UN ROLE DANS LA SELECTION DES CARGOS.....	49

ABRÉVIATIONS

5-FU : 5-fluoro-uracile

ACVRII : activin receptor type II

AP : adaptator protein

Arf : ADP ribosylation factor

ArfGAP : Arf GTPase activating protein

BFA : brefeldin A

CALM : Clathrin Assembly Lymphoid Myeloid Leukemia

CB : clathrin box

CCP : clathrin coated pit

CCR : cancer colorectal

CCV : clathrin coated vesicule

CHC : clathrin heavy chain

CIN : chromosomal instability

CLC : clathrin light chain

CME : clathrin-mediated endocytosis

COP : coat protein complex

DG : diacylglycérol

EGFR : epidermal growth factor receptor

ERC : endosomal recycling compartment

FIP : family interacting protein

FOLFOX : 5-FU, leucovorine, oxaliplatine

GDP : guanine diphosphate

GEF : guanine exchange factor

GTP : guanine triphosphate

HNPCC : human non polyposis colorectal cancer

IDL : insertion-deletion loop

InSiGHT : international society for gastrointestinal hereditary tumours

LOH : loss of heterozygosity

MDCK : Mardin-Darby Canine Kidney

MEC : matrice extracellulaire

MLH : MutL homolog

MLL : Myeloid/Lymphoid Leukemia

MMP : métalloprotéases

MMR : mismatch repair

MP : membrane plasmique

MSH : MutS homolog

MSI : microsatellite instability

MSS : microsatellite stable

MT : microtubule

MVB : multivesicular body

NMD : non-sense mediated mRNA decay

PA : phosphatidic acid

PAF : polypose adénomateuse familiale

PCNA : proliferating cell nuclear antigen

PCR : polymerase chain reaction

PI(3)P : phosphoinositide-3-phosphate

PIP2 : phosphoinositide-4,5-biphosphate

PIP3 : phosphoinositide-3,4,5-phosphate

PIP4 : phosphoinositide-4-phosphate

PIP5K : PIP5 kinase

PLD : phospholipase D

poly(A)₁₀ : 10 répétitions adénines

RE : reticulum endoplasmique

SILAC : stable isotope labeling by amino acids in cell culture

SMAP1 : small ArfGAP1

TEM : transition épithélio-mésenchymateuse

Tf : transferrine

TfR : transferrin receptor

TGFβRII : transforming growth factor β receptor type II

TNM : tumor node metastasis

Ub : ubiquitine

AVANT PROPOS

Mon projet de Thèse porte sur l'**Étude fonctionnelle de SMAP1 : un nouveau gène à la croisée du trafic vésiculaire et de l'oncogenèse MSI**. Ce projet a vu le jour à la suite d'une étude globale menée au laboratoire consistant à mettre en évidence de nouveaux gènes cibles de la cancérogenèse MSI(Microsatellite Instability). Le gène SMAP1 (Small ArfGAP 1) a été identifié comme étant très fréquemment muté dans les CCR (cancers colorectaux) MSI. Ainsi, l'équipe a décidé de mener un projet afin de caractériser les conséquences fonctionnelles de ces mutations, chose inédite au laboratoire dédié à l'analyse moléculaire des tumeurs MSI. À mon arrivée au laboratoire, des collaborations avaient été établies avec Christophe Lamaze à l'Institut Curie pour l'étude de l'endocytose et avec Marie-Lise Lacombe et Mathieu Boissan, du CdR Saint-Antoine pour l'étude de la migration/invasion.

La caractérisation moléculaire des mutations de SMAP1 a été menée lorsque j'étais au sein de l'équipe d'Alex Duval alors que la partie fonctionnelle a été réalisée lorsque j'ai rejoint l'équipe d'Yves Le Bouc au cours de ma seconde année de Thèse. Les expériences d'endocytose clathrine-dépendante du récepteur de la transferrine n'étant pas l'expertise du laboratoire, j'ai eu l'opportunité de l'importer au laboratoire après l'avoir prise en main au sein de l'équipe de Philippe Chavrier à l'Institut Curie.

Dans ce manuscrit, je commencerai par une revue bibliographique des différents aspects de mon travail, partie dans laquelle j'ai essayé de vous présenter les données de la littérature nécessaires et suffisantes à la compréhension de mon travail de Thèse. Dans un second temps, je détaillerai les résultats obtenus lors de ma Thèse sous forme d'article et de résultats complémentaires. Pour finir, je

discuterai mes résultats et les mettrai en relation avec les données de la littérature afin de définir des modèles moléculaires du ou des rôle(s) cellulaire(s) de SMAP1 et décrirai les perspectives possibles de cette étude. Des annexes sont présentes à la fin de mon manuscrit correspondant à une revue générale sur les CCR MSI à laquelle j'ai participé en cours de ma Thèse et un article portant sur l'étude de la division cellulaire des cellules souches intestinales que j'ai initié lors de mon stage de Master 2 réalisé à l'Institut de Génomique Fonctionnelle à Montpellier au sein de l'équipe de Philippe Jay.

Je vous en souhaite bonne lecture.

INTRODUCTION

PARTIE I. INTRODUCTION

Le cancer colorectal (CCR) représente un enjeu majeur de santé publique dans les pays industrialisés et notamment en France. En effet, il est la troisième cause de décès par cancer en France et représente plus de 36 000 nouveaux cas par an (Institut National du Cancer, Novembre 2010). Les CCR peuvent être dissociés en deux groupes selon le mécanisme moléculaire à l'origine de leur survenue (Figure 1).

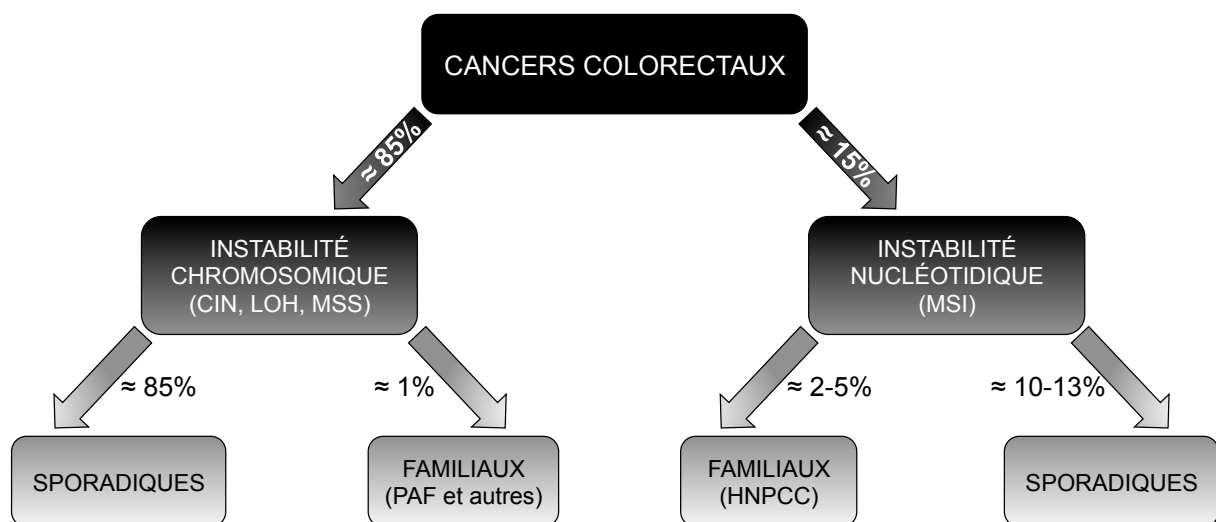


Figure 1 : Distribution des différents types de cancers colorectaux. CIN : Chromosomal INstability ; LOH : Loss Of Heterozygosity ; MSS : MicroSatellite Stable ; MSI : MicroSatellite Instable ; PAF : Polypose Adénomateuse Familiale ; HNPCC : Human Non Polyposis Colorectal Cancer

La majorité des tumeurs (≈ 85%) présente une instabilité chromosomique, elles sont appelées CIN (*Chromosomal INstability*), LOH (*Loss Of Heterozygosity*) ou encore MSS (*MicroSatellite Stable*). Elles ont un contenu en ADN anormal et présentent des pertes et des gains de fragments chromosomiques. Elles comportent le plus souvent des mutations dans les gènes *APC*, *TP53* et *KRAS*. Dans 12-15% des cas, les CCR sont dits MSI (*MicroSatellite Instability*). La tumeur présente une instabilité génétique

au niveau de séquences nucléotidiques répétées, les microsatellites. Ces séquences sont distribuées dans tout le génome et représentent 3% des séquences génomiques. Les microsatellites peuvent être intergéniques ou intragéniques, ces derniers pouvant être codants (exons) ou non codants (introns, promoteurs, UTR). De par leur nature, ces séquences sont des points chauds de mutation lors de la réplication. En effet, l'ADN polymérase « dérape » sur ces séquences ce qui aboutit à des mutations par insertion/délétion des motifs répétés. Les mutations survenant au cours de la réplication sont normalement reconnues et réparées par le système de réparation des mésappariements de base (MMR pour Methyl-directed Mismatch Repair). Dans une cellule MMR-déficiente le taux de mutations augmente d'un facteur 100 à 1000 comparé à une cellule dont le système MMR est intact. De nombreux gènes contenant des microsatellites codants sont mutés dans les tumeurs MSI mais un faible nombre a fait l'objet d'études fonctionnelles permettant de les impliquer directement à la tumorigenèse. En effet, seules les mutations des gènes du *TGFβRII* (*Transforming Growth Factor β Receptor type II*) [18] et de l'*ACVRII* (*Activin Receptor type II*) [19, 20] ont été clairement impliquées dans la cancérogenèse colorectale MSI.

Au laboratoire, a été identifié un nouveau gène cible muté dans 73% des CCR MSI [21] : Small ArfGAP 1 (*SMAP1*). Durant ma thèse, j'ai étudié les conséquences fonctionnelles des mutations de *SMAP1* dans l'oncogenèse MSI.

I. Oncogénèse MSI (Microsatellite Instability)

1. Découverte de l'oncogénèse MSI

La notion d'instabilité nucléotidique fut abordée pour la première fois en 1992 par Manuel Perucho lors d'un congrès international sur les oncogènes (« *Human oncogenes : the second decade* ») ; leurs travaux n'ont été publiés qu'en juin 1993, les éditeurs étant réticents au caractère inédit de la découverte [22]. Comme de nombreuses découvertes scientifiques, celle de « l'oncogénèse MSI » a été fortuite. En effet, c'est en recherchant une perte chromosomique afin de définir le gène responsable d'une forme familiale de CCR que deux autres équipes, celle de Bert Vogelstein et Kenneth Kinzler à Baltimore et celle de Richard Kolodner à Boston, ont publié en mai 1993, un mois avant la publication des travaux de Perucho, la démonstration d'un nouveau mécanisme moléculaire de transformation tumorale [23] [24] [25].

Les tumeurs de phénotype MSI présentent une instabilité nucléotidique au niveau de séquences particulières du génome, les microsatellites. Ces séquences répétées de motifs nucléotidiques de 1 à 6 paires de bases sont des points chauds de mutations lors de la réplication. Les mutations par insertion/délétion sont normalement réparées par le système MMR. Dans un contexte MSI, ce système est défaillant et les mutations ne sont pas réparées créant de nouveaux allèles.

2. Le système de réparation des mésappariements de bases (MMR)

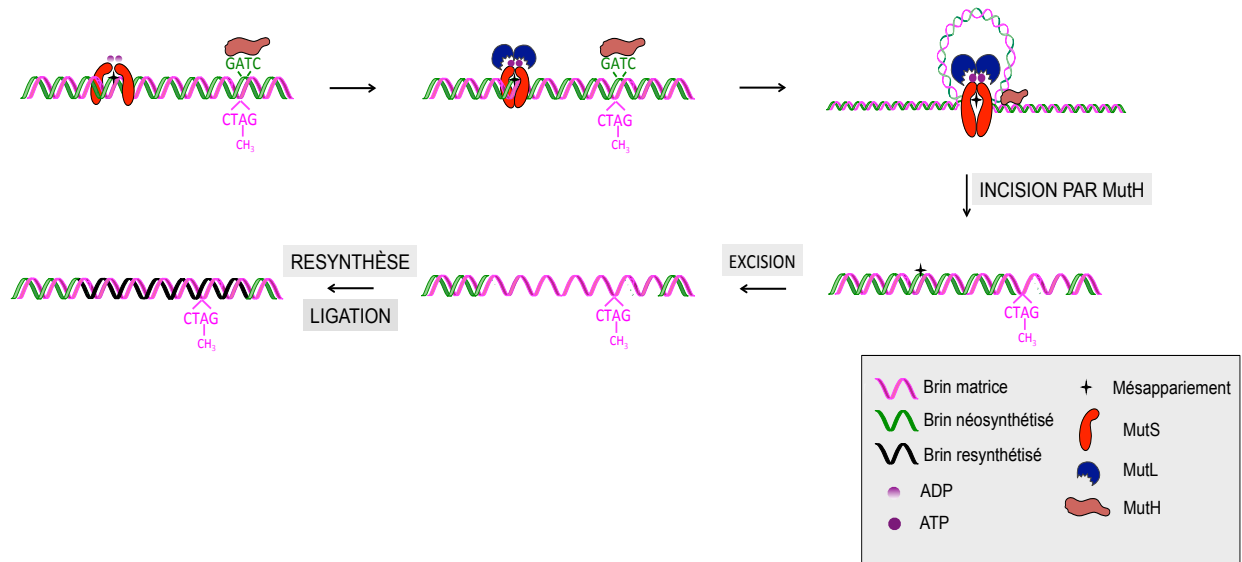
La fidélité de la réplication de l'ADN est en grande partie assurée par l'ADN polymérase elle-même. En effet, elle ne commet en moyenne qu'une erreur sur 10^5 bases appariées. De plus, l'ADN polymérase a une capacité de relecture et de

correction grâce à son activité 3'-5' exonucléase ; le taux d'erreur est alors de $1/10^7$. Les erreurs qui persistent à cette autocorrection sont réparées par le système MMR ; le taux d'erreur est alors de $1/10^{9-10}$.

Le système MMR est impliqué dans la réparation des simples mésappariements de bases et des boucles formées à la suite d'insertions/délétions (IDL, *Insertion/Deletion Loop*). Ce système a d'abord été décrit chez la bactérie *Escherichia coli* [26]. Il est composé de trois protéines principales MutS, MutL et MutH, dont il tire le nom, système MTHLS. La protéine MutS est une ATPase impliquée dans la reconnaissance des erreurs de réplication de type mésappariements de bases ou insertions/délétions. MutS agit sous forme d'homo-dimère pour pouvoir agir. Une fois liée à l'ADN, le chargement d'une molécule d'ATP aboutit à un changement de conformation permettant la liaison de la protéine MutL et la formation d'une boucle d'ADN autour du mésappariement [27]. La protéine MutL est une ATPase qui permet de coupler la reconnaissance de l'erreur de réplication et l'étape d'excision [28]. La discrimination entre les brins matrice et néosynthétisé est rendue possible par la présence de séquences GATC méthylées sur l'adénine, cette modification n'intervenant qu'après la réplication, elle identifie le brin matrice. Ainsi, l'endonucléase MutH se lie aux séquences GATC non-méthylées sur l'adénine du brin néosynthétisé [29] et digère l'ADN au niveau de ces séquences. L'activité endonucléase de MutH est stimulée par MutL quand celle-ci est liée à MutS et favorise également l'hydrolyse de l'ATP de MutS nécessaire à la translocation de l'homo-dimère MutS le long de la molécule d'ADN jusqu'au complexe ADN/MutH le plus proche [30]. L'excision du brin d'ADN nécessite la présence de l'hélicase MutU (ou UvrD ou ADN hélicase II) qui est recrutée par MutL [31] [32]. MutH étant capable de digérer l'ADN en 3' ou en 5' du mésappariement, l'excision se fera par des

exonucléases 3'-5' ou 5'-3' [33]. Le brin digéré est resynthétisé par l'ADN polymérase III et reliqué [26] (Figure 2A).

A



B

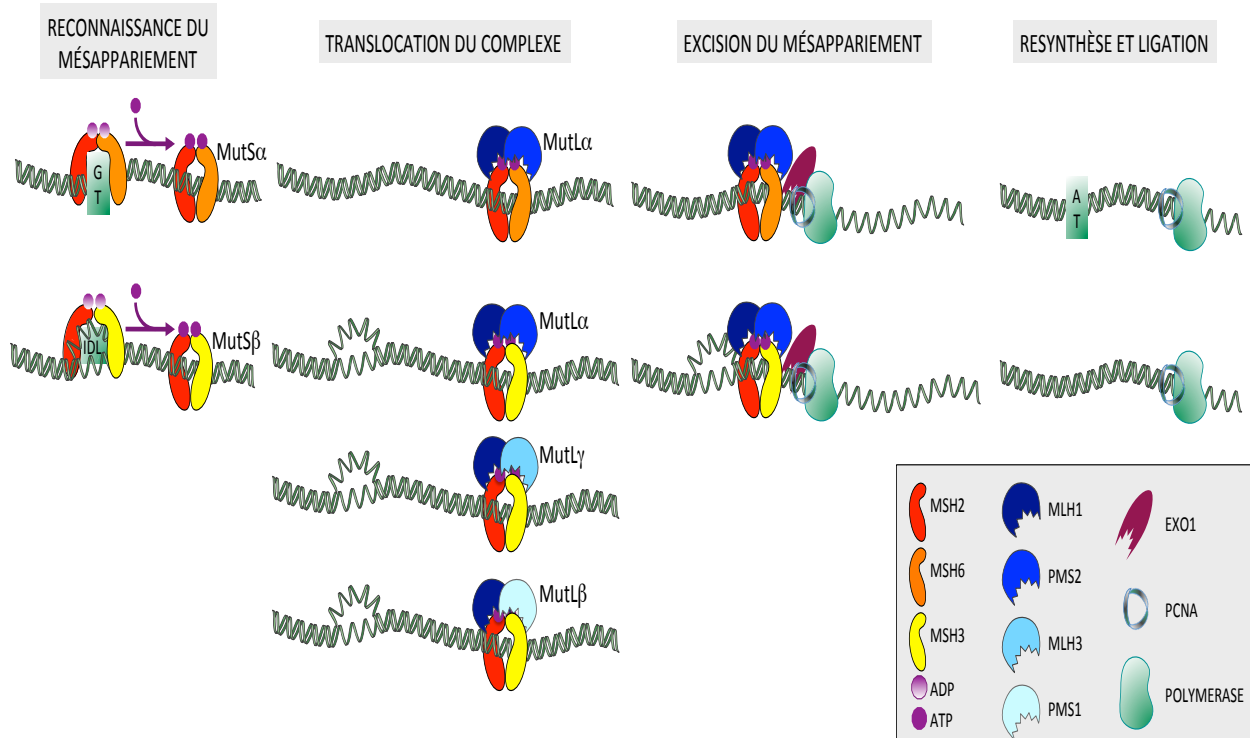


Figure 1 : Le système de réparation des mésappariements de bases (MMR) chez la bactérie (A) et chez l'homme (B). D'après Jacob et al., 2002, Biochimie [11] et Zaanen et al., 2010, Cell Oncol [17].

Ce système MMR a également été identifié chez les eucaryotes, notamment la levure *Saccharomyces cerevisiae* [34] [35] [36]. Dans cet organisme, six homologues de la protéine MutS bactérienne ont été identifiés, MSH1 (MutS Homolog) impliquée dans l'intégrité de l'ADN mitochondrial [35] [36], MSH2, MSH3 [37], MSH4 et MSH5 impliquées dans la méiose [38], et MSH6 [39]. La protéine MSH2 est capable de former des hétérodimères avec les protéines MSH6 pour former le complexe MutS α et MSH3 pour former le complexe MutS β [40] [41] [42]. Le complexe MutS α est capable de reconnaître les mésappariements de base et les insertions/délétions d'une à quelques bases alors que le complexe MutS β est lui capable de reconnaître les insertions/délétions de plusieurs bases [43]. Ces complexes existent aussi bien chez *S. cerevisiae* que chez l'Homme. La protéine MutL bactérienne a quatre homologues chez la levure, MLH1-3 (*MutL Homolog*) [44] et PMS1 (*PostMeiotic Segregation*) [45] et quatre chez l'Homme, MLH1, MLH3, PMS1 (=MLH2 chez la levure) et PMS2 (=PMS1 chez la levure). La protéine MLH1 est capable de former les complexes MutL α avec hPMS2 (PMS1 chez la levure), MutL β avec PMS1 et MutL γ avec MLH3. Le complexe MutL α se lie au complexe MutS α pour réparer les simples mésappariements de bases et les insertions/délétions d'une à deux bases et à MutS β pour la réparation des insertions/délétions $\geq$ à deux bases, les complexes MutL β et MutL γ se lient, plus rarement, à MutS β afin de réparer les insertions/délétions de plusieurs nucléotides (Figure 2B). Contrairement à *E. coli*, chez les eucaryotes, la discrimination entre les brins matrice et néo-synthétisé ne dépend pas de la méthylation de l'ADN ; le mécanisme permettant de différencier les brins matrice et néo-synthétisé n'est pas formellement identifié à ce jour bien qu'une protéine candidate ait été désignée : la protéine PCNA (*Proliferating Cell Nuclear Antigen*). Cette protéine est impliquée dans la réplication de l'ADN en augmentant la processivité de l'ADN polymérase c'est-à-dire sa capacité à synthétiser l'ADN sans

se détacher de la matrice. PCNA est capable de se lier à l'extrémité N-terminale des protéines MSH6 et MSH3 et permettrait ainsi de guider les complexes MutS α ou MutS β vers l'extrémité libre de la molécule d'ADN en cours de synthèse [46]. PCNA est également capable de lier l'exonucléase 5'-3' EXO1 qui elle-même interagit avec MSH2, MSH3 et MLH1 chez l'Homme [47]. Cette exonucléase ne serait pas la seule impliquée dans l'excision du brin portant le mésappariement. Le brin d'ADN est resynthétisé par l'ADN polymérase δ et peut être ϵ [48] et ligué probablement par l'ADN ligase I qui est souvent associée à l'ADN polymérase δ et à PCNA [49].

3. Le syndrome HNPCC (Hereditary Non Polyposis Colorectal Cancer) et l'instabilité des microsatellites

L'existence d'une prédisposition génétique au développement de tumeurs épithéliales dans certaines familles a été décrite pour la première fois en 1913 par le Dr Warthin. En 1966, le docteur Henry T. Lynch a étudié deux grandes familles atteintes de CCR, de cancers de l'endomètre et de l'estomac. Cette étude lui a permis d'établir le caractère autosomique dominant de la transmission de cette prédisposition [50] [51] qui est nommée syndrome de Lynch ou HNPCC (Hereditary Non Polyposis Colorectal Cancer). Ce syndrome héréditaire est différent de la polypose adénomateuse familiale (PAF), autre prédisposition génétique, décrite chez des patients porteurs le plus souvent de mutations germinales dans le gène *APC* et présentant de nombreux polypes intestinaux bénins qui dégénèrent en tumeur maligne. Le syndrome HNPCC représente la prédisposition familiale la plus importante des CCR (2-5%, 1% pour la PAF, Figure 1). Ce n'est qu'au début des années 90, 25 ans après la description clinique du syndrome HNPCC, que les études moléculaires de cette prédisposition génétique vont être menées. Les

analyses de liaison réalisées sur des familles répondant aux critères d'Amsterdam de 1991 (au moins trois personnes atteintes dans deux générations successives, avec un diagnostic avant l'âge de 50 ans), ont permis de lier la maladie aux chromosomes 2p [23] [24] ou 3p [52] selon les familles étudiées. La recherche d'une perte d'hétérozygotie dans ces régions candidates dans les tumeurs des patients a permis de mettre en évidence une instabilité au niveau des marqueurs nucléotidiques utilisés pour l'analyse de liaison. En effet, en analysant les profils de migration de ces marqueurs sur gel après PCR (*Polymerase Chain Reaction*), une instabilité nucléotidique a été observée chez des patients atteints du syndrome HNPCC ainsi que dans des cas sporadiques [23] [22]. Cette instabilité ne concerne pas uniquement les dinucléotides (CA)_n mais également les (CT)_n [22], les trinucléotides (CAG)_n [23] et (GGC)_n [22]. L'instabilité des microsatellites observée par ces équipes consiste essentiellement en des délétions de motifs nucléotidiques. À ce stade, l'équipe de Manuel Perucho émet l'hypothèse selon laquelle ces délétions sont les conséquences d'un défaut dans une protéine essentielle à la fidélité de la réplication ou à la réparation des erreurs de réplication [22].

Cette hypothèse ouvre la porte à la recherche des gènes humains homologues des gènes bactériens du système MuthLS. C'est ainsi qu'à la fin de l'année 1993, l'équipe de Richard Kolodner publie un article sur la découverte d'un gène lié au syndrome HNPCC localisé en 2p21-22 : hMSH2 (*human MutS Homolog 2*) [53]. Ce gène est l'homologue humain du gène bactérien *MutS*. Afin d'impliquer de façon certaine le gène hMSH2 dans la prédisposition au syndrome HNPCC, l'équipe de Kolodner a étudié ce gène dans 26 tumeurs de patients HNPCC. Ils ont pu identifier une mutation qui engendre la traduction d'une protéine anormale à la suite d'un épissage modifié. Cette mutation est présente à l'état homozygote dans la tumeur et hétérozygote dans le sang des patients [53]. L'équipe de Vogelstein a analysé le

gène hMSH2 dans des familles plus grandes que celles étudiées par Kolodner ce qui a permis de confirmer la présence de mutations qui aboutissent à la synthèse d'une protéine tronquée ou à une protéine non fonctionnelle [54]. Les mutations constitutionnelles de MSH2 sont responsables d'un tiers des syndromes HNPCC (Figure 3 ; *International Society for Gastrointestinal Hereditary Tumours* {InSiGHT}, Novembre 2011). Ces mutations sont distribuées sur la totalité du gène et la majorité consiste en des substitutions de base (InSiGHT, <http://www.insight-group.org/mutations/>).

Le second gène codant une protéine du système MMR qui a été identifié comme lié au locus HNPCC est le gène MLH1, homologue du gène MutL bactérien. Ce gène est localisé sur le chromosome 3p21.3, responsable du syndrome HNPCC dans certaines familles [55] [56]. La protéine MLH1 est la protéine nécessaire à la formation des complexes MutL chez les eucaryotes. L'absence de protéine MLH1 dans la cellule aboutit à un phénotype d'instabilité sévère. La plupart des mutations germinales décrites à ce jour consistent en des mutations tronquantes, dues à des décalages du cadre de lecture ou à des mutations dans les sites d'épissage [55] [11]. Les mutations faux-sens décrites touchent en majorité le site ATPase ou le site de liaison aux autres protéines MLH (MLH3, PMS1 et PMS2) [11], régions très conservées entre les espèces. Les mutations germinales du gène MLH1 sont responsables de 40% des syndromes HNPCC (Figure 3 ; InSiGHT, Novembre 2011). Ainsi, les mutations des gènes MSH2 et MLH1 sont responsables à eux seuls de plus de 70% des syndromes HNPCC.

Les autres gènes homologues des gènes bactériens identifiés comme étant impliqués dans le phénotype MSI, bien qu'ayant des effets moindres que MSH2 et MLH1, sont MSH6 [57] et hPMS2 [58]. Les mutations germinales du gène MSH6 sont responsables de près de 20% des syndromes HNPCC et celles du gène PMS2

représentent $\approx 8\%$ des syndromes HNPCC. Les gènes MLH3 et PMS1 ne sont mutés que dans un peu plus de 1% des cas (Figure 2, InSiGHT, Novembre 2011).

La transformation tumorale requiert la perte du second allèle par mutation somatique ou LOH ; ce schéma suit la théorie de Knudson selon laquelle un second événement, la perte de l'allèle normal, est nécessaire à la transformation tumorale.

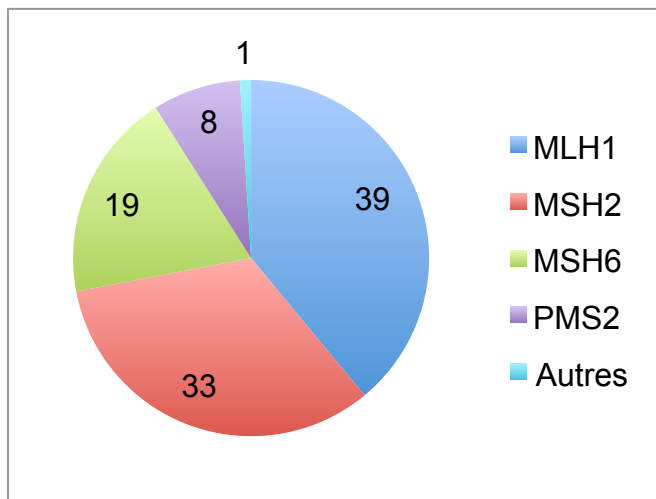


Figure 3 : Répartitions des mutations germinales des gènes du système MMR chez les patients atteints du syndrome HNPCC. (InSiGHT, Novembre 2001)

4. Les cancers sporadiques

Le phénotype MSI a également été décrit dans des cancers survenant chez des patients ne présentant pas de prédisposition génétique, ce sont les cas sporadiques [22] [25] qui représentent 10 à 13% de l'ensemble des CCR. La survenue des tumeurs est en moyenne plus tardive dans les cas sporadiques que chez les patients atteints du syndrome HNPCC (72 ans au lieu de 45 ans). Les cancers de phénotype MSI représentent environ 15% des CCR, des cancers de l'estomac et de l'endomètre.

De manière étonnante, aucune mutation des gènes du système MMR n'a été identifiée dans les tumeurs sporadiques MSI [59]. En effet, c'est une inactivation bi-allélique du gène MLH1 par méthylation des îlots CpG en 3' [60] qui est responsable

du développement de ces tumeurs [61] [62] [63]. Cette inhibition transcriptionnelle aboutit à l'absence de protéine MLH1 dans la cellule et donc à un système MMR déficient.

5. Caractéristiques cliniques des tumeurs MSI

La différence entre les CCR MSI, sporadiques ou familiaux, et les tumeurs MSS ne se limite pas uniquement aux caractéristiques moléculaires. En effet, d'un point de vue clinique, les CCR MSI sont des tumeurs volumineuses qui se développent au niveau du colon droit pour la majorité des cas sporadiques, sont peu différenciées, mucineuses, présentent une forte infiltration lymphocytaire et sont peu invasives à distance [22] [64] [17]. Les tumeurs MSI sont de meilleur pronostic que les tumeurs MSS [25] [65] [66] et répondent différemment au traitement par chimiothérapie [67] [68] [69].

6. Le traitement adjuvant des tumeurs stade III par chimiothérapie

Les traitements chimiothérapeutiques des patients atteints de tumeurs MSI ne diffèrent pas de ceux atteints de tumeurs MSS. Les années 2000 ont vu se développer plusieurs études cliniques portant sur différents traitements dont la chimiothérapie adjuvante qui consiste à traiter les patients après résection de la tumeur par chirurgie afin de diminuer les risques de rechute. Un certain nombre d'études ont été menées sur les chimiothérapies adjuvantes basées sur le 5-fluorouracile (5-FU), un analogue de pyrimidine qui bloque la réplication, afin de comparer les réponses aux traitements des patients atteints d'un CCR MSS ou MSI (sporadique ou familial) de stade III selon la classification TNM, c'est-à-dire avec envahissement ganglionnaire mais sans métastases à distance (Table 1). Plusieurs

études ont mis en évidence des différences de réponse à la chimiothérapie selon que les tumeurs soient de phénotype MSS ou MSI [67] [68] [69]. En effet, le traitement par 5-FU seul après chirurgie des patients atteints d'une tumeur MSI de stade III ne leur est pas bénéfique comparé aux patients atteints d'une tumeur MSS [70]. Depuis 2004, les patients opérés pour un CCR de stade III sont traités par le protocole FOLFOX qui consiste en l'administration de 5-FU en association avec l'oxaliplatine, qui potentialise l'effet du 5-FU. L'oxaliplatine est un agent alkylant qui fait majoritairement des pontages intrabrins ou interbrins entre deux guanines bloquant ainsi la réplication de l'ADN. Le bénéfice du traitement adjuvant par le protocole FOLFOX est beaucoup plus marqué pour les patients atteints de cancer du côlon MSI de stade III que pour ceux atteints d'un cancer du colon MSS pour lesquels l'allongement de la survie sans progression de la maladie est modeste [71] [69]. Néanmoins, cette étude ayant été réalisée sur un nombre relativement limité de patients (n=233) et de manière rétrospective, il est indispensable que ces observations soient confirmées dans une étude randomisée prospective pour pouvoir être intégrées dans l'arbre décisionnel de la prise en charge des patients atteints d'un cancer du côlon de stade III.

Groupe selon la classification TNM			
STADE	Stade T	Stade N	Stade M
I	T1 ou T2	N0	M0
IIa	T3	N0	M0
IIb	T4	N0	M0
IIIa	T1 ou T2	N1	M0
IIIb	T3 ou T4	N1	M0
IIIc	Tx	N2	M0
IV	Tx	Nx	M1

Tableau 1: Classification TNM des tumeurs. T : tumeur primaire ; N : ganglions lymphatiques régionaux ; M : métastases. T1 : tumeur envahissant la muqueuse ou la sous-muqueuse, T2 : tumeur envahissant la couche musculaire externe sans la dépasser, T3 : tumeur envahissant les tissus non pariétaux péri-coliques ou péri-anaux, T4 : tumeur envahissant les tissus voisins, Tx : quel que soit le stade T. N0 : aucun envahissement ganglionnaire, N1 : envahissement de 1 à 3 ganglions, N2 : envahissement de 4 ganglions et plus, Nx : quel que soit le stade N. M0 : aucune métastase à distance, M1 : présence de métastase(s) à distance.

7. La notion de gène cible

La déficience en système MMR n'est pas un phénomène transformant en soi. En effet, ce sont les mutations accumulées dans les gènes dits « cibles » qui ont un pouvoir néoplasique. Les gènes cibles de l'oncogénèse MSI contiennent un microsatellite codant ou non codant. Les microsatellites codants sont en général des mononucléotides alors que les microsatellites introniques peuvent être composés de motifs di-nucléotidiques. Plus le nombre de répétitions nucléotidiques est élevé plus la fréquence de mutations est élevée. Les mutations par insertion/délétion des microsatellites mono-nucléotidiques codant aboutissent à un décalage du cadre de lecture, et le plus souvent à l'apparition d'un codon STOP prématuré au niveau du transcrit qui peut être traduit en une protéine tronquée présentant un néo-peptide à l'extrémité COOH (Figure 4).

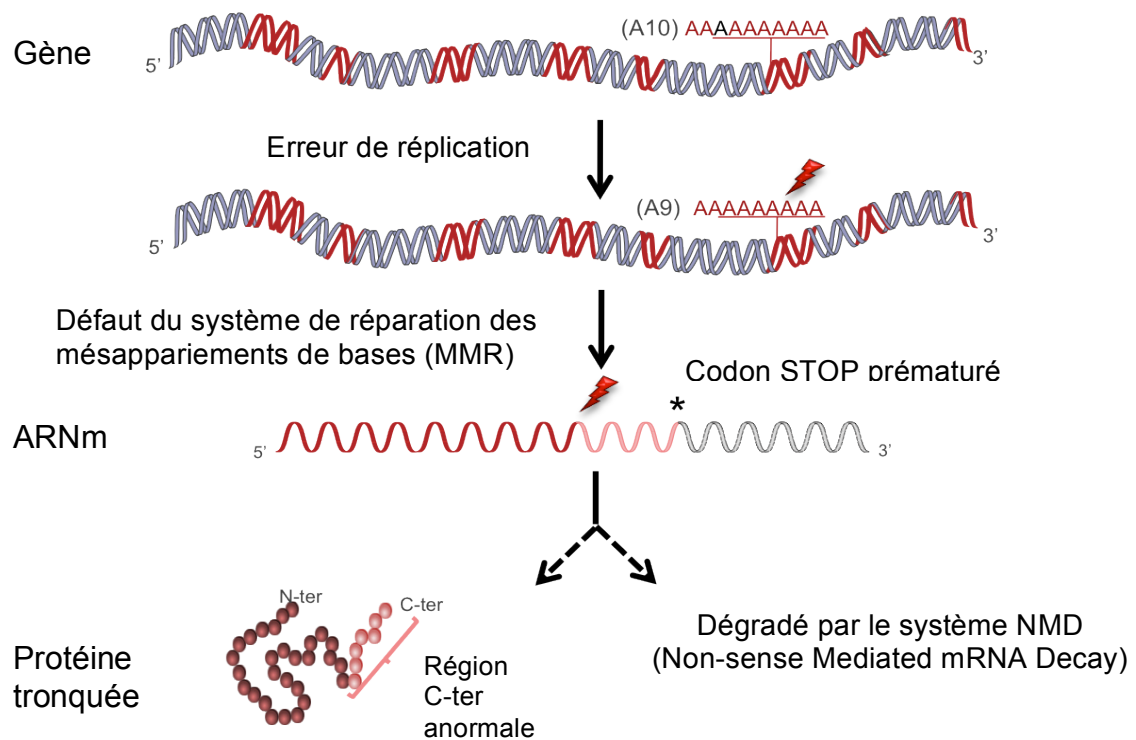


Figure 4: L'oncogenèse MSI.

Sur la base de la définition d'un gène cible, plus de 500 gènes sont potentiellement cibles de l'oncogenèse MSI (SelTarBase ; <http://www.seltarbase.org/?extension=latest&topic=main>). De ce fait, l'enjeu de l'étude des tumeurs MSI est de distinguer les mutations dites « bruit de fond » qui surviennent du fait de la déficience du système MMR des mutations effectivement sélectionnées pour la transformation tumorale. Très peu de gènes mutés dans les tumeurs MSI ont été directement impliqués dans l'oncogenèse. Effectivement, seuls les gènes des récepteurs de type 2 du TGF β (TGF β R2) [18] et de l'activine (ACVR2) [20] ont fait l'objet d'études fonctionnelles. Afin de contourner les difficultés inhérentes aux études fonctionnelles, deux modèles statistiques ont été établis : (i) l'un est basé sur les fréquences de mutations observées dans les tumeurs MSI pour

un gène donné selon la taille du microsatellite [2] (Figure 5A) et (ii) l'autre sur les fréquences de mutations observées dans les tumeurs MSI dont la taille de la répétition est $\geq$ à 7 nucléotides [4] (Figure 5B). Ainsi, d'après le modèle de Woerner, un gène ayant un microsatellite de 8 nucléotides est considéré comme cible si la fréquence de mutations est $\geq$ à 30%, alors que pour un gène ayant un microsatellite de 10 nucléotides cette fréquence doit être au moins de 56% (Figure 5A). Les gènes cibles sont différents d'un tissu à l'autre du fait de l'expression tissu-spécifique de certains gènes. Dans le cas des CCR MSI et selon le modèle de Duval, la fréquence de mutations des gènes cibles est $\geq$ à 30%. Deux gènes sont plus fréquemment mutés : les gènes *TGF β R2* et *ACVR2*, ayant une répétition de 10 adénines, dont les fréquences de mutation sont de 85% et 68% respectivement (Figure 5B).

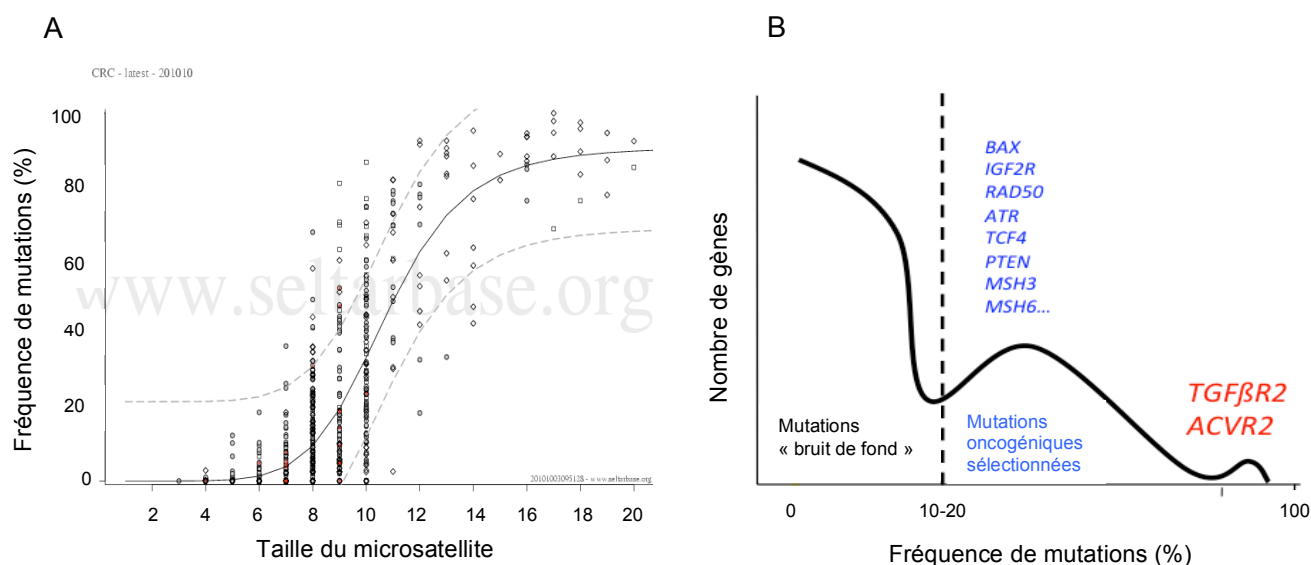


Figure 5 : Fréquences de mutations des gènes cibles de l'oncogénèse colorectale MSI.
 (A) Modèle statistique de Woerner [2]. (B) Modèle statistique de Duval [4].

Jusque récemment, les gènes étudiés étaient ceux qui présentaient un microsatellite dans une région codante car il est difficile de prévoir les conséquences de mutations

dans les régions introniques. Cependant, plusieurs études ont mis en évidence les conséquences oncogéniques des mutations des microsatellites introniques, situés pour la plupart dans les sites accepteurs d'épissage, dans différents gènes [72] [73] [74]. En effet, les gènes ATM et MRE11 impliqués dans la réparation des cassures double brin de l'ADN, possèdent des microsatellites non codants de 15 et 11 répétitions de thymine respectivement [73] et le gène HSP110, codant une protéine chaperonne, a un microsatellite de 17 répétitions de thymine [74]. Les mutations de ces microsatellites entraînent un épissage aberrant des ARNm aboutissant à la synthèse de protéines tronquées qui, dans le cas de HSP110, joue un rôle dominant négatif [74].

Dans la cellule, il existe un système de surveillance des transcrits, le système NMD (*Non-sense Mediated mRNA Decay*). Ce système consiste en un complexe protéique capable de détecter les codons STOP prématurés et de dégrader les ARN anormaux [75] (Figure 4). Afin d'identifier de nouveaux gènes cibles de l'oncogenèse colorectale MSI, la stratégie du laboratoire fut d'inhiber le système de surveillance des ARNm anormaux, le système NMD, par ARN interférence spécifique du gène Upf1, facteur essentiel du NMD, dans une lignée cellulaire de CCR MSI, la lignée HCT116 [21]. En comparant les niveaux d'expression des gènes dans la lignée parentale et la lignée transfectée avec le siARN Upf1, notre laboratoire a pu identifier 1251 gènes surexprimés après inhibition de Upf1 (rapport $\geq 1,5$). Parmi ces gènes, 472 présentent un microsatellite codant d'au moins 7 répétitions de nucléotides et 22 d'entre eux ont un rôle potentiel dans l'oncogenèse colorectale. La présence de mutation dans le microsatellite codant a été recherchée dans ces 22 gènes ; 9 gènes présentent une mutation à l'état homo- ou hétérozygote dans la lignée cellulaire HCT116, dont le gène du TGF β R2 déjà défini comme gène cible, permettant de valider l'approche du laboratoire. La recherche de mutations de ces 9 gènes dans 44

tumeurs colorectales primaires MSI a permis de définir les fréquences de mutation de ces gènes dans le but de les considérer ou non comme cibles de l'oncogenèse MSI. Comme décrit dans la littérature, des mutations du gène TGF β R2 ont été identifiées dans plus de 90% des 44 tumeurs. Quatre gènes jamais décrits comme cibles de l'instabilité MSI ont été identifiés dont le gène SMAP1 (Small ArfGAP1) muté dans 73% des tumeurs analysées. Cette fréquence de mutations place SMAP1 dans le groupe de gènes très fréquemment mutés dans les CCR MSI (Figure 5B). Ces résultats nous ont motivé à développer le projet d'étude fonctionnelle des mutations de SMAP1 dans l'oncogenèse colorectale MSI.

En résumé, l'étude de cas familiaux de cancers a permis la découverte d'un nouveau mécanisme moléculaire initiateur de tumeurs, l'instabilité des microsatellites. Ce phénotype de tumeurs a également été découvert dans des cas sporadiques. Il est dû à l'absence d'un système MMR efficace et donc à l'accumulation de mutations essentiellement par insertion/délétion dans les séquences microsatellites. Le phénotype MSI a été décrit dans les tumeurs de l'endomètre, de l'estomac et dans les CCR. La majorité des CCR MSI (72%) n'expriment plus les gènes MLH1 ou MSH2, essentiels au système MMR. L'étude des CCR MSI a permis de mettre en évidence le caractère particulier de ces tumeurs. En effet, elles représentent un groupe à part entière, aussi bien d'un point de vue moléculaire que clinique. Elles sont de meilleur pronostic que les tumeurs MSS et répondent différemment aux traitements par chimiothérapie. L'étude moléculaire de ces tumeurs a permis de mettre en évidence des gènes cibles de la cancérogenèse colorectale MSI mais trop peu d'études fonctionnelles sont réalisées afin de comprendre leur implication dans la tumorigenèse. Ainsi, la définition des gènes cibles n'est réalisée, pour la plupart des gènes, que sur des données de fréquence de mutations selon les deux modèles

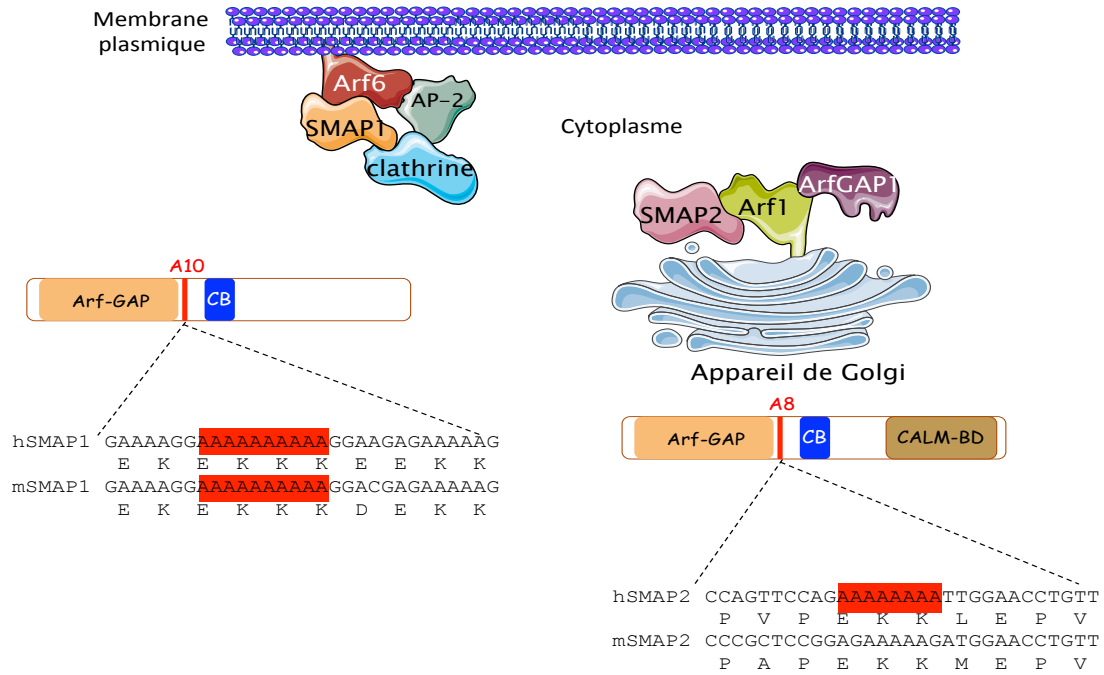
statistiques de Woerner et Duval. Ces gènes cibles comportent des séquences microsatellites mutées par insertion/délétion dans les cellules tumorales dû à l'absence d'un système MMR efficient aboutissant à un décalage du cadre de lecture. La protéine traduite est tronquée ou absente suite à la dégradation de l'ARNm anormal par le système NMD. C'est l'absence de protéine ou l'expression d'une protéine tronquée qui est à l'origine du phénomène néoplasique.

II. Le gène SMAP1 (Small ArfGAP1) : un nouveau gène cible de l'oncogenèse MSI

1. Le gène SMAP1

Le gène SMAP1 (Small ArfGAP1) est situé sur le chromosome 6q13 et est constitué de 11 exons dont un alternatif, l'exon 5 [76]. Il possède un microsatellite de 10 répétitions adénines (A10) dans l'exon 6 (Figure 6A). Il appartient à la famille des ArfGAP (ADP ribosylation factor GTPase Activating Protein) qui comptent à ce jour 31 gènes regroupés en 10 sous-familles selon une nomenclature consensus définie lors d'une conférence internationale regroupant les acteurs principaux de la recherche sur ces protéines [6] (Figure 6B).

A



B

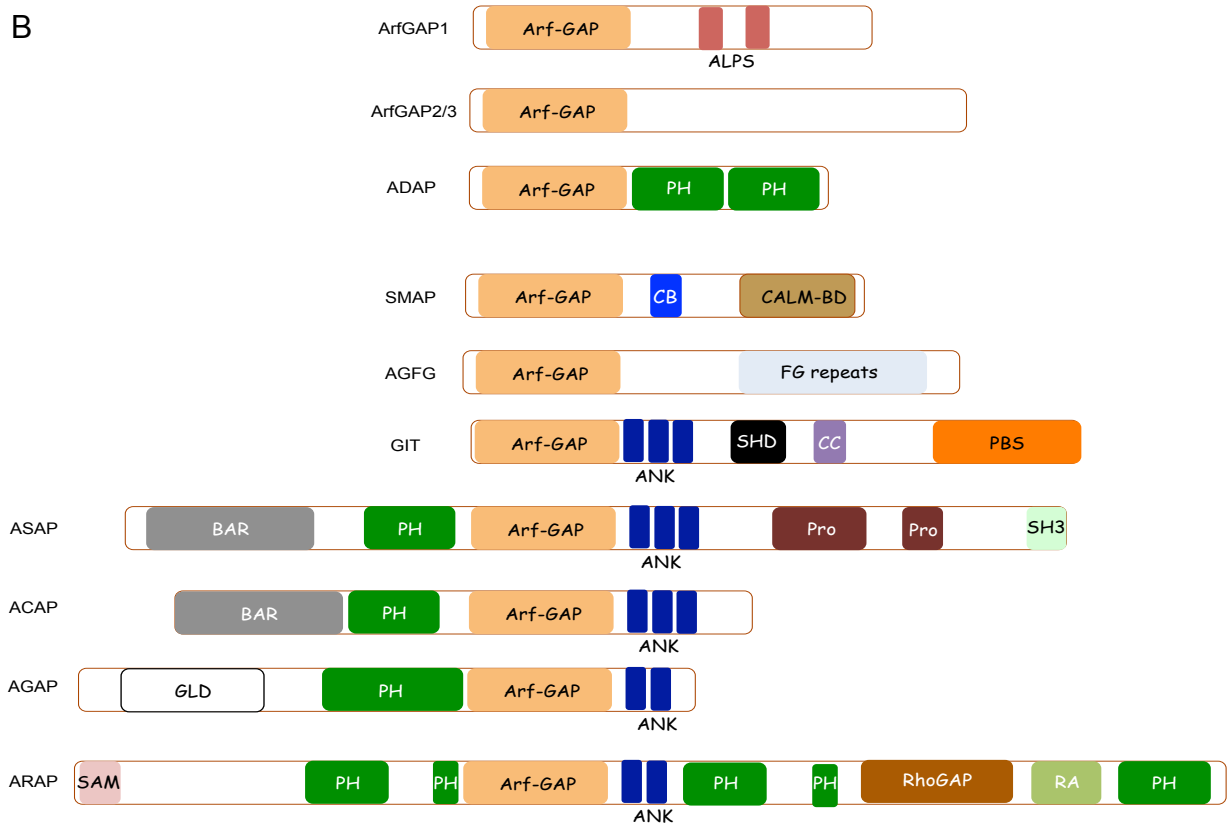


Figure 5 : Les protéines ArfGAP. (A) Structure des domaines fonctionnels et localisation sub-cellulaires des protéines SMAP1 et SMAP2. Les alignements de séquences des protéines humaines et murines sont également représentés. (B) Structure des domaines fonctionnels de chaque sous-famille de ArfGAP. ALPS : ArfGAP1 lipid-packing sensor ; ArfGAP : ADP ribosylation factor GTPase Activating Protein ; ANK : Ankirin repeat ; BAR : Bin/Amphiphysin/Rvs ; CALM-BD: CALM binding domain ; CB : Clathrin Box ; CC : Coiled-Coil ; FG repeats : plusieurs copies du motif XXFG ; GLD : GTP-binding protein-like domain ; PBS : Paxillin Binding Site ; PH :Pleckstrin homology domain ; RA : Ras association motif ; RhoGAP : RhoGAP domain ; SAM : Sterile α -motif ; SH3 :Src homology 3 domain ; SHD : Spa-homology domain. D'après Kahn et al., 2008, *J Cell Biol* [6].

Le gène SMAP1 a été le premier gène identifié de la sous-famille SMAP [77] qui ne comprend que deux membres : SMAP1 et SMAP2 [78]. Il code la protéine SMAP1, une ArfGAP spécifique de la protéine Arf6 impliquée dans l'hydrolyse du GTP d'Arf6 et donc dans la régulation de son cycle GDP/GTP (Figure 7), qui sera détaillé plus loin dans mon manuscrit.

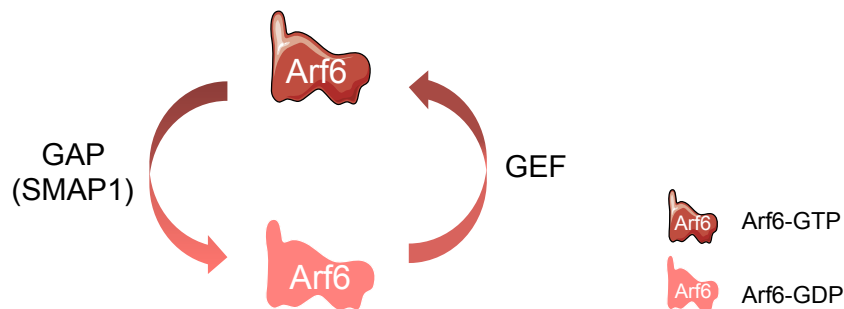


Figure 6: Cycle GDP/GTP d'Arf6. GAP : GTPase Activating Protein ; GEF : Guanine Exchange Factor.

2. La protéine SMAP1

La protéine SMAP1 est constituée d'un domaine ArfGAP permettant l'hydrolyse du GTP d'Arf6 en GDP et d'un domaine de liaison à la clathrine, protéine impliquée dans l'endocytose. La protéine SMAP1 serait impliquée dans l'endocytose clathrine-dépendante de protéines membranaires tels que le récepteur à la transferrine (TfR) [77] ou la E-cadhérine [79]. Les protéines SMAP sont les seules de la famille ArfGAP à pouvoir lier la chaîne lourde de la clathrine *via* leur domaine LLGLD (*clathrin box*, CB). Les protéines SMAP1 et SMAP2 partagent 50% d'identité sur toute leur longueur et 85% dans leur domaine ArfGAP [78] (Figure 8). À la différence de SMAP1, SMAP2 possède en plus des domaines ArfGAP et CB un domaine de liaison à la protéine CALM (*Clathrin Assembly Lymphoid Myeloid Leukemia*) qui

aurait un rôle dans la régulation de la taille des vésicules recouvertes de clathrine [1]. SMAP2 est une ArfGAP spécifique d'Arf1 localisée à la membrane des endosomes qui a un rôle dans la régulation du transport rétrograde trans-golgien [80]. *In vitro*, SMAP1 et SMAP2 sont toutes deux capables d'hydrolyser le GTP d'Arf1 et d'Arf6 [78]. Ainsi, la spécificité de chaque protéine décrite *in vivo* serait due à la seule localisation subcellulaire de SMAP1 et de SMAP2. En effet, il a récemment été décrit que la région C-terminale de SMAP2 est nécessaire et suffisante à sa localisation subcellulaire et à sa spécificité [81].

hSMAP1	MATRS	SCREKAQKLNEQH	LI	LSKLL	REEDNKY	CADCEAKGPRWASWNIGVFICIRCAGIH
hSMAP2	MTGKS	SVKD-----VDRY	QAVLANLL	LEEDNKF	CADCQSKGPRWASWNIGVFICIRCAGIH	
hSMAP1	RNLGVHISRVKSVNLDQWTAE	EQIQCMQDMGNT	KARLLYEANLPEN	FRRPQT	DQAVEFFIR	
hSMAP2	RNLGVHISRVKSVNLDQWTQ	EQIQCMQEMGNGKANRLYEAYLPET	FRRPQID	PAVEGFIR		
hSMAP1	DKYEKKKY	YDK----	NAIAITNKEKEKKKEEKKREKEPEKPAKPLTAEKLQKKDQOLEPK			
hSMAP2	DKYEKKKYMDRSLDINAFRKE	KDDKWKRGSEPVPEKKLE	EPVVF	EKVMPQKKEDPQLPRK		
hSMAP1	KSTSPKKAAEPTV	DLLGLDGP	PAVAPVTNGNT	TVPPLNDDL	LDIFGPMISNPLPATVMPPAQ	
hSMAP2	---SSPKSTAPVM	DLLGLDAP	AVACSIANSKTS	-----NTLEKDLDLLASVPSPSSSGSRK		
hSMAP1	GTPSAPAAATLSTV	TSGDL	DLFTEQTTKSEEVAKKQLSKDSILSLYGTGTIQ	QQSTPGVF		
hSMAP2	VVGSMPTAGSAGSV	PEN-LNLFPEPGSKSEEIGKKQLSKDSILSLYGS-QTP	QMPTQAMF			
hSMAP1	MGPTNIPFTSQAP	PAAFQGFP	SMGVPVPAAPGLIGNVMGQSP	SMMVG-MPMPNGFMGNAQT		
hSMAP2	MAPAQMAYPTAYP	SFPGVT	PPNSIMGSMMPPPVGMVAQPGASGMVAPMAMPAGYMGGMQA			
hSMAP1	GVMP	PLPQN	VVGFP	QGG--MVGQMGAPQSKFGLPQAQQPQWSLSQMNQQMAGMSISSATPTA		
hSMAP2	SMMGV	PNGMMTTQ	QAGYMAGMAAMPQTVYGVQPAQQLQWNL	TQMTQQMAGMNFYGANGMM		
hSMAP1	GF	GQPS	STTAGWS	GSSSGQTLSTQLWK		
hSMAP2	NY	GQSMS	---GGNGQAANQTLSPQMWK			

Figure 7 Alignement des protéines SMAP1 et SMAP2. Les séquences identiques sont surlignées en jaune et le domaine ArfGAP est borné par des crochets noirs.

3. Les fonctions cellulaires de SMAP1

Le gène SMAP1 a d'abord été décrit comme étant impliqué dans l'érythropoïèse [82] [83]. En effet, la protéine SMAP1 a été identifiée comme étant une glycoprotéine membranaire exprimée par les cellules stromales et étant impliquée dans l'effet stimulant de ces cellules sur l'érythropoïèse [83], d'où la dénomination initiale de ce gène, *Stromal Membrane-Associated Protein 1*. En 2005, le gène SMAP1 a été identifié comme étant fusionné avec le gène MLL (*Myeloid/Lymphoid Leukemia*) chez un patient atteint d'une leucémie [84]. Cette translocation n'a été décrite que chez ce patient et le rôle de SMAP1 dans l'érythropoïèse n'a pas été confirmé depuis.

Depuis 2005, le peu de littérature concernant SMAP1 (4 articles de la même équipe) porte sur son rôle dans l'endocytose clathrine-dépendante. En effet, une étude a montré que l'inhibition de l'expression de SMAP1 par ARN interférence, sa surexpression ou l'expression d'un mutant dont l'activité GAP n'est plus fonctionnelle, SMAP1(R61Q), aboutit à un défaut d'internalisation clathrine-dépendante du TfR dans la lignée cellulaire HeLa [77]. La surexpression de SMAP1 inhibe l'endocytose en empêchant le recrutement du complexe adaptateur AP-2 et de la chaîne lourde de la clathrine alors que le mutant SMAP1(R61Q) l'inhibe en empêchant le seul recrutement de la clathrine. Ainsi, d'après Tanabe et al., l'effet de SMAP1 sur l'endocytose clathrine-dépendante n'est pas contrôlée par son activité sur Arf6. Or, plusieurs articles de différentes équipes montrent que l'internalisation du TfR ne dépend pas d'Arf6 contrairement à son recyclage, cette notion sera discutée plus loin [10] [85]. De plus, toute cette étude a été menée en microscopie à fluorescence, méthode qualitative qui ne permet pas de quantifier les protéines internalisées ou recrutées. De plus, la forme transfectée de SMAP1 correspond à la protéine murine qui certes partage 86% d'homologie avec la protéine humaine [76] mais qui ne possède pas l'exon 5 alternatif (Gene/NCBI, identifiant gène 60682 pour

le gène humain et 983686 pour le gène murin). Cette même équipe a publié en 2008, un étude portant sur le rôle de SMAP1 dans l'endocytose clathrine-dépendante de la E-cadhérine, protéine majeure des jonctions adhérentes [79]. Dans cette étude, les auteurs ont travaillé sur les cellules MDCK (*Mardin-Darby Canine Kidney*) qui sont un modèle de cellules épithéliales polarisées de rein de chien dans lesquelles ils ont transfecté l'ADN complémentaire du gène SMAP1 murin, les protéines murine et canine partageant 89% d'identité [79]. Dans cet article, les auteurs concluent que la surexpression de la protéine SMAP1 inhibe l'internalisation clathrine-dépendante de la E-cadhérine au même titre que le TfR. Ainsi, les cellules sur-exprimant SMAP1 migrent moins que les cellules MDCK parentales. Cette observation vaut également pour les cellules exprimant le mutant SMAP1 R61Q dont l'activité GAP est nulle. Ainsi, selon cette étude, l'augmentation d'Arf6-GTP résultant de l'expression de SMAP1(R61Q) inhibe l'internalisation de la E-cadhérine. Or, dans plusieurs autres articles, l'activation d'Arf6 semble être indispensable à l'internalisation de la E-cadhérine [86] [87] et dans les tumeurs mammaires, il a été démontré que le niveau d'activation d'Arf6 corrèle avec l'invasivité des cellules [88].

Ainsi, les discordances entre les quelques travaux menés par la même équipe sur la protéine SMAP1 et les données de la littérature ne nous permettent pas de conclure sur le (ou les) processus cellulaire(s) dans le (ou lesquels) la protéine SMAP1 serait impliquée. Son rôle dans l'inactivation d'Arf6 par l'hydrolyse de son GTP en GDP [77] permet néanmoins d'émettre des hypothèses sur ces processus cellulaires. En effet, la protéine Arf6, petite protéine G de la super-famille Ras, est très étudiée (509 articles à ce jour) et a été impliquée dans plusieurs mécanismes cellulaires qui seront détaillés dans la suite de ce manuscrit.

III. La protéine Arf6 : rôles physiopathologiques

1. Les protéines Arf

La protéine Arf6 fait partie de la famille des protéines Arf (ADP ribosylation factor) et de la super-famille des petites protéines G Ras. Les protéines Arf, chez les eucaryotes, comptent 6 membres regroupés en 3 classes selon leur identité de séquence [89] [90]:

Classe I : Arf1, Arf2 et Arf3 partagent plus de 96% d'identité, Arf2 étant absente chez l'Homme,

Classe II : Arf4 et Arf5 partagent 90% d'identité et 81% d'identité avec Arf1 [91],

Classe III : Arf6, la plus divergente des protéines Arf, partage 66 à 70% d'identité avec les autres Arf (Tableau 2).

Classe	Nom	Localisation membranaire	Fonctions cellulaires	Références
I	Arf1	Golgi et endosomes	Transport Golgi-RE et intra-Golgi	D'Souza-Schorey et al, 2006, Mol Cell Biol
			Structure du Golgi	Fucini et al, 2000, J Biol Chem Fucini et al, 2002, Mol Biol Cell
			Régulation de la liaison de l'actine à la membrane du Golgi	
	Arf2	Absente chez l'Homme		Chun et al, 2008, Mol Biol Cell
	Arf3	Golgi	Formation des vésicules à partir du Golgi	Manolea et al, 2010, Mol Biol Cell
II	Arf4	Compartiment intermédiaire entre le RE et le Golgi	Tri de la rhodopsine dans le Golgi	Chun et al, 2008, Mol Biol Cell Deretic et al, 2005, PNAS
			Biogenèse des vésicules chromogranine A+ à partir du Golgi	Sadataka, 2011, FEBS
	Arf5	Compartiment intermédiaire entre le RE et le Golgi	Biogenèse des vésicules chromogranine A+ à partir du Golgi	Chun et al, 2008, Mol Biol Cell Sadataka, 2011, FEBS J
III	Arf6	Membrane plasmique et endosomes	Recyclage de protéines membranaires	Prigent et al, 2003, J Cell Biol D'Souza-Schorey et al, 2006, Mol Cell Biol Klein et al, 2006, J Biol Chem Montagnac et al, 2011, Curr Biol
			Remodelage de l'actine	Donaldson, 2003, J Biol Chem D'Souza-Schorey et al, 2006, Mol Cell Biol Klein et al, 2006, J Biol Chem
			Cytocinèse	D'Souza-Schorey et al, 2006, Mol Cell Biol Montagnac et al, 2009, Curr Biol Schweitzer et al, 2011, Semin Cell Dev Biol

Tableau 2: Les classes de protéines Arf

Les protéines Arf ont d'abord été décrites et nommées pour leur rôle de cofacteur lors de l'ADP-ribosylation de la sous-unité α des protéines G_s catalysée par la toxine du choléra [92]. Elles sont aujourd'hui étudiées pour leur rôle dans le trafic intracellulaire et la formation des vésicules de transport. Ce sont des protéines ubiquitaires, très conservées chez les eucaryotes. Elles sont myristoylées à leur extrémité N-terminale qui forme une hélice amphipathique permettant l'ancrage aux membranes [13]. Elles passent d'une conformation inactive, quand elles sont liées au GDP (*Guanine DiPhosphate*), à une conformation active, quand elles sont liées au GTP (*Guanine TriPhosphate*). Le cycle GTP-GDP est permis par l'action de protéines spécifiques, les GEF (*Guanine nucleotide Exchange Factor*) qui change le GDP en GTP et les GAP (*GTPase Activating Protein*) qui hydrolyse le GTP en GDP (Figure 7).

Des rôles spécifiques ont été décrits pour chaque protéine Arf mais de plus en plus de publications suggèrent qu'elles n'agiraient pas de manière indépendante [91] [93]. En effet, il a été démontré qu'Arf6-GTP permet le recrutement de la protéine GEF ARNO qui active à son tour Arf1 dans les endosomes [93]. La protéine Arf1 est localisée à la membrane de l'appareil de Golgi par son extrémité N-terminale amphiphile et ses interactions avec d'autres protéines de la membrane, considérées comme ses récepteurs (p23, p24 et la membrine) [89]. La protéine Arf1 est impliquée dans la formation des vésicules de transport à partir de l'appareil de Golgi. Elle recrute, une fois activée, les complexes protéiques COPI, protéines de structure enveloppant les vésicules de transport.

- Le cycle GTP/GDP

Contrairement à la plupart des protéines G, les protéines Arf n'ont pas d'activité GTPase intrinsèque détectable [94] [95]. Ainsi, les ArfGAP sont indispensables à

l'hydrolyse du GTP des protéines Arf et donc à leur inactivation (Figures 6B et 7). L'échange du GDP en GTP est catalysé par les protéines GEF. Toutes les ArfGEF identifiées à ce jour possèdent un domaine catalytique d'environ 200 acides aminés, le domaine Sec7 (Sec7d) [16] qui consiste en 10 hélices α , très conservé entre les espèces et suffisant à l'échange du GDP en GTP [96]. Il existe cinq familles de GEF classées selon leur structure et leur sensibilité à la brefeldin A (BFA) (Figure 9). La BFA est un métabolite fongique qui se lie au domaine catalytique Sec7 des GEF aboutissant à la formation d'un complexe abortif Arf-GDP-BFA-Sec7d empêchant le déplacement du GDP [97] [90].

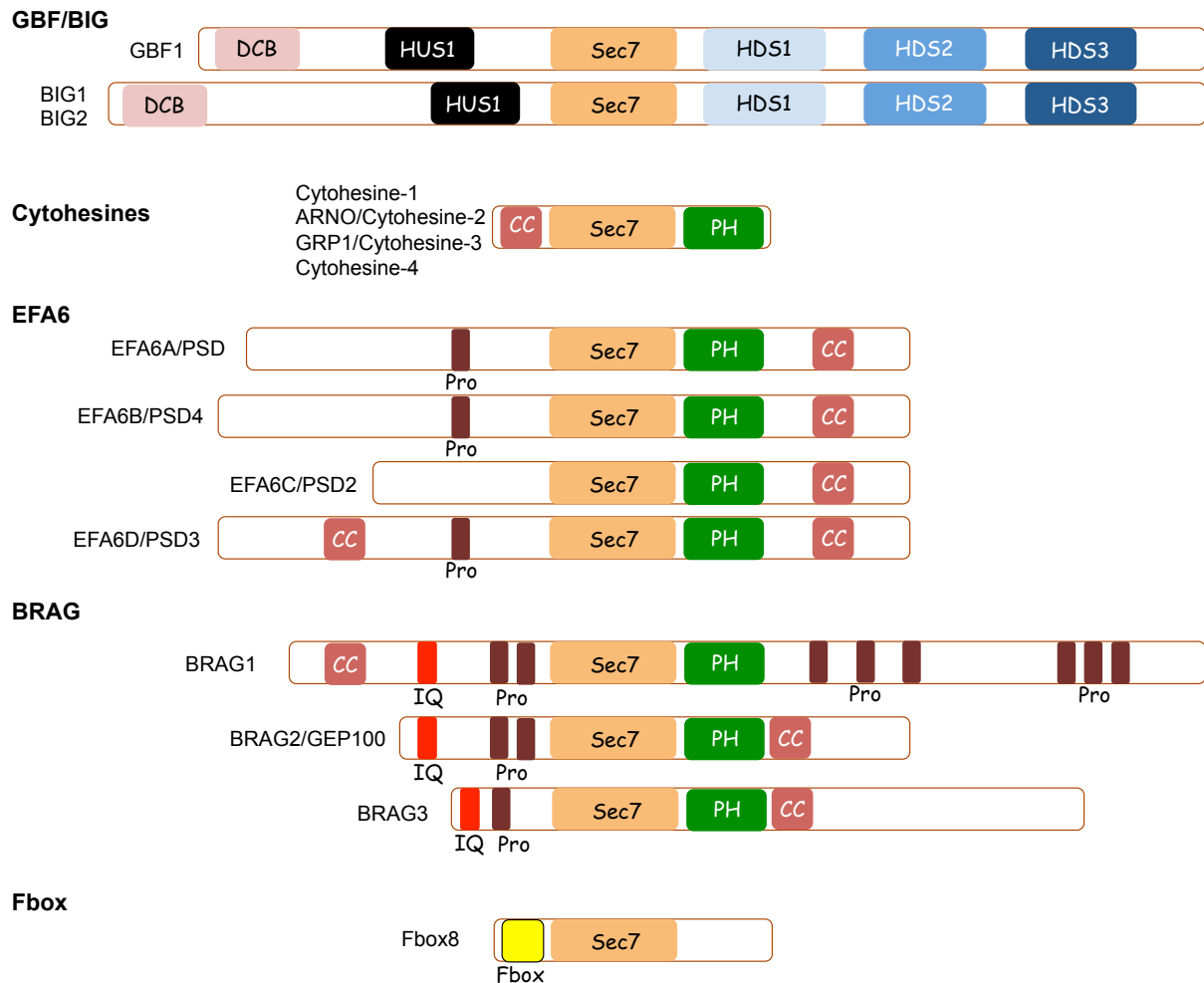


Figure 8: Les cinq familles des protéines ArfGEF. Structure des domaines fonctionnels des protéines ArfGEF. GBF : Golgi BFA-resistance factor ; BIG : BFA-inhibited GEF ; ARNO : Arf Nucleotide Binding site Opener ; EFA6 : Exchange Factor for Arf6 ; BRAG : Brefeldin-resistant ArfGEF ; DCB : Dimerization and Cyclophilin Binding; Sec7 : domaine catalytique ; CC : coiled-coil ; PH : domaine pleckstrin homology ; Pro : région riche en proline ; IQ : site de liaison à la calmoduline ; Fbox : domaine impliqué dans l'interaction avec la E3-ubiquitine ligase. *D'après Casanova, 2007, Traffic [16].*

Les protéines Arf1-5 sont solubles lorsqu'elles sont liées au GDP et membranaires quand elles sont liées au GTP. Arf6 est la seule protéine qui est stablement ancrée aux membranes quelque soit son état d'activation [91]. Des études de cristallographie ont permis de mettre en évidence les structures tridimensionnelles des protéines Arf1 et Arf6 selon qu'elles soient liées au GDP ou au GTP, plusieurs régions interviennent lors de l'activation par les GEF [95]. Liée au GDP, l'hélice N-terminale bloque la région « interswitch » dans une conformation rétractée qui empêche les protéines Arf de se lier aux effecteurs (Figure 10).

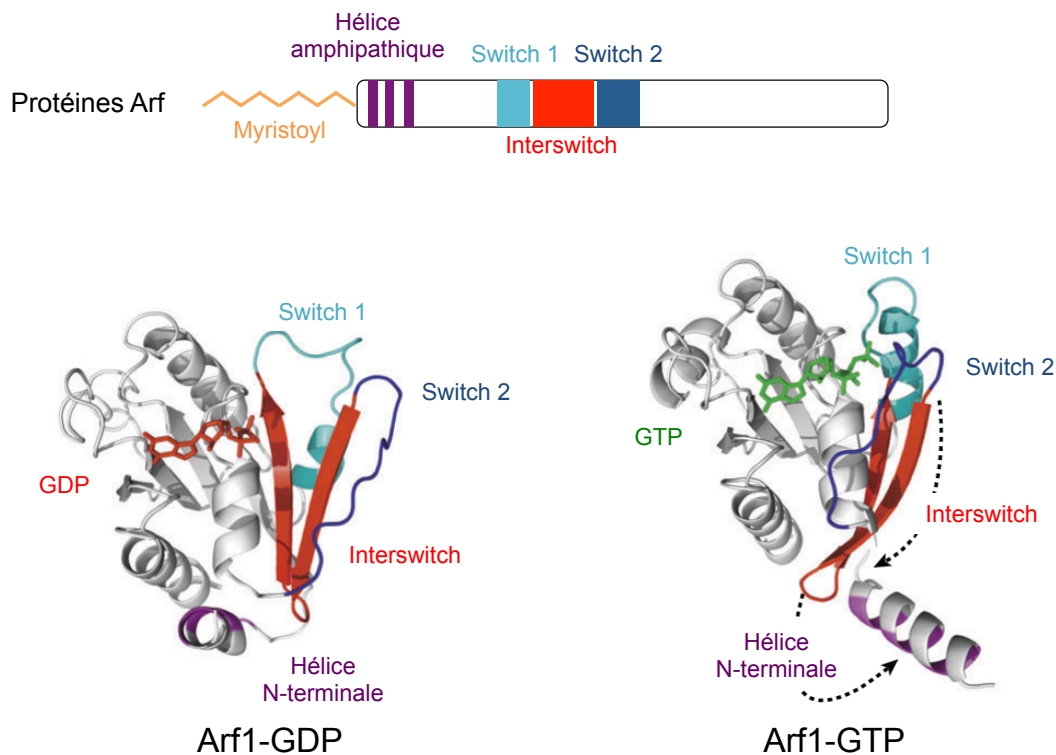


Figure 9: Le cycle GDP/GTP des protéines Arf. La liaison du GTP aboutit au déplacement de la région interswitch permettant le rapprochement des régions switch et au déplacement de l'hélice N-terminale qui devient apparente et qui peut se lier aux membranes. Les flèches en pointillés noirs représentent les mouvements de la région interswitch et de l'hélice N-terminale dus à la liaison du GTP. D'après Gillingham et Munro, 2007, *Annu Rev Cell Dev Biol* [13].

L'interaction de la région N-terminale avec les membranes *via* le PIP2 (Phosphatidyl Inositol 4,5 Phosphate) est indispensable à la liaison des protéines GEF aux Arf [96]. Le changement de conformation des Arf permet aux GEF de se lier aux régions « switch » I et II catalysant le changement du GDP en GTP [96]. La liaison du GTP accentue le déplacement des régions « switch » I et II permettant la liaison aux effecteurs [95]. C'est donc sous leur forme active que les protéines Arf transduisent un signal intracellulaire. Les protéines Arf1 et Arf6 ont une structure relativement similaire quand elles sont dans leur conformation active. La spécificité de liaison des effecteurs semble être due à la seule localisation subcellulaire des protéines Arf [98] puisqu'*in vitro*, plusieurs molécules effectrices sont capables d'interagir avec plusieurs Arf [99]. Il a été démontré, dans le cas d'Arf1, que l'échange du GDP en GTP permet un meilleur ancrage d'Arf1 à la membrane par le déplacement de l'extrémité N-terminale [100] et le recrutement des effecteurs tels que COPI (Coat Protein Complex I), complexe protéique de membrane de sept sous-unités [101], AP-1, AP-3 et AP-4 [102], complexes tétramériques adaptateurs nécessaires au recrutement de la clathrine. Le rôle des protéines Arf a été étudié en utilisant des mutants mimant la forme active ou inactive. Or il a été démontré que le cycle GDP/GTP est indispensable aux effets biologiques des Arf, notamment d'Arf6 [103] [104]. En effet, l'expression du mutant constitutivement actif Arf6(Q67L) ou du mutant inactif Arf6(T27N) bloque les effets d'Arf6. Plus récemment, un mutant dont le cycle GDP/GTP est plus rapide a été créé, Arf6(T157A) et permet d'étudier plus précisément les effets cellulaires d'Arf6. Ainsi, la régulation du cycle GTP/GDP est indispensable au bon fonctionnement des mécanismes cellulaires dans lesquels les protéines Arf sont impliquées.

2. La protéine Arf6

La protéine Arf6 est le membre le plus divergent de la famille des protéines Arf. Elle constitue à elle seule la classe III des protéines Arf. Contrairement aux autres membres de la famille, la protéine Arf6 est membranaire quelque soit son état d'activation. Elle est localisée à la membrane plasmique quand elle est liée au GDP et au niveau des endosomes quand elle est activée [85]. Plusieurs rôles ont été décrits dans divers mécanismes cellulaires tels que l'endocytose, aussi bien clathrine-dépendante qu'indépendante, le remodelage de l'actine, la migration et l'invasion cellulaires et la cytokinèse, étape finale de la mitose [89]. SMAP1 étant décrit comme ayant un rôle dans l'endocytose clathrine-dépendante (CME, *Clathrin Mediated Endocytosis*) du TfR [77] et de la E-cadhérine [79], je m'intéresserai plus spécifiquement, dans la suite de mon manuscrit, au rôle d'Arf6 dans l'endocytose clathrine-dépendante et dans la migration/invasion cellulaire.

2.1. Arf6 et endocytose clathrine-dépendante

Afin de survivre, la cellule a besoin de capter des éléments de la matrice extracellulaire et de renouveler ses protéines et lipides membranaires. Pour ce faire, plusieurs voies sont possibles selon la ou les protéine(s) à transporter et le type cellulaire: l'endocytose dépendante de protéines de structure telles que la clathrine [105] [106] et la cavéoline [107], l'endocytose indépendante de telles protéines de structure, la pinocytose, la transcytose ou la phagocytose.

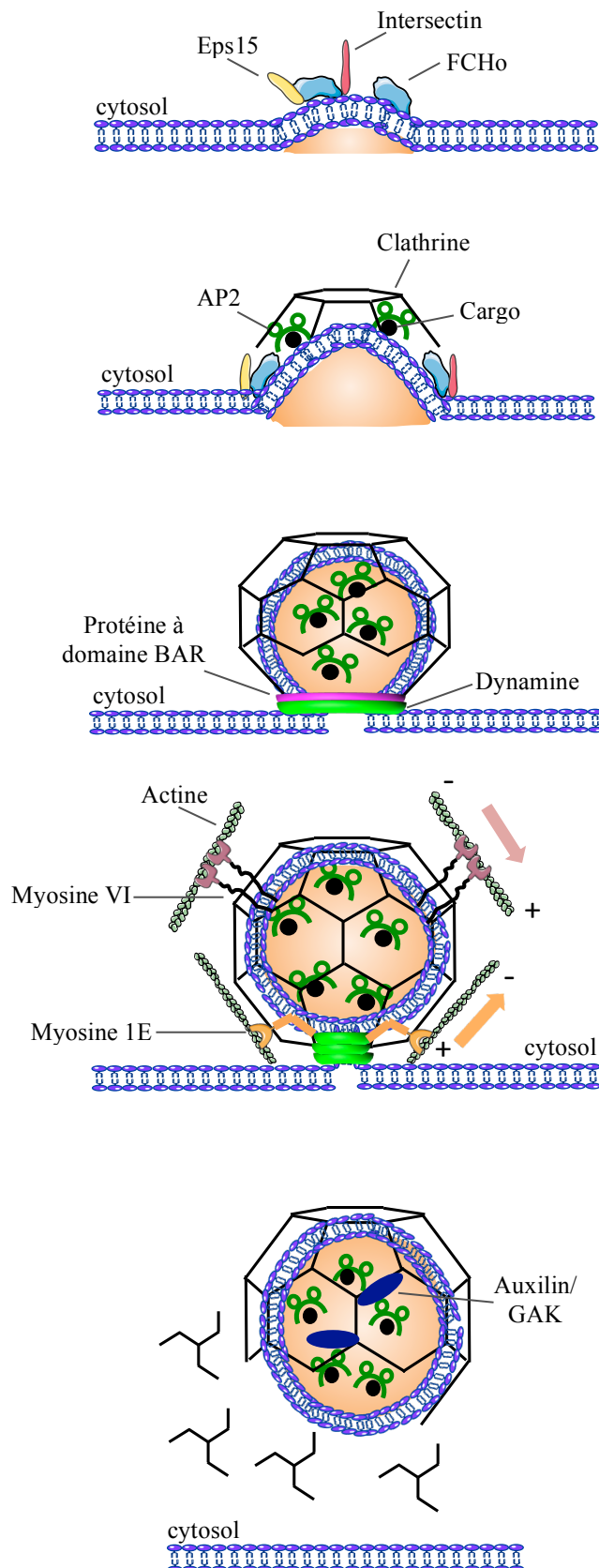
Parmi les deux voies d'endocytose dépendantes de protéines de structure connues à ce jour, l'endocytose clathrine-dépendante est la première à avoir été caractérisée par Barbara Pearse en 1975 [105]. En effet, bien que la description des « caveolae » (du latin « petites cavernes»), invaginations membranaires en forme de puits, date

des années 50 [108], la caractérisation structurelle n'a été décrite que 40 ans plus tard par la découverte de la cavéoline [107]. La CME a donc été, pendant de nombreuses années, considérée comme voie unique d'internalisation de molécules par des puits recouverts de protéines.

2.1.1. Endocytose clathrine-dépendante

Le type d'endocytose le plus étudié et le mieux caractérisé à ce jour est la CME. Ce mécanisme d'internalisation de molécules doit son nom à la protéine de structure qui recouvre les invaginations membranaires, la clathrine. La formation de ces invaginations puis des vésicules est un mécanisme complexe qui met en jeu plusieurs protéines à différentes étapes du processus. L'endocytose clathrine-dépendante se déroule en cinq étapes [1] (Figure 11):

- la nucléation : formation d'une invagination membranaire appelé le puits
- la sélection du cargo
- l'assemblage du manteau de clathrine
- la scission de la vésicule
- le retrait du manteau de clathrine.



La nucléation : Les protéines FCHo se lient à des régions membranaires riches en PIP2 afin d'accentuer la courbure de la MP. Elles recrutent les protéines Eps15 et Intersectin qui recrutent à leur tour le complexe AP2. La formation du CCP est initiée.

La sélection des cargos : Le complexe AP2 se lie à des séquences peptidiques spécifiques de récepteurs membranaires et recrute d'autres protéines adaptatrices reconnaissant d'autres cargos. Les protéines de la nucléation sont expulsées au bord du CCP.

L'assemblage du manteau de clathrine: Les triskélia de clathrine sont recrutés par les complexes AP2 et se polymérisent en un manteau de clathrine qui stabilise la courbure de la MP. Les protéines à domaine BAR et la dynamine commencent à entrer en action.

La scission de la vésicule: Les dimères de dynamine sont recrutés par les protéines à domaine BAR. L'hydrolyse du GTP permet le glissement des molécules de dynamine les uns sur les autres aboutissant à la constriction du cou de la CCV. La scission est achevée par l'action des myosines VI et 1E sur les filaments d'actine.

Le désassemblage du manteau de clathrine: L'auxilin/GAK se lie aux régions N-terminales de la clathrine et recrute l'ATPase Hsc70 qui initie le désassemblage du manteau de clathrine. L'activité phosphatase de la synaptojanin sur le PIP3 et le PIP2 déstabilise les liaisons des protéines adaptatrices à la MP.

Figure 10: L'endocytose clathrine-dépendante. D'après McMahon, 2011, *Nat Rev Mol Cell Biol* [1].

- La nucléation

La première étape de la CME est la nucléation. Elle consiste en la mise en place de plusieurs types de protéines adaptatrices qui permettront le recrutement de la clathrine (Figure 11). La nucléation permet la formation de puits recouverts de clathrine (CCP, *Clathrin Coated Pits*). Ce processus serait aléatoire [109] et la stabilisation des CCP se ferait par le recrutement des protéines adaptatrices. Ce recrutement est permis par différents changements moléculaires localisés : la composition lipidique de la membrane plasmique (MP), les modifications conformationnelles ou post-traductionnelles des récepteurs membranaires activés par leur ligand [110].

Le changement conformationnel des récepteurs activés permet le démasquage de séquences peptidiques reconnues par les protéines adaptatrices (Table 3) [8] [110]. Ces protéines sont de deux types [110]:

- la protéine hétéro-tétramérique, AP2 (*Adaptor Protein*), constituée de deux grandes sous-unités α et β 2, d'une sous-unité moyenne μ 2 et enfin d'une petite sous-unité σ 2 ;
- les protéines dites monomériques, ARH (*Autosomal Recessive Hypercholesterolemia protein*), Dab2 (*Disabeled 2*), β -arrestine, numb, epsine, Eps15 (*EGFR Pathway Substrate 15*) et CALM (*Clathrin Assembly Lymphoid Myeloid Leukemia*).

Il a récemment été démontré [111] que la nucléation dépend d'un complexe protéique mettant en jeu, en plus des protéines adaptatrices, une famille de protéines à domaine F-BAR (FCH-Bin/Amphiphysin/Rvs), FCHo 1/2 (*Fer/Cip4 Homology domain only protein*). Les protéines FCHo sont ubiquitaires et localisées aux CCP, dans les régions enrichies en phosphatidylinositol-4,5-biphosphate (PIP2) auquel

elles se lient. Elles seraient indispensables à la formation des CCP, puisque leur inhibition par siARN diminuent de manière drastique le nombre de CCP à la MP et l'endocytose du TfR [111]. Les protéines FCHo possèdent un homodimère de domaine F-BAR permettant d'accentuer la courbure de la MP au niveau des CCP aussi faible soit elle [111]. Les protéines FCHo permettent le recrutement des protéines adaptatrices Eps15 et intersectin 1/2 aux CCP qui elles-mêmes permettent le recrutement du complexe AP2 et de la clathrine [111].

- La sélection des cargos

Cette étape met en jeu les protéines adaptatrices présentes au moment de la nucléation, ainsi ces deux étapes ne sont pas complètement dissociables (Figure 11). La liaison du ligand à son récepteur aboutit à un changement conformationnel au niveau de la queue cytoplasmique. Les protéines adaptatrices peuvent ainsi reconnaître et se lier aux séquences peptidiques démasquées ou aux polyubiquitines dans le cas du récepteur de l'EGF (*Epidermal Growth Factor*) [112]. Les protéines Eps15 et epsin sont capables de reconnaître les cargos ubiquitylés par leurs motifs d'interaction à l'ubiquitine (UIM). Ces protéines adaptatrices sont capables de trier les cargos vers une endocytose clathrine-indépendante ou une CME selon qu'ils sont mono- ou polyubiquitylés respectivement [113] [112]. Eps15 interagit également avec des protéines de tri plus tardif, comme Hrs et les protéines ESCRT [114]. La liaison de l'epsin à la MP permet, *via* son domaine ENTH (*Epsin N-Terminal Homology*), d'augmenter la courbure du CCP [1].

Dans le cas des protéines membranaires non ubiquitylées, le message d'internalisation est intrinsèque à la protéine. Le complexe AP2, dans sa conformation active, reconnaît des séquences peptidiques consensus (Table 3). En effet, il a été démontré que la liaison d'AP2 aux cargos dépend de la phosphorylation

de la thréonine 156 de la sous-unité $\mu 2$ [115] [116]. Cette phosphorylation est réalisée par la kinase AAK1 (*α -adaptin-Associated Kinase 1*) [117] et permet le changement de conformation de la sous-unité $\mu 2$, rendant accessible le domaine de liaison aux cargos et augmentant l'affinité d'AP2 pour le PIP2 [114]. La kinase AAK1 est localisée dans les CCP et son activité est stimulée par la clathrine [8] permettant l'enrichissement des CCP en cargos. Le complexe AP2/cargo active la PIP5 kinase (PIP5K) [118] qui permet la phosphorylation du phosphatidylinositol-4-phosphate (PIP) en PIP2. De plus, Arf6-GTP active la PIP5K et la phospholipase D produisant de l'acide phosphatidique qui à son tour joue le rôle de cofacteur dans l'activation de la PIP5K [89]. Ainsi, la quantité de PIP2 augmente localement et permet la croissance du CCP.

Protéine adaptatrice	AP2	ARH	Dab2	β -arrestin	Numb	Epsin	Eps15
Cargo sélectionné	RTK	LDLR	Megalin et LDLR	RCPG	Notch	Récepteurs mono-Ub	Récepteurs poly-Ub
Séquence reconnue	YXX Φ ou [D/E]XXXL[L/I]	FXNPXY	FXNPXY	Phosphorylation	NPXY	Ubiquitine	Ubiquitine

Tableau 3: Les principales protéines adaptatrices qui ont un rôle dans la sélection des cargos. X : acide aminé quelconque ; Φ : résidu hydrophobe volumineux. AP2 : Adaptor Protein 2 ; ARH : Autosomal Recessive Hypercholesterolemia ; Dab2 : Disabled 2 ; Eps15 : EGFR Pathway Substrate 15 ; RTK : Récepteur Tyrosine Kinase ; LDLR : Récepteur aux Low-Density-Lipoprotein ; RCPG ; Récepteur Couplé aux Protéines G ; Ub : Ubiquitine

- L'assemblage du manteau de clathrine

Une fois les cargos sélectionnés, le manteau de clathrine peut être assemblé (Figure 11). La forte concentration locale de complexes AP2, notamment *via* la protéine Eps15 capable de lier plusieurs complexes AP2 à la fois, permet le recrutement et la polymérisation de la clathrine. La liaison d'AP2 à la clathrine déplace Eps15 qui se retrouve expulsé au bord de la vésicule en formation, où elle pourra recruter de

nouvelles molécules AP2, permettant ainsi la croissance de la vésicule recouverte de clathrine (CCV, *Clathrin-Coated Vesicle*) [114]. Le manteau de clathrine permet la stabilisation de la courbure de la MP mais n'en est pas à l'origine [1].

La clathrine est la première protéine de structure de l'endocytose identifiée [105], et a fait l'objet de nombreuses recherches (6196 articles référencés sur PubMed à ce jour). La clathrine forme autour de la vésicule à internaliser une structure tridimensionnelle en forme de panier, le « manteau de clathrine » (Figure 12C). Ce dernier est constitué de plusieurs triskélia de clathrine chacun formé de trois chaînes lourdes (CHC, $\approx 190\text{kDa}$) (Figure 12A) et de trois chaînes légères (CLC, $\approx 25\text{kDa}$) [119] [120] [121] (Figure 12B).

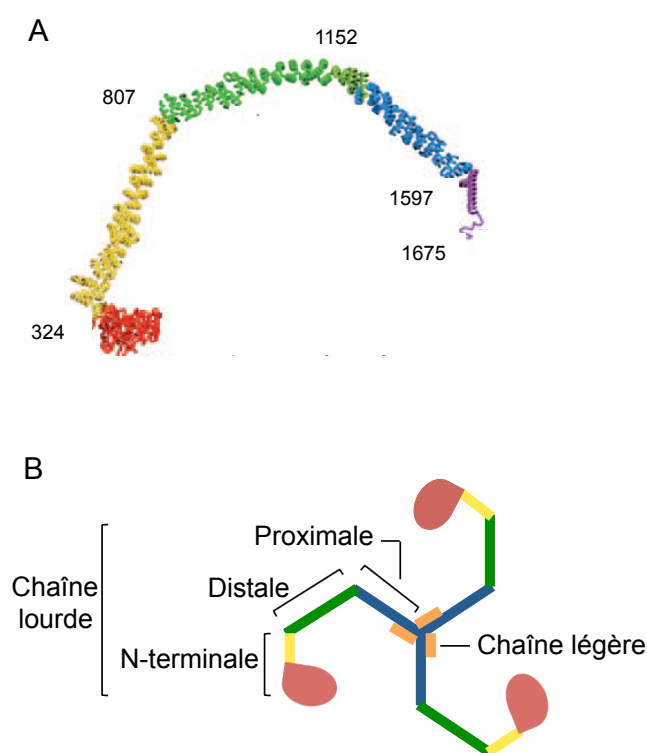


Figure 11: Représentation schématique de la clathrine. A. Monomère de chaîne lourde B. Un triskélion de clathrine formé de 3 chaînes lourdes et de 3 chaînes légères. C. Panier de clathrine avec un triskélion en bleu. D'après Edeling et al., 2006, *Nature Review* [8].

La trimérisation des CHC a lieu au niveau de la partie C-terminale de chaque monomère. Les triskélia de clathrine s'assemblent en « panier » de manière

spontanée en présence de calcium et de magnésium à un pH compris entre 6 et 6,5 *in vitro* [119]. La taille et la forme des paniers varient selon le type cellulaire [106]. L'amélioration des techniques de microscopie électronique a permis de visualiser les paniers de clathrine avec une résolution de plus en plus importante (Figure 13A). Ainsi, à ce jour, il est possible de visualiser la structure exacte d'un panier de clathrine (Figure 13) [8]. Les CLC se retrouvent sur la face externe du panier (Figure 13B) alors que les régions N-terminales des CHC se retrouvent sur la face interne (Figure 13C) où elles interagissent avec le domaine « clathrin-box » de protéines liées directement ou indirectement à la membrane plasmique. C'est le cas de l'auxilline, protéine nécessaire au retrait du « manteau de clathrine » [8]. Le rôle de la CLC est très peu connu, sa position externe lui permettrait d'interagir avec des protéines cytosoliques [122].

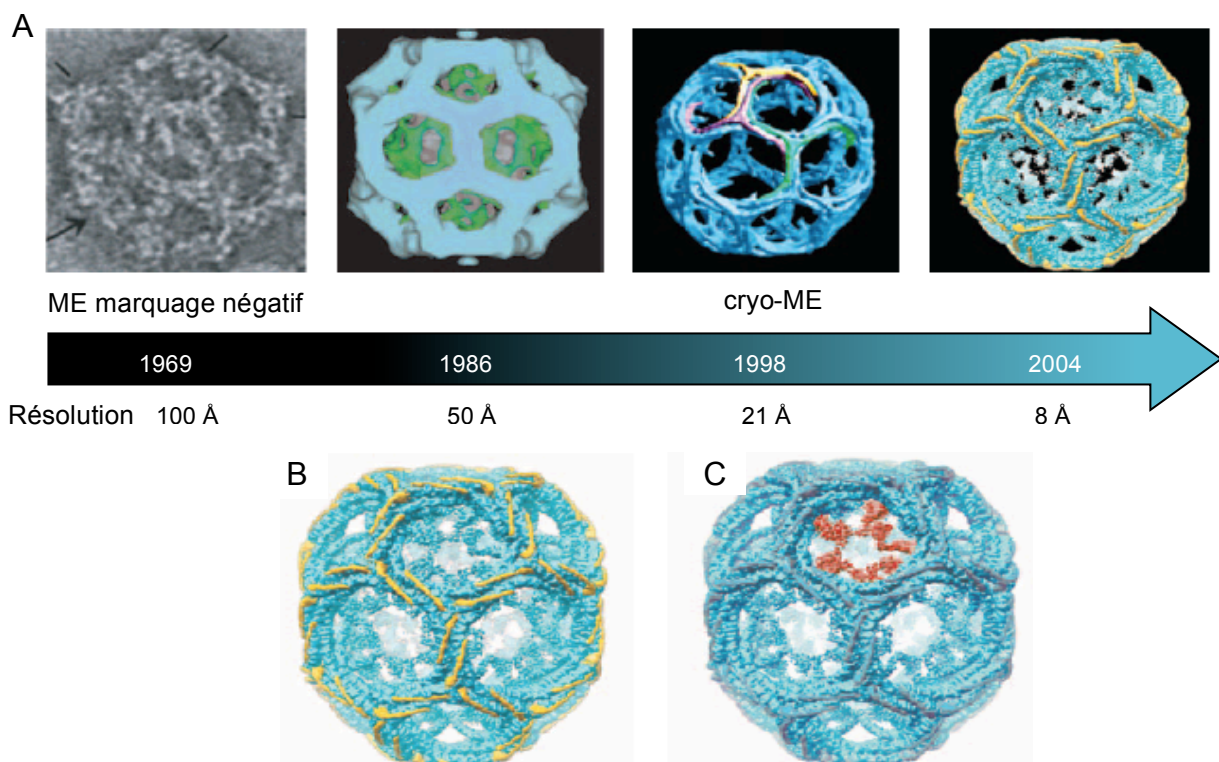


Figure 13: Structure tridimensionnelle d'un panier de clathrine. A. Évolution des techniques de microscopie électronique et amélioration de la résolution. B. Reconstruction tridimensionnelle d'un panier de clathrine. En bleu, les chaînes lourdes et en jaune, les chaînes légères. C. Reconstruction tridimensionnelle d'un panier de clathrine. En bleu, les chaînes lourdes et en rouge les régions N-terminales. D'après Edeling et al., 2006, *Nature Review* [8].

- La scission de la CCV

Le mécanisme de scission de la CCV n'est pas très bien connu à ce jour. Une fois la CCV formée, des protéines contenant des domaines BAR, permettant la courbure de la MP, liées à la MP depuis la formation du CCP entrent en action (Figure 11). C'est le cas de l'amphiphysin, de la SNX9 (*Sorting Nexin 9*), et de l'endophilin. Ces trois protéines ont un domaine SH3 capable de lier la dynamine, petite protéine G impliquée dans la scission des vésicules aussi bien dans la CME que dans d'autres mécanismes d'endocytose. Les protéines à domaine BAR sont recrutées par AP2 et la clathrine au bord du CCP, endroit de forte courbure, dès l'étape de nucléation mais ne sont pas actives, car elles s'auto-inhibent. L'amphiphysin est capable de lier la clathrine, *via* son domaine N-terminal, et la dynamine, par l'intermédiaire de son domaine SH3 C-terminal, mais ces interactions sont mutuellement exclusives [123]. Ainsi, au moment de la nucléation, l'amphiphysin recrutée par AP2 et la clathrine ne peut pas lier la dynamine. À la fin du processus d'internalisation, une fois la membrane suffisamment courbée par l'amphiphysin, les dimères de dynamine sont recrutés par l'amphiphysin et déplacent la clathrine [123] [114]. Une fois la dynamine recrutée, l'hydrolyse de son GTP permet un glissement des dimères les uns sur les autres aboutissant à la constriction du cou de la CCV et à la libération de la CCV dans le cytoplasme [124]. Cette dernière étape est controversée. En effet, d'autres études ont montré que la seule constriction du cou de la CCV n'est pas suffisante à la scission [125], suggérant l'existence d'autres forces. Il a récemment été publié que les myosines VI [126] et 1E [127] seraient impliquées dans la CME et produiraient les forces nécessaires à la scission [110].

- Le désassemblage du manteau de clathrine

Une fois la CCV détachée de la MP, le manteau de clathrine doit être retiré, ainsi l'endosome peut fusionner avec le compartiment cible (Figure 11). Après la scission de la CCV, le manteau de clathrine est discontinu, puisqu'à l'endroit du cou de la CCV, l'assemblage n'a pas pu avoir lieu [1]. C'est à cet endroit que le désassemblage serait initié [1]. Après la scission, l'endophilin recrute la synaptojanin, une phosphoinositide phosphatase qui convertit le PIP3 en PIP2 et le PIP2 en PI(4)P. Ainsi, la concentration de PIP2 à la membrane de l'endosome diminue, perturbant ainsi les liaisons des protéines adaptatrices telle que le complexe AP2. Ce changement lipidique à la membrane de l'endosome permet également le recrutement de protéines du désassemblage : auxilin/GAK. L'auxilin, protéine spécifique du tissu nerveux, et la protéine GAK (*cyclin G-Associated Kinase*), protéine ubiquitaire, sont des protéines homologues [114]. Elles sont capables de lier le PI(4)P de la membrane par leur domaine PTEN-like (*Phosphatase and Tensin homolog*) et la région N-terminale de la clathrine par leur « clathrin box » [128]. Auxilin/GAK recrutent, en présence d'ATP, la protéine chaperonne et ATPase Hsc70 (*Heat Shock Cognate 70*), dont elles sont les co-facteurs [129] [130]. Une fois le complexe auxilin (GAK)/Hsc70 formé, le désassemblage du manteau de clathrine peut être achevé par la rupture des interactions clathrine-clathrine.

2.1.2. Le recyclage des récepteurs membranaires

L'endocytose couple l'internalisation et le recyclage des protéines et des lipides membranaires. En effet, avant même le Grenelle de l'environnement, la cellule a mis en place des mécanismes de recyclage des protéines membranaires qui permettent d'économiser l'énergie nécessaire à la transcription et à la traduction. Ces

mécanismes de recyclage sont essentiels à la cellule et entrent en jeu dans divers processus cellulaires comme l'invasion, la morphogenèse ou encore la cytokinèse [10]. Les compartiments cellulaires du recyclage sont en partie communs aux cargos internalisés de manière clathrine-dépendante ou -indépendante. Le recyclage à la MP peut être dissocié en deux voies : le recyclage rapide et le recyclage lent. Ces deux voies de recyclage n'engagent pas les mêmes acteurs moléculaires ni les mêmes compartiments intracellulaires (Figure 14).

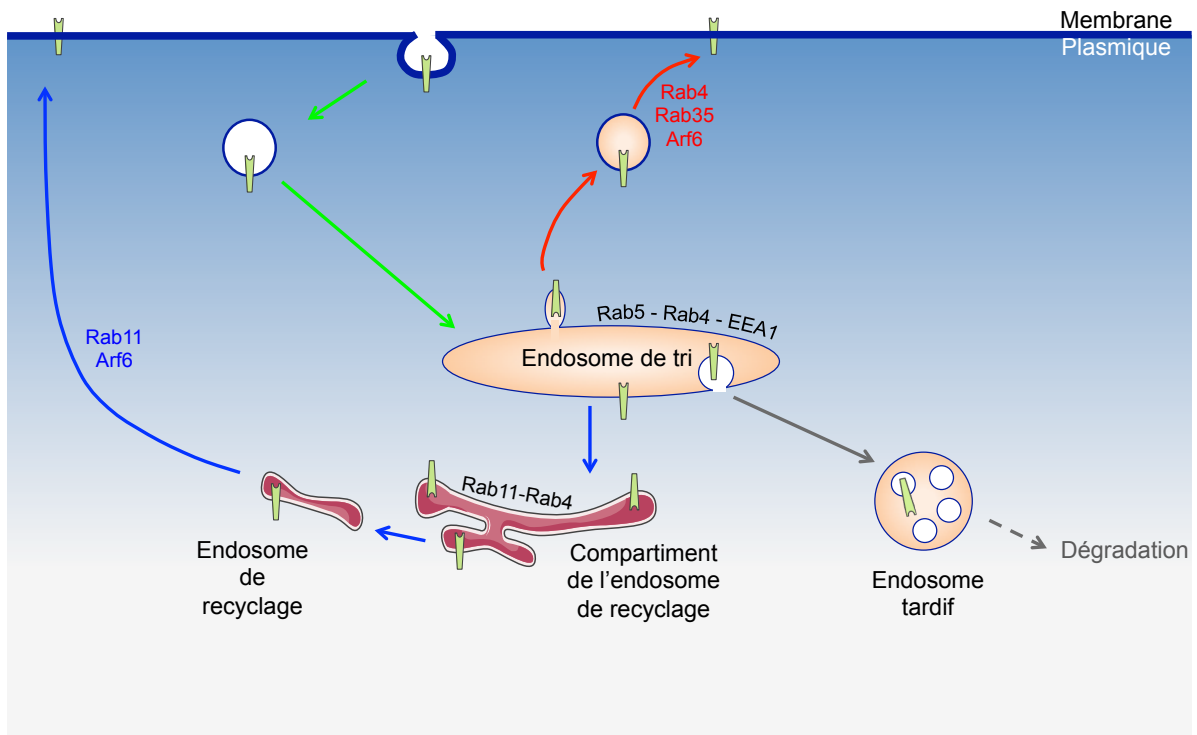


Figure 14: Le recyclage des protéines membranaires internalisées par la voie clathrine-dépendante. Les flèches vertes représentent l'internalisation clathrine-dépendante ; les flèches rouges représentent le recyclage rapide ; les flèches bleues représentent le recyclage lent ; les flèches grises représentent les voies de dégradation. *D'après Grant, 2009, Nat Rev Mol Cell Biol* [10].

- Le recyclage rapide

Le recyclage dit « rapide » est mesurable au bout de quelques minutes après le début de l'internalisation. Les premières études réalisées portaient surtout sur le TfR dont la phase rapide de recyclage a lieu dès cinq minutes [131]. L'endosome, une fois libéré de son manteau de clathrine, fusionne avec d'autres endosomes puis avec l'endosome de tri [132]. Cette organelle est souvent englobée avec l'endosome de recyclage dans le terme « endosome précoce » [132] [133]; ainsi dans la suite de mon manuscrit, je ferai la distinction entre ces deux organelles et ne citerai pas l'endosome précoce. L'endosome de tri est la première organelle à recevoir les cargos internalisés à partir de la MP [133] (Figure 14). Comme son nom l'indique, l'endosome de tri est impliqué dans le tri des cargos qu'il reçoit. Trois destinations sont possibles : la MP, le compartiment de l'endosome de recyclage (ERC, *endosomal recycling compartment*) ou l'endosome tardif [132]. Contrairement au tri qui a lieu lors de l'internalisation qui est dépendant de signaux (séquences peptidiques ou modifications post-traductionnelles), le tri réalisé par l'endosome dépend de sa géométrie [134] [132]. En effet, l'endosome de tri est un organite qui forme plusieurs tubules dans lesquels se retrouvent concentrés les récepteurs membranaires débarrassés de leur ligand par le pH acide de l'endosome de tri (pH $\approx$ 6) [133]. Cet enrichissement en protéines membranaires dans les tubules de l'endosome de tri est dû à un ratio {surface membranaire/volume} plus important que dans la partie vésiculaire de l'organelle [132, 133]. Ainsi, les protéines membranaires sont par défaut recyclées à la MP alors que le tri vers l'endosome tardif est dépendant de la présence d'un signal de localisation [132]. L'endosome de tri est une organelle qui mature en endosome tardif [134]. Il ne reçoit d'endosomes que pendant quelques minutes, après lesquelles son pH s'acidifie, passant de $\approx$ 6 à $\approx$ 5,5, empêchant la fusion de nouvelles vésicules. Cette étape serait le début de la maturation mais ce mécanisme reste encore peu connu à ce jour [132].

D'un point de vue moléculaire, la fusion des endosomes avec l'endosome de tri est un processus qui met en jeu plusieurs types de protéines. La régulation de la fusion est réalisée par des protéines de la famille Rab, petites protéines G impliquées dans de nombreux processus cellulaires. Au niveau de l'endosome de tri est localisée l'isoforme Rab5 [135] qui est impliquée dans la régulation de l'entrée des protéines cargos dans l'endosome de tri, dans la formation de PI(3)P, phosphoinositide enrichi dans l'endosome de tri, et dans la mobilité de cette organelle le long des microtubules et des filaments d'actine [133]. Comme toute protéine G, la protéine Rab5 cycle entre sa forme active, liée au GTP, et sa forme inactive, liée au GDP. C'est une fois active que la protéine Rab5 est localisée à la membrane des endosomes de tri et peut recruter ces effecteurs [133]. L'un des premiers effecteurs recrutés est la PI(3)P-kinase qui génère le PI(3)P à la membrane de l'endosome de tri et facilite ainsi le recrutement d'autres protéines qui se lient simultanément à Rab5 et au PI(3)P [136]. C'est le cas de la protéine EEA-1 (*Early Endosomal Antigen-1*) dont la liaison à Rab5 et au PI(3)P est essentielle à la fusion homotypique de l'endosome de tri, c'est-à-dire la fusion entre plusieurs endosomes de tri [137]. Cette fonction de la protéine EEA-1 est due à sa capacité à lier directement les protéines t-SNARE (*target-SNAP Receptor*), syntaxine 6 et 13 [138, 139] et ainsi à les recruter à la membrane.

La protéine Rab5 n'est pas la seule protéine de la famille Rab à être localisée à la membrane de l'endosome de tri, c'est également le cas de la protéine Rab4 [140] ou encore de la protéine Rab35 [141]. Rab4 est impliquée dans la régulation de la sortie des protéines triées vers la MP ou vers l'endosome de recyclage [133]. Quant à Rab35, elle a été décrite comme impliquée dans le recyclage rapide du TfR [141].

En plus des protéines de la famille Rab, d'autres petites protéines G sont impliquées dans le recyclage rapide, c'est le cas des protéines de la famille Arf et plus

précisément Arf6. En effet, il a été démontré qu'Arf6 est impliquée dans le recyclage rapide du TfR et que l'hydrolyse du GTP d'Arf6 est nécessaire à ce recyclage [85]. Le mécanisme moléculaire permettant cet effet serait lié à la capacité d'Arf6 d'interagir indirectement avec des protéines motrices spécifiques des microtubules (MT). En effet, deux nouveaux effecteurs d'Arf6-GTP ont récemment été identifiés : JIP3 et JIP4 (*c-Jun N-terminal-kinase Interacting Protein*) [3]. Ces protéines ont été décrites comme ayant un rôle dans la cascade JNK/MAPK (*Mitogen-Activated Protein Kinases*) [142] mais également comme ayant la capacité de lier les protéines motrices des MT : la kinésine 1, *via* sa chaîne légère, et le complexe dynactine, *via* ses sous-unités p150^{glued} et p50 [3]. JIP4 et JIP3 sont des protéines homologues respectivement ubiquitaire et neuronale. L'interaction avec Arf6-GTP se fait par l'un de leurs deux domaines de type « leucine-zipper » (LZ II) et est spécifique de la forme active d'Arf6, les protéines JIP3/4 étant incapables d'interagir avec Arf1 et Arf5 quelque soit leur état d'activation [3]. Le domaine LZ II est également celui par lequel JIP3/4 interagit avec la chaîne légère de la kinésine 1 [143], les interactions étant mutuellement exclusives. Ainsi, le complexe Arf6-GTP/JIP4 empêche la liaison de JIP4 à la kinésine 1 et permet celle du complexe dynactine, *via* ses sous-unités p150^{glued} et p50. Ainsi, quand Arf6 est activé, la vésicule cargo se dirige vers l'extrémité négative des MT, c'est-à-dire au centre de la cellule. Une fois le GTP d'Arf6 hydrolysé par une GAP spécifique, la liaison d'Arf6 avec JIP4 est rompue et JIP4 est libre d'interagir avec la kinésine 1 permettant le déplacement de la vésicule cargo vers l'extrémité positive des MT, c'est-à-dire vers la périphérie, permettant ainsi le recyclage de la vésicule à la MP [3] (Figure 15).

- Le recyclage lent

Cette étape du recyclage des protéines membranaires est observée 15 à 30 minutes après le début de l'internalisation dans le cas du TfR [131]. Ce « retard » par rapport au recyclage rapide est dû au fait que les protéines, plutôt que de recycler directement à partir de l'endosome de tri, passent par un autre compartiment intracellulaire, l'endosome de recyclage (ERC) (Figure 14). Cette organelle résulterait de la maturation de l'endosome de tri qui donnerait d'une part un compartiment multi-vésiculaire, le *multivesicular body* (MVB), généré à partir de la partie principale de l'endosome de tri, et d'autre part l'ERC, généré à partir des tubules de l'endosome de tri [132]. L'ERC est un organite complexe formé de plusieurs tubules maintenus par le réseau de MT plus ou moins proche du centromère selon le type cellulaire [132] [10]. Lors de sa maturation, l'endosome de tri perd l'expression de Rab5, se met à exprimer Rab11, autre membre de la famille Rab spécifique de l'ERC, et continue d'exprimer Rab4 [144]. Les effecteurs de Rab11, définis comme les Rab11FIP (*Rab11-Family Interacting Protein*), jouent un rôle important dans le positionnement de l'ERC dans la cellule, c'est le cas de la protéine Rab11FIP3 qui interagit à la fois avec Rab11 et Arf6 [10]. Les protéines Rab11 et Rab4 doivent être activées pour permettre le recyclage à la MP du TfR [145].

Le rôle d'Arf6 dans le recyclage lent est en partie dû à sa capacité à activer la phospholipase D (PLD), présente au niveau de l'ERC, qui génère l'acide phosphatidique (PA) et le diacylglycérol (DG) à partir de la phosphatidylcholine. Le PA serait impliqué dans la formation des vésicules de recyclage alors que le DG serait impliqué dans la fusion des vésicules avec la MP [10]. De plus, la régulation de la polymérisation de l'actine et du mouvement des vésicules cargos le long des MT confèrent à Arf6 un rôle prédominant dans le déplacement des vésicules de recyclage de l'ERC à la MP [10] [3] (Figure 15). Le recyclage du TfR à la MP est également régulé par une Arf6-GAP, ACAP1, qui est capable de reconnaître une

séquence peptidique de la partie cytoplasmique du TfR et ainsi permettre son recyclage [146]. Les mutations de cette séquence retardent le recyclage du récepteur mais ne l'abolissent pas, suggérant l'existence d'autres mécanismes [146].

- Recyclage et cytokinèse

Au cours des premières étapes de la mitose (prophase-métaphase), la surface membranaire de la cellule en division diminue, ceci étant la conséquence du blocage du recyclage des vésicules à la MP pendant que l'internalisation continue [147]. À partir de l'anaphase et ce jusqu'à la télophase, les deux cellules filles croissent et encore plus pendant la dernière étape de la mitose, la cytokinèse [147]. Cette étape s'achève par la séparation des deux cellules filles au moment de l'abscission. Dans son souci de parcimonie, la cellule a mis en place des mécanismes moléculaires lui permettant d'économiser la production de membranes.

La CME est indispensable à la cytokinèse [148]. La CME a lieu aux pôles de la cellule tout au long de mitose [147], les vésicules sont transportées le long du fuseau mitotique jusqu'au pont intercellulaire (*midbody*), région centrale de la cellule en division, où le recyclage a lieu [148]. La reprise du recyclage des vésicules issues de l'ERC est indispensable à la cytokinèse et requiert les protéines Rab11 et Arf6 [149] [9]. La protéine Rab11-FIP3, qui interagit à la fois avec Rab11 et Arf6 formant ainsi un complexe ternaire, est indispensable à l'étape d'abscission puisque son inhibition bloque la séparation des deux cellules filles [149]. L'expression d'un mutant constitutivement actif d'Arf6, Arf6(Q67L), inhibe l'étape d'abscission aboutissant à des cellules binuclées [148]. Cette observation pourrait s'expliquer par l'effet d'Arf6 sur le déplacement des vésicules cargos le long des microtubules. En effet, l'interaction d'Arf6-GTP avec la protéine JIP4 favorise la liaison de cette dernière au complexe dynactine déplaçant les vésicules vers le centre de la cellule le long des

MT donc à l'encontre du recyclage à la MP [3]. Cependant, au moment de la cytokinèse, un pic furtif d'Arf6-GTP est détecté [150]. Cette augmentation d'Arf6-GTP localisée au pont intercellulaire est nécessaire à la fusion des vésicules de recyclage puisqu'Arf6-GTP est capable de lier la sous-unité Sec10 de l'exocyste, complexe octamérique impliqué dans l'accrochage des vésicules à la MP [151] (Figure 15). Cette interaction est renforcée par la liaison de Rab11 à la sous-unité Sec15 de l'exocyste [152].

Les vésicules issues de l'ERC ne sont pas les seules vésicules impliquées dans la cytokinèse. En effet, des vésicules issues de l'endosome de tri ou encore de l'appareil de Golgi sont nécessaires à la cytokinèse [10]. Les vésicules issues de l'endosome de tri sont Rab35-positives [141]. Cet autre membre de la famille des petites protéines G Rab est essentiel à l'étape d'abscission. Étant issues de l'endosome de tri, les vésicules Rab35-positives sont enrichies en PIP2, phosphoinositide nécessaire à la séparation des deux cellules filles par le recrutement d'autres protéines au niveau du pont intercellulaire [141]. L'enrichissement en PIP2 peut également s'expliquer par la présence du pic furtif d'Arf6-GTP au pont intercellulaire, Arf6-GTP activant la PIP5-kinase aboutissant à la production de PIP2 [89].

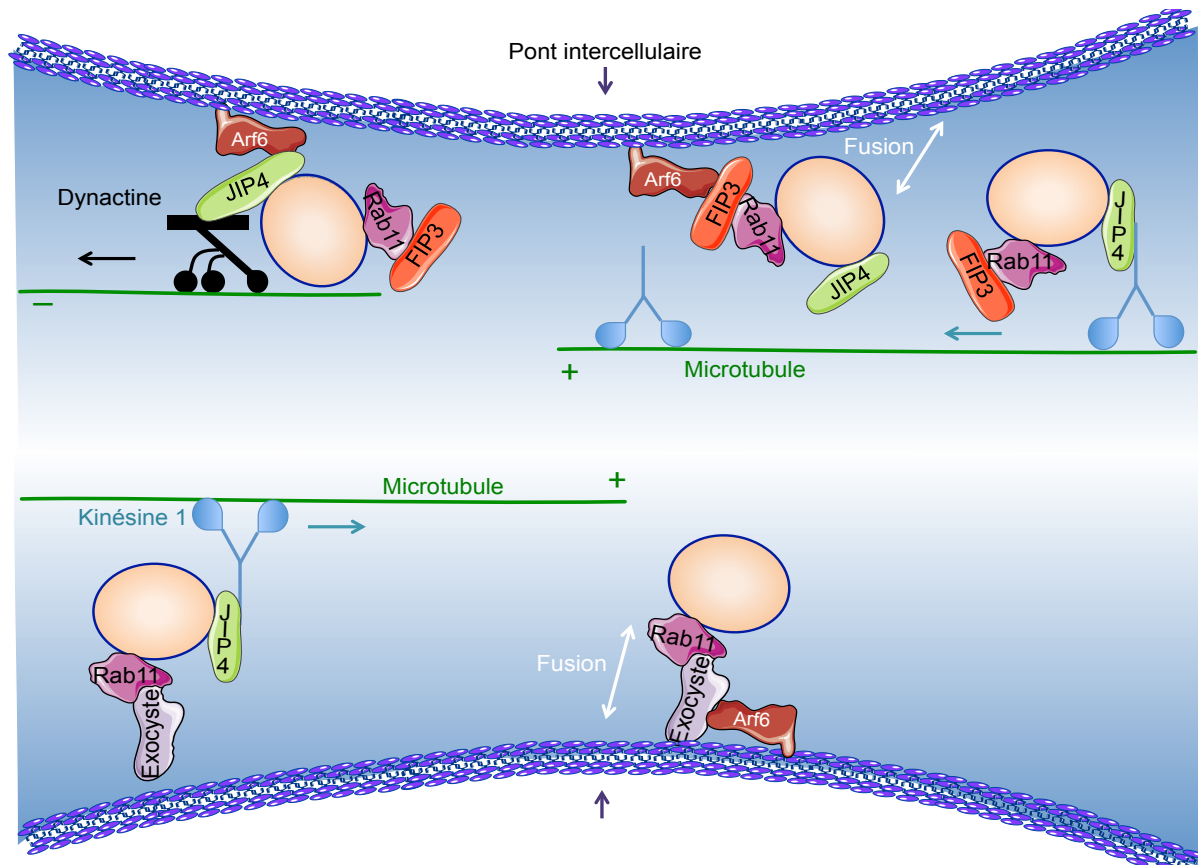


Figure 15: Trafic vésiculaire au pont intercellulaire lors de la cytokinèse. D'après Montagnac et al., 2009, *Curr Biol* [3] et Montagnac et al., 2008, *Curr Opin Cell Biol* [9].

En résumé, le bon fonctionnement du cycle GDP/GTP d'Arf6 est nécessaire au recyclage des vésicules membranaires. Son interaction avec les protéines JIP3/4 est indispensable à ses effets sur les différentes voies de recyclage. Selon la protéine Rab avec laquelle Arf6 est associée et selon la phase dans laquelle se trouve la cellule (interphase ou mitose), Arf6 sera impliquée dans le recyclage rapide (Rab4, Rab5 en interphase), dans le recyclage lent (Rab4, Rab11 en interphase) ou dans la cytokinèse (Rab11 en mitose). Son implication dans le métabolisme des lipides, notamment des phosphoinositides, lui permet de recruter directement ou indirectement les effecteurs nécessaires au recyclage. Une fois les vésicules

transportées à la MP, la liaison d'Arf6 à l'exocyste est indispensable à la fusion des membranes.

2.2. Rôle d'Arf6 dans la migration et l'invasion cellulaires

Aussi bien lors de processus physiologiques (développement, cicatrisation) que pathologiques (cancers), les cellules épithéliales sont amenées à migrer et à envahir d'autres tissus. Pour cela, les cellules réalisent la transition épithélio-mésenchymateuse (TEM) qui consiste en la perte des jonctions adhérentes et de la polarité cellulaire et l'acquisition d'un pouvoir migratoire et invasif [153]. La TEM est essentielle au développement et à la différenciation tissulaire et est un processus réversible. En effet, une fois les cellules arrivées dans leur niche, elles réalisent la transition mésenchymato-épithéliale (TME) afin de s'implanter [153]. Ce processus indispensable à la vie peut être détourné de sa fonction première et être à l'origine du développement de métastases à partir des tumeurs primaires [153].

Les mécanismes moléculaires à l'origine de la TEM pathologique sont nombreux et complexes. Dans la suite de mon manuscrit, je développerai les processus moléculaires qui impliquent la protéine Arf6.

De par ses rôles dans divers mécanismes cellulaires, la protéine Arf6 est un acteur important de la migration et de l'invasion cellulaires. Elle est suractivée dans les cancers du sein [88], les gliomes [154] et les mélanomes [155] et augmentent le pouvoir invasif et métastatique de ces tumeurs.

Comme décrit précédemment, la protéine Arf6 est impliquée dans le recyclage de protéines membranaires (Figures 14 et 15) parmi lesquelles la E-cadhérine, protéine impliquée dans les interactions cellule-cellule par les jonctions adhérentes [156] et

l'intégrine $\beta 1$, protéine impliquée dans les interactions cellule-matrice extracellulaire [157]. Ces deux protéines ont un rôle dans l'acquisition du phénotype invasif des cellules tumorales.

2.2.1. L'endocytose de la E-cadhérine

Les cellules épithéliales sont maintenues cohésives par la présence de différents types de jonctions intercellulaires dont les jonctions adhérentes. Ces jonctions ont pour composant principal la E-cadhérine, protéine membranaire qui fait des liaisons homophiliques Ca^{2+} -dépendantes avec les protéines des cellules adjacentes. Les jonctions adhérentes sont composées de plusieurs autres protéines du côté cytoplasmique telles que la β -caténine qui interagit directement avec la E-cadhérine, l' α -caténine qui interagit avec la β -caténine et le réseau d'actine sous-membranaire [5], la p120-caténine qui interagit avec la E-cadhérine pour stabiliser les jonctions [14] et régulerait l'activité des protéines Rho impliquée dans la polymérisation de l'actine [158] (Figure 16).

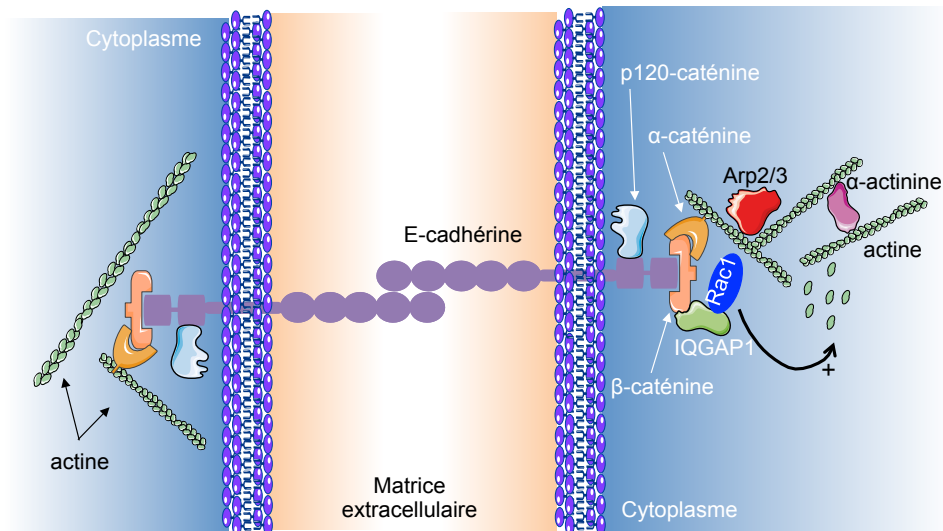


Figure 16: Les jonctions adhérentes. Deux cellules épithéliales adjacentes se contactent via leurs protéines E-cadhérines qui forment des liaisons homotypiques Ca^{2+} -dépendantes. La β -caténine se lie à la queue cytosolique de la E-cadhérine et à l' α -caténine qui fait le lien entre la E-cadhérine et les filaments d'actine. La protéine IQGAP1 interagit avec la β -caténine et la protéine Rac1 qui permet la polymérisation des filaments d'actine. Le branchage du réseau d'actine est permis par les protéines Arp2/3 et sa stabilisation est assurée par l' α -actinine. D'après Conacci-Sorell et al., 2002, J Clin Invest [5] ; Bryant et al., 2004, Trends Cell Biol [14] ; Izumi et al., 2004, J Cell Biol [15].

Lors de la TEM, les jonctions adhérentes sont désassemblées, la E-cadhérine est internalisée, dégradée par le protéasome et réprimée au niveau transcriptionnel notamment par SNAIL [159] et SLUG [160]. En effet, la β -caténine est transférée au noyau où elle régule la transcription de gènes impliqués dans l'invasion notamment les facteurs de répression SNAIL et SLUG [161]. Il a récemment été démontré qu'Arf6 est impliquée dans le recyclage de la E-cadhérine à la MP mais également dans la régulation de sa dégradation par le protéasome [162] (Figure 17). En effet, Arf6-GTP interagit avec la protéine arfaptine-2 qui inhibe le protéasome 26S [163].

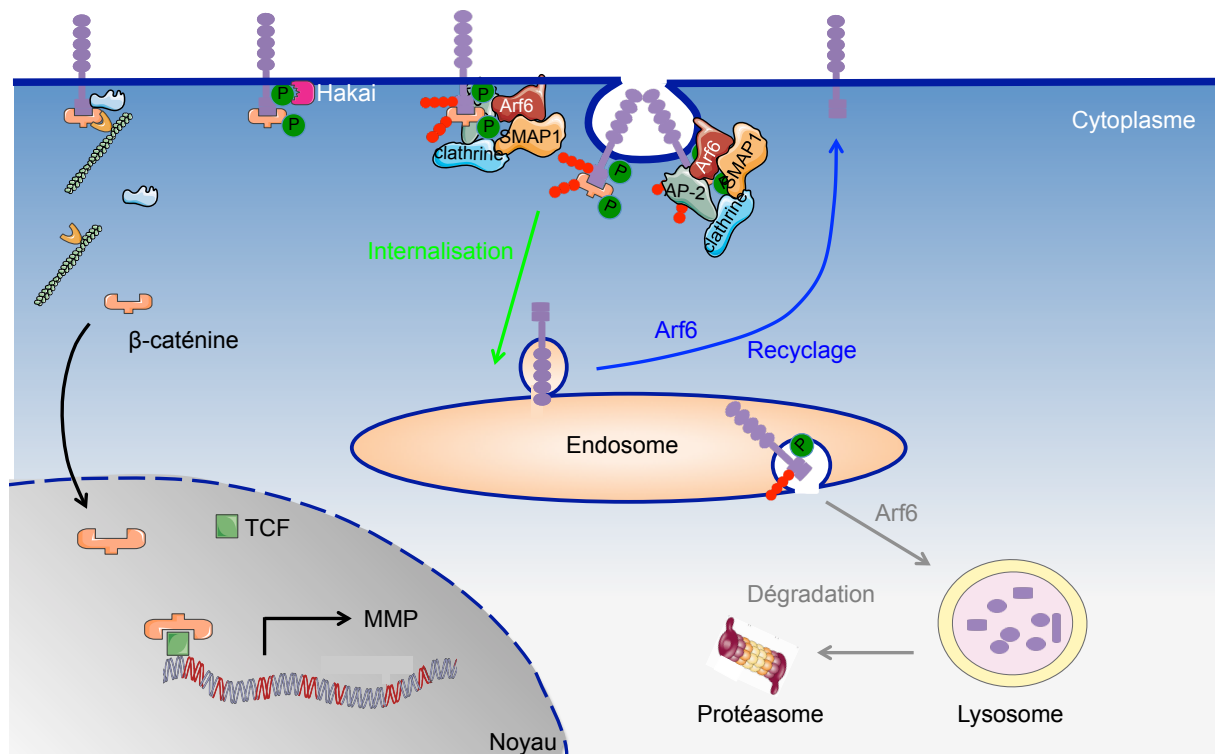


Figure 17: Endocytose clathrine-dépendante de la E-cadhérine lors de la transition épithélio-mésenchymateuse. L'internalisation de la E-cadhérine est contrôlée par Arf6 et SMAP1. La E-cadhérine mono-ubiquitinylée (Ub) est adressée au lysosome et au protéasome *via* Arf6 pour dégradation. La E-cadhérine non Ub est recyclée à la MP. La β-caténine est transloquée au noyau où elle interagit avec TCF4 (T Cell Factor) pour activer la transcription de gènes impliqués dans la TEM. D'après D'Souza-Schorey, 2005, Trends Cell Biol [164] et Figueiredo et al., 2011, PLoS One [162].

Les signaux d'internalisation de la E-cadhérine sont la phosphorylation [165] et l'ubiquitination (Ub) [166] (Figure 17). La phosphorylation de la E-cadhérine et des caténines notamment la β-caténine est contrôlée par les protéines tyrosine kinase de la famille Src et par d'autres récepteurs tyrosine kinase tels que l'EGFR (*Epidermal Growth Factor Receptor*) [165]. La phosphorylation de la E-cadhérine (E-cadhérine-P) par Src a lieu dans le domaine de liaison de la p120-caténine qui est déplacée du complexe E-cadhérine/caténines. La β-caténine également phosphorylée n'est plus capable de lier l'α-caténine ; ainsi, la E-cadhérine n'est plus associée au réseau d'actine. La E-cadhérine-P recrute la protéine E3-Ub ligase Hakai qui ubiquitinye la

E-cadhérine et la β -caténine aboutissant à l'internalisation du complexe [166] (Figure 17).

Le mécanisme moléculaire du recyclage de la E-cadhérine est peu connu à ce jour. La E-cadhérine serait dé-ubiquitinylée au niveau de l'endosome de recyclage et transportée dans des vésicules Arf6- et Rab11-positives vers la MP [162] [167]. Arf6 n'a pas pour seul rôle le recyclage de la E-cadhérine ; en effet, il a été démontré que l'activation d'Arf6 est nécessaire à l'internalisation de la E-cadhérine [86] [14] [168] [87]. L'expression du mutant Arf6(T27N), inactif, inhibe l'internalisation de la E-cadhérine et ce malgré sa phosphorylation par Src, signal activateur de l'internalisation de la E-cadhérine [86]. Ainsi, l'ubiquitination du complexe E-cadhérine/ β -caténine n'est pas suffisante à son internalisation mais est nécessaire à son transport vers le lysosome et donc à sa dégradation [168]. Dans l'étude de Kon et al., l'augmentation d'Arf6-GTP par l'expression du mutant SMAP1 R61Q, dont l'activité GAP est nulle, inhibe également l'internalisation de la E-cadhérine, au même titre que l'augmentation d'Arf6-GDP par la surexpression de SMAP1 [79]. Dans les tumeurs du sein, le niveau d'Arf6-GTP corrèle avec le pouvoir migratoire des cellules, plus Arf6 est actif et plus la cellule est capable de migrer [88]. Cette observation serait due à l'activité d'Arf6 sur le remodelage de l'actine par l'intermédiaire de Rac1, petite GTPase impliquée dans la polymérisation de l'actine, et à l'activité d'Arf6 sur le métabolisme lipidique développé précédemment [89]. En effet, l'interaction d'Arf6-GTP avec Nm23-H1, une nucléoside diphosphate kinase, premier gène suppresseur de métastases identifié [169] [170], résulte en la séquestration de Tiam1, une GEF de Rac1, par Nm23-H1 l'empêchant ainsi d'activer Rac1 [171] (Figure 18A). L'expression de Nm23-H1 permet de stabiliser les jonctions adhérentes en stabilisant la E-cadhérine à la MP [172]. En effet, les cellules présentes au front de migration d'une tumeur perdent l'expression de Nm23-

H1 libérant Tiam1 qui est alors libre d'activer Rac1 permettant la polymérisation de l'actine aboutissant à la formation des lamellipodes et invadopodes, structures membranaires de la migration [173] (Figure 18B). Cet effet est accentué par l'internalisation de la E-cadhérine due à l'augmentation d'Arf6-GTP [87]. Le transport de Rac1 jusqu'à ces structures serait facilité par Arf6 [174]. Rac1 et Arf6 sont colocalisés aux invadopodes et permettent la polymérisation de l'actine et le renouvellement des structures de la migration. Cependant, d'après les travaux de Tanabe et al., la surexpression de SMAP1 n'a aucune conséquence sur la formation de protrusions chimio-induites. Ainsi, les auteurs concluent que SMAP1 ne serait pas la Arf6GAP impliquée dans la régulation du remodelage de l'actine [77].

Afin que les cellules deviennent invasives, c'est-à-dire qu'en plus de migrer elles soient capables de digérer la matrice extracellulaire (MEC), elles doivent exprimer des protéases telles que les métalloprotéases (MMP). Ces protéines, sécrétées ou membranaires, digèrent les différents composants de la MEC, collagène, fibronectine... La translocation nucléaire de la β -caténine suite au désassemblage des jonctions adhérentes permet l'activation transcriptionnelle des gènes codant les MMP2, 7 et 9 (<http://www.stanford.edu/~rnusse/pathways/targets.html>) (Figure 17). Arf6 n'aurait aucun rôle dans la sécrétion de ces protéases [88].

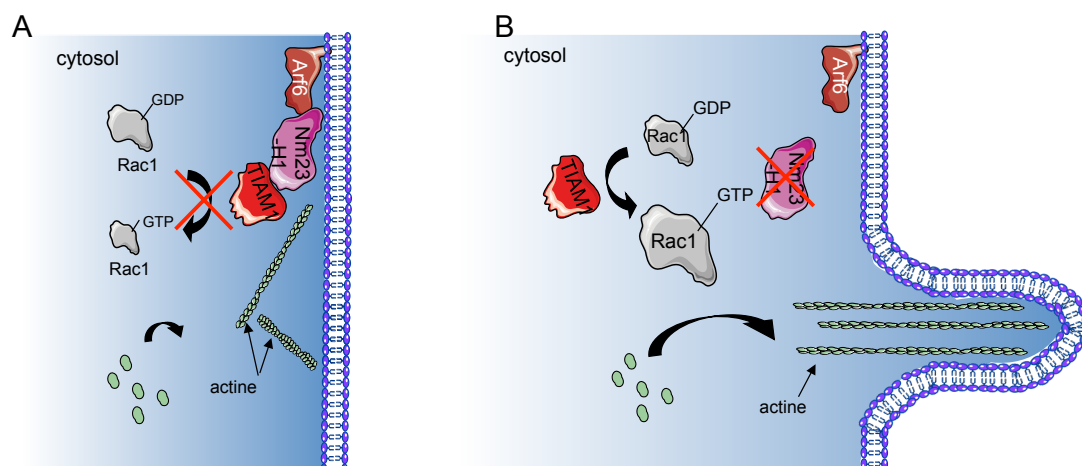


Figure 18: Effet de Nm23-H1 dans l'invasion cellulaire. A. La protéine Nm23-H1 séquestre Tiam1 empêchant l'activation de Rac1 donc la polymérisation de l'actine nécessaire à la formation des invadopodes. B. La perte de l'expression de Nm23-H1 au front de migration libère Tiam1 qui active Rac1 permettant la formation d'invadopodes.

2.2.2. L'endocytose de l'intégrine $\beta 1$

En plus de son action sur les jonctions adhérentes, Arf6 est impliqué dans l'endocytose des intégrines, protéines membranaires récepteurs des composants de la MEC [175] [157]. Les intégrines sont des hétérodimères composés d'une sous-unité α et d'une sous-unité β impliquées dans l'adhésion à la MEC [175]. Les intégrines sont présentes à la MP sous forme active ou inactive. L'activation des intégrines est contrôlée par des signaux extracellulaires complexes, c'est la signalisation « *inside-out* » qui permet le changement conformationnel du complexe α - β d'une conformation fermée, inactive, à une conformation ouverte, active [12]. Une fois sous la forme active, les intégrines peuvent interagir avec les éléments de la MEC aboutissant au rassemblement de plusieurs intégrines et ainsi à la formation

des complexes focaux d'adhésion. Les intégrines rassemblées sont responsables de la signalisation « *outside-in* » [12] (Figure 19).

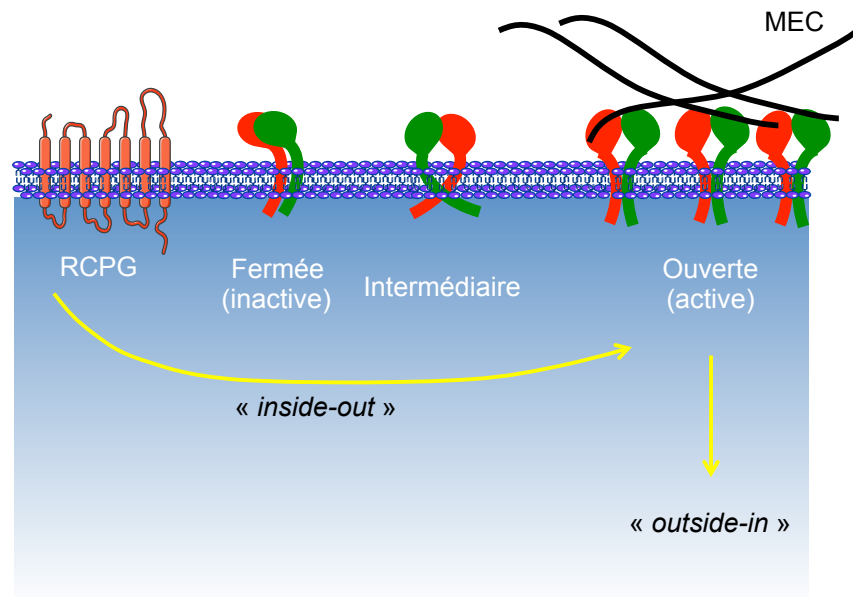


Figure 19: Activation des intégrines. Des signaux extracellulaires permettent le changement conformationnel des intégrines passant d'une conformation fermée à une conformation active qui peut interagir avec les éléments de la MEC permettant le rassemblement des intégrines et donc le renforcement des complexes focaux d'adhésion. D'après Margadant et al., 2011, *Curr Opin Cell Biol* [12]

La sous-unité $\beta 1$ est celle qui interagit avec le plus grand nombre de sous-unités α [175]. Le trafic de cette sous-unité sera particulièrement étudié dans la suite de mon manuscrit.

La sous-unité $\beta 1$ des intégrines, comme la plupart des protéines membranaires, est internalisée puis recyclée à la MP. Durant la migration, son endocytose est polarisée, elle est internalisée à partir de l'arrière de la cellule et est recyclée au front de migration [157]. L'internalisation de $\beta 1$ a lieu de manière constitutive, en absence de signal [175], et est clathrine-dépendante [7]. La sous-unité $\beta 1$ est ensuite transportée jusqu'à l'endosome de recyclage, à partir duquel elle est recyclée au front de migration de la cellule sous l'effet de signaux externes tel que l'EGF. Il a récemment

été montré que l'endocytose de l'intégrine dépend de son état d'activation [7]. En effet, l'internalisation est plus efficace pour la forme active alors que le recyclage de la forme inactive est plus rapide. Le recyclage de la forme inactive est contrôlé par l'endosome de tri *via* Rab4 (recyclage rapide) alors que celui de la forme active est géré par l'endosome de recyclage *via* Arf6 et Rab11 [7] (Figure 20). La forme inactive de l'intégrine $\beta 1$ est recyclée au front de migration, où elle colocalise avec Arf6, pour être activée et former de nouveaux complexes d'adhésion [175]. L'activité d'Arf6 sur le remodelage de l'actine est indispensable au recyclage de l'intégrine $\beta 1$ au front de migration [175] [7]. En effet, l'expression du mutant d'Arf6 (Q37E/S38I) qui inhibe la régulation du réarrangement du réseau d'actine [176], bloque le recyclage de l'intégrine $\beta 1$ au même titre que le traitement par la cytochalasine D, inhibiteur de la polymérisation de l'actine [175].

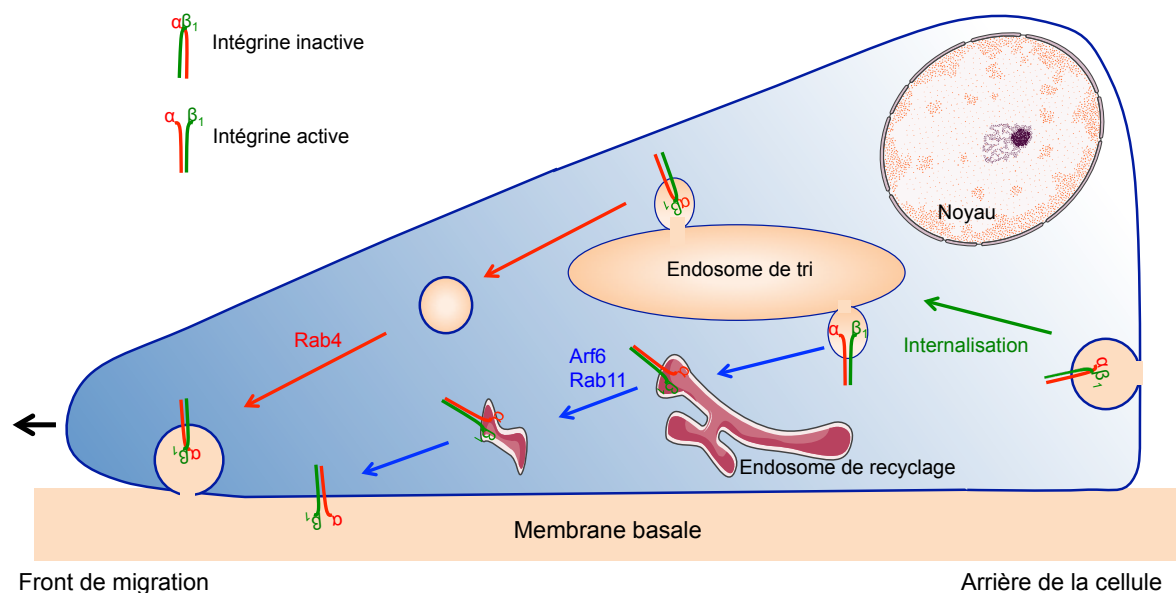


Figure 20: Endocytose de l'intégrine $\beta 1$ au cours de migration cellulaire. L'intégrine $\alpha/\beta 1$ est internalisée à l'arrière de la cellule pour être recyclée au front de migration. La flèche verte représente l'internalisation. Les flèches rouges représentent le recyclage rapide de l'intégrine $\alpha/\beta 1$ inactive. Les flèches bleues représentent le recyclage lent de la forme active. D'après Arjonen, 2012, *Traffic* [7].

En résumé, l'effet d'Arf6 sur la migration cellulaire est en partie dû à son action sur les jonctions adhérentes et sur le recyclage des intégrines. En effet, l'activation d'Arf6 est nécessaire au désassemblage des jonctions adhérentes *via* l'internalisation de la E-cadhérine [87] et l'inactivation de Rac1 [171]. Arf6 est également impliqué dans le recyclage à la MP de la E-cadhérine et dans sa dégradation par le protéasome [162]. Le déséquilibre du cycle GDP/GTP par l'expression de mutants d'Arf6 [86] [14] ou de SMAP1 [79] aboutit à l'inhibition de l'internalisation de la E-cadhérine.

En ce qui concerne l'intégrine $\beta 1$ qui est la sous-unité β qui interagit avec le plus grand nombre de sous-unités α , la régulation de son trafic vésiculaire est nécessaire au bon fonctionnement de la machinerie d'adhésion cellulaire. Arf6 est nécessaire à son recyclage au front de migration des cellules en mouvement, notamment par son activité sur le réseau d'actine [175].

Ainsi, le déséquilibre du cycle GDP/GTP d'Arf6 en faveur de la forme active aboutit à la TEM mais n'est pas suffisante à l'acquisition du pouvoir invasif puisqu'Arf6 n'est pas impliqué dans la sécrétion des MMP, essentielles à la dégradation de la MEC [88].

IV. Objectifs de l'étude

En conclusion de cette revue bibliographique, les fonctions cellulaires d'Arf6 sont multiples. Ainsi, la régulation de son cycle GDP/GTP est essentielle à son fonctionnement normal. À ce jour, plusieurs protéines GAP et GEF ont été décrites comme étant impliquées dans la régulation du cycle d'Arf6. L'une des protéines GAP

la plus récemment identifiée est la protéine SMAP1 qui est spécifique d'Arf6. De manière intéressante, la protéine SMAP1 a été identifiée, dans notre laboratoire, comme très fréquemment mutée dans les CCR MSI.

Les objectifs de ma Thèse ont été :

- la caractérisation des mutations de SMAP1, à savoir SMAP1 est-il un gène cible de l'oncogenèse MSI ? ;
- l'étude de l'impact des mutations de SMAP1 sur ses fonctions cellulaires. En effet, la protéine SMAP1 a été décrite comme étant impliquée dans l'internalisation clathrine-dépendante du TfR et de la E-cadhérine mais ces résultats sont difficiles à expliquer au vu des données de la littérature concernant Arf6 ;
- l'analyse des conséquences des mutations sur le pouvoir invasif des cellules tumorales, les tumeurs MSI étant faiblement métastatiques à distance.

RÉSULTATS

PARTIE II. RÉSULTATS

I. Article

Small ArfGAP 1 (SMAP1), une nouvelle protéine activatrice de GTPase spécifique d'Arf6 impliquée dans l'oncogenèse MSI

Les cancers colorectaux (CCR) instables sur les microsatellites (MSI) représentent 15% des CCR. Ils résultent de l'inactivation du système de réparation des mésappariements de bases (MMR) impliqué dans la réparation des erreurs de réplication commises par l'ADN polymérase. En absence de système MMR, la cellule accumule des mutations en particulier au niveau de séquences nucléotidiques répétées, les microsatellites. De par leur nature, ces séquences sont des points chauds de mutations par insertion/délétion qui conduisent à un décalage du cadre de lecture et à la synthèse d'une protéine tronquée. L'enjeu de l'étude des tumeurs MSI est de distinguer les mutations dites « bruit de fond », dues à la déficience du système MMR, des mutations sélectionnées par les cellules tumorales pour leur développement. Très peu de ces gènes ont fait l'objet d'une étude fonctionnelle.

Au sein du laboratoire, un nouveau gène cible a été identifié, le gène SMAP1 (*Small ArfGAP1*) muté dans 73% des CCR MSI. Ce gène code une protéine de la famille ArfGAP spécifique d'Arf6, une petite protéine G impliquée dans de nombreux mécanismes cellulaires tels l'endocytose clathrine-dépendante et le cycle cellulaire. La protéine SMAP1 catalyse l'hydrolyse du GTP d'Arf6 en GDP et a été décrite comme étant impliquée dans la régulation d'Arf6 lors de l'endocytose clathrine-dépendante du récepteur de la transferrine (TfR).

Dans cette étude, nous avons caractérisé les mutations de SMAP1 dans les tumeurs MSI et leurs conséquences sur son expression. Nous avons montré par séquençage que les mutations de SMAP1 sont spécifiques des lignées cellulaires dérivées de tumeurs MSI de diverses origines tissulaires et ne surviennent que dans le microsatellite de 10 répétitions adénine (A10). De plus, l'expression du transcrit de SMAP1 est diminuée dans les CCR MSI aussi bien que dans les lignées MSI comparées aux lignées non-MSI dû à la dégradation de l'ARNm par le système NMD (*Non-sense mediated mRNA Decay*). Au niveau protéique, aucune bande n'est détectable en western blot dû à la dégradation de la protéine tronquée par le protéasome. Ainsi, les mutations de SMAP1 sont de type perte-de-fonction. La fréquence de mutations de SMAP1 diminue en fonction du stade de la tumeur ; plus la tumeur est invasive plus la fréquence de mutations est faible. Quelque soit le statut mutationnel de SMAP1, la réponse des patients à la chimiothérapie est identique.

Les protéines SMAP1 et Arf6 étant impliquées dans l'endocytose clathrine-dépendante, nous avons analysé les conséquences fonctionnelles des mutations de SMAP1 sur l'endocytose du TfR. Pour cela, nous avons dérivé des clones à partir de la lignée HCT116 mutée pour SMAP1 ré-exprimant la forme sauvage de SMAP1. La restauration de l'expression de SMAP1 corrige le défaut du recyclage rapide du TfR observé dans les cellules HCT116 mais n'a pas de conséquence sur son internalisation.

La protéine Arf6 est également impliquée dans l'étape de cytokinèse, dernière phase du cycle cellulaire. L'une des caractéristiques cliniques des tumeurs MSI est leur volume plus grand que celui des tumeurs non-MSI. Ainsi, nous avons analysé l'effet des mutations de SMAP1 sur le cycle cellulaire. La complémentation de SMAP1

dans les cellules HCT116 ralentit la phase G2M du cycle cellulaire et par conséquent diminue la vitesse de prolifération des cellules et leur capacité clonogénique.

En conclusion, nos résultats montrent que le gène SMAP1 est un gène cible de l'oncogenèse MSI et dont les mutations résultent en une perte-de-fonction. D'un point de vue clinique, l'absence de SMAP1 semble avoir un rôle protecteur et d'un point de vue moléculaire, l'absence de SMAP1 retarde le recyclage rapide du TfR et accélère la vitesse de prolifération cellulaire, en concordance avec l'une des caractéristiques cliniques des tumeurs MSI qu'est leur taille volumineuse. La protéine SMAP1 étant impliquée dans la régulation du cycle GTP/GDP d'Arf6, son absence affecterait les fonctions cellulaires d'Arf6, suggérant que SMAP1 participerait à une nouvelle voie d'oncogenèse caractérisée par des altérations du trafic membranaire.

**Small Arf-GAP1 (SMAP1), a novel Arf6-specific GTPase activating protein
involved in MSI oncogenesis**

Fatiha Sangar, Ann-Sofie Schreurs, Oriane Mauger, Aziz Zaanani, Catherine Miquel,
Jean-François Fléjou and Françoise Praz

Affiliations of authors: INSERM, UMR_S 938, Saint-Antoine Research Center, F-75012, Paris, France (FS, ASS, OM, AZ, JFF, FP); UPMC Univ Paris 06, UMR_S 938, Saint-Antoine Research Center, F-75012, Paris (FS, ASS, OM, AZ, JFF, FP); Department of Pathology, Saint-Antoine Hospital, AP-HP, F-75012, Paris, France (JFF); Department of Pathology, Sainte-Anne Hospital, University Paris Descartes, F-75014, Paris, France (CM).

Correspondence to: Françoise Praz, PhD, (e-mail: francoise.praz@upmc.fr), Centre de Recherche Saint-Antoine, INSERM-UPMC UMR_S 938, Bâtiment Kourilsky, 184 rue du Faubourg Saint-Antoine, 75571 Paris cedex 12, France.

Abstract

Background. *Small Arf-GAP1* (SMAP1), a GTPase-activating protein specific for Arf6, a small GTPase acting on membrane trafficking and actin remodeling, is frequently mutated in colorectal cancers (CRC) displaying microsatellite instability (MSI).

Methods. *SMAP1* mutations have been characterized (genotyping, sequencing) in 89 MSI CRCs, including 41 stage II, 40 stage III and 8 stage IV, and 46 cell lines of which 25 displayed MSI. *SMAP1* expression levels have been measured at the mRNA and protein levels (qRT-PCR, western blots). Endocytosis of transferrin receptor (TfR) (FACS, immunofluorescence), proliferation and cell cycle analysis (FACS) have been performed on the *SMAP1*-null HCT116 cells and three isogenic clones stably re-expressing wild-type *SMAP1*.

Results. *SMAP1* mutation frequency was inversely correlated with disease stage ($P = .03$), but had no effect on disease-free survival of patients with stage III CRC receiving fluorouracil-based adjuvant treatment. Mutations occurred in MSI tumors of various tissue origins, consisted exclusively in indel in the 10-adenine repeat, and resulted in a significant decrease in *SMAP1* transcript levels in cell lines derived from MSI CRCs ($P = .006$). Truncated *SMAP1* protein was sensitive to proteasome degradation, and undetectable in cells with *SMAP1* mutations. *SMAP1* inactivation had no effect on TfR internalization but slowed down TfR recycling to the membrane ($P = .020$). Consistent with the characteristic clinical features of MSI tumors, *SMAP1* mutations increased cell proliferation ($P = .008$) by shortening the G2/M phase.

Conclusion. Our observations suggest that *SMAP1* inactivation in MSI CRC contributes to the emerging oncogenic pathway involving abnormal Arf6 regulation.

Introduction

Colorectal cancers (CRCs) displaying microsatellite instability (MSI) represent 15% of all CRCs, the second leading cause of death from cancer in Western countries. MSI results from the inactivation of the mismatch repair (MMR) system, which is dedicated to post-replicative correction of polymerase errors arising during replication (1). In absence of an efficient MMR system, tumors accumulate numerous mutations especially in microsatellites because these simple sequence repeats are hotspots for indel mutations during replication (2, 3). Coding microsatellites are most often mononucleotide repeats, in which indel mutations cause a translation frameshift and a premature termination codon (PTC) (4). Abnormal transcripts containing a PTC are generally recognized and degraded by the nonsense-mediated mRNA decay (NMD) system, but mRNA decay may be partially efficient, leading to the synthesis of a truncated protein containing a C-ter neopeptide (5). Among the numerous mutations arising in MMR-deficient tumors, driver mutations are positively selected during MSI-driven oncogenesis, defining the so-called real target genes (2, 3). According to a statistical model based on regression analysis of their mutation frequency in regard to the repeat tract length (3), genes carrying an (A)₁₀ repeat may be considered as real target genes if their mutation frequency exceeds 56% in MSI CRCs. This is notably the case of the transforming growth factor β receptor 2 gene whose coding (A)₁₀ repeat is mutated in approximately 80% MSI CRC and leads to loss of proliferation control (6).

We have recently identified several novel putative MSI-driven target genes including small Arf-GAP1 (small ADP ribosylation factor-GTPase activating protein, known as *SMAP1*) (7). In the series of 44 primary MSI CRCs initially studied, *SMAP1* coding

(A)10 repeat displayed mutations in 73% cases, a mutation frequency among the highest in MSI CRCs and far above the threshold defined for real target genes (3, 7).

SMAP1 is an ArfGAP that acts on Arf6 (8), a small GTPase involved various cellular functions, including clathrin-dependent endocytosis, actin remodeling and cytokinesis (9). The *SMAP1* gene located on chromosome 6q13 contains 11 exons, with an alternatively spliced exon 5 and a coding (A)10 repeat in exon 6 (10). The longer isoform of *SMAP1* encodes a 467-amino acid protein composed of an ArfGAP domain, a C4 type zinc finger, a clathrin binding domain (CB) and a CALM (Clathrin Assembly Lymphoid Myeloid Leukemia) protein binding domain (Figure 1). The shorter *SMAP1* transcript devoid of exon 5 encodes a 440-aa long protein that lacks 27 aa downstream the Arf-GAP domain with yet unknown specific function assigned. The murine *Smep1* gene homolog, originally named Stromal-Membrane Associated Protein 1, was initially identified as playing a role in stromal cells to stimulate erythropoiesis; it shares 93% identity at the protein level with its human homolog and a conserved coding (A)10 repeat (11). *SMAP1* shows a high degree of homology with *SMAP2*, a recently identified *SMAP* gene family member acting as an Arf1-specific ArfGAP in the clathrin-dependent early endosome to trans-Golgi network traffic (12). *SMAP2* gene harbors a coding (A)8 repeat, not conserved in mice. At the protein level, SMAP1 and SMAP2 share 63% overall similarities, with their Arf-GAP domains exhibiting 90% aa identity (Figure 1). SMAP1 binds directly to the clathrin heavy chain through its clathrin box sequence and regulates clathrin-dependent endocytosis of transferrin (Tf) receptor (TfR) (8) and E-cadherin (13).

In this study, we report that *SMAP1* mutations arise frequently in human MSI tumors of various tissue origins, both in solid tumors and in lymphomas. In MSI CRCs, *SMAP1* mutation frequency was inversely related to tumor TNM stage, being highest in stage II CRCs that are devoid of local or distant metastasis, suggesting

that *SMAP1* mutations may counteract the tumor metastatic potential. Using various cell lines, we have shown that MSI-driven *SMAP1* frameshift mutations lead to decreased levels of abnormal transcripts and to subsequent degradation of residual truncated protein by the proteasome, resulting in loss-of-function. *SMAP1* inactivation alters several important cellular functions, such as TfR recycling, cell proliferation and clonogenic capacities, as well as cell cycle.

Materials & Methods

Patients, treatment and follow up

This study included 89 patients who underwent surgical resection of histologically proven MSI CRC mainly at the Saint-Antoine hospital or at the Institute Mutualiste Montsouris (Paris, France) between March 1994 and October 2007, and for whom clinical data and tumor tissue could be retrieved. Tumor type and grade of differentiation were determined according to the World Health Organization criteria. According to the TNM (Tumor-Node-Metastasis) classification; there were 41 stage II, 40 stage III and 8 stage IV cases. Among patients with stage III CRC, 33 received adjuvant chemotherapy consisting in 5-fluorouracil and leucovorin (FL) either alone (n = 20) or in combination with oxaliplatin (FOLFOX, n = 13); 3 patients did not receive adjuvant chemotherapy and the information was not available for 4 patients. Chemotherapy started at the Saint-Antoine Hospital within seven weeks after surgery, where follow-up was further conducted, as previously reported (14). Patient follow-up was defined as the time between surgery and the last hospital contact or disease recurrence.

Immunohistochemical analysis.

Blocks of formalin-fixed paraffin-embedded CRC tissue comprising an area of normal colonic mucosa adjacent to the tumor were selected in each case. Sections of 4 µm from the paraffin-embedded tissue samples containing normal and tumor areas were deparaffinized in xylene and rehydrated in pure ethanol, endogenous peroxidase was blocked using 3% hydrogen peroxide in methanol for 30 min. Antigen retrieval was performed in citrate buffer (pH 6.0). Sections were then incubated for 2

hours at room temperature with antibody to SMAP1 (dilution 1/75, HPA030574, Atlas Antibodies), and revealed using the Biogenex Detection kit (Biogenex). SMAP1 expression was considered negative when there was a complete absence of cytoplasmic staining of neoplastic cells, provided that there were adjacent lymphocytes, normal epithelial cells, endothelial cells or fibroblasts showing positive staining.

Cell line cultures, proliferation and clonogenic assays

HeLa and HCT116 cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing stable glutamine (PAA Laboratories GmbH) supplemented with 10% heat-inactivated fetal calf serum (FCS), penicillin (100 µg/mL), and streptomycin (100 units/mL) at 37°C, in a 5 % CO₂ atmosphere. Cell lines were authenticated using 5 highly polymorphic dinucleotide (CA)_n repeats, which were amplified in a single PCR reaction using a fluorescent primer for each amplicon (D2S123, D5S107, D17S1791, D17S250 and D18S1127). The PCR products were denatured and migrated using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems), as described for genotyping; genetic profiles were compared to the reference panels established for the HeLa and HCT116 cell lines. For cell proliferation assays, cells were seeded in 6-well plates at 12000 cells per well and grown in complete medium during 24h or 48h after seeding. Then, cells were harvested and counted using the Scepter counter (Millipore). For clonogenic assays, cells were seeded at 50 cells per well in 24-well plates, grown during 6 days at 37°C. Cells were then fixed in ice-cold 100% methanol during 10min at -20°C and stained with 1% toluidine blue in 1% borax. Clones were counted using the G:BOX chemiluminescence image analyzer and the GeneTools software (Syngene, Ozyme).

Cell cycle analysis

Cells were seeded in 6-well plates at 3×10^5 cells per well in complete medium. After 48 hours, cells were treated with nocodazole at 50ng/ml during 16h, then washed and incubated in complete medium at 37°C. Cells were harvested and counted at various time points; 10^6 cells were fixed in 2mL ice-cold ethanol 70%, permeabilized with 1% NP40, treated with 1mg/mL RNase A and labeled with 50µg/mL propidium iodide. Cells were then analyzed by flow cytometry using LSRII cytometer (Becton Dickinson). Cell cycle data were calculated using ModFit software.

DNA extraction

Genomic DNA was extracted from frozen tumors or cultured cell lines using the NucleoSpin Tissue kit (Macherey-Nagel). When frozen tumor samples were not available, DNA was extracted from paraffin-embedded tumor tissue sections using the NucleoSpin FFPE DNA kit (Macherey-Nagel), after xylene deparaffinization.

***SMAP1* and *SMAP2* genotyping**

Indel mutations occurring in the *SMAP1* (A)₁₀ and *SMAP2* (A)₈ repeats were screened by genotyping genomic DNA using one fluorescent primer (sequences are available as supplementary data). Adequate dilutions of the fluorescent PCR products were mixed with formamide and GeneScan 400HD ROX-labeled size standards (Applied Biosystems), heat-denatured and run on 36 cm capillaries containing GS Performance Optimized Polymer 7, on an ABI 3100 Genetic Analyzer and analyzed using the GeneMapper 3.7 software (Applied Biosystems).

***SMAP1* sequencing**

The PCR products corresponding to each exon and the flanking sequences (see Supplementary Table for primer sequences) were purified using a Multiscreen-PCR filtration plate (Millipore). The sequencing reaction was performed with the Big Dye Terminator v1.1 kit (Applied Biosystems) and one of the PCR primers. A second step of purification was performed in a Multiscreen-DV plate (Millipore) containing fine G50 Sephadex resin (Amersham-Biosciences, GE Healthcare Europe GmbH). The sequencing reaction products were run on an ABI PRISM 3100 Genetic Analyzer, and analyzed using the Sequencing Analysis version 5.2 software (Applied Biosystems).

RNA extraction and reverse transcription

RNA extraction was performed on cells homogenized with the Mixer Mill MM300 (Qiagen) using the RNeasy Mini kit with an additional DNA digestion step by DNase I before the RNA elution step (Qiagen). Reverse transcription into cDNA was obtained by incubating RNA during 10 min at 25°C and 2 h at 37°C with the High Capacity cDNA RT kit utilizing random primers (Applied Biosystems).

Analysis of *SMAP1* transcript expression by quantitative RT-PCR (RT-qPCR)

SMAP1 transcript expression was determined with the *SMAP1* Assay on demand gene expression (Hs 00222567 assay) and a TaqMan endogenous control, and was further normalized using 18S as a reference (Hs 99999901 assay). Real-time PCR reactions were performed using the ABI Prism 7900HT Sequence Detection System (PE Applied Biosystems) and the TaqMan PCR master mix (Applied Biosystems). Amplification plots and predicted threshold cycle (Ct) values were obtained with the Sequence Detector Software (SDS 2.1, Applied Biosystems). For each cell line, the ratios between *SMAP1* and 18S were calculated as the mean

of the ratios of independent triplicate experiments and expressed relative to the HeLa cell line used as a reference.

Constructions of the wild-type and A9 mutated SMAP1 HA-tagged expressing vectors

We cloned in pcDNA3.1-Hygro⁺ expression vector (Life Technologies), the SMAP1 cDNA originated from pENTR221-SMAP1 vector (clone IOH27289, Gateway, Life Technologies) after performing site-directed mutagenesis to correct the C1181A mutation, corresponding to Pro394His substitution. The SMAP1-A9 mutated form of SMAP1, corresponding to a 1bp deletion in the (A)₁₀ repeat, was obtained by site-directed mutagenesis using the 5'-GACAAAATAAATTGGAGAAAGAAAAGGAAAAAAAAAAGGAAGAGAAAAGAGAGAAAAG-3' oligonucleotide and its complementary sequence, as primers. In order to add the HA tag sequence at the 3' end, the SMAP1 cDNA fragments were amplified by PCR. The wild-type SMAP1-HA and mutant SMAP1-A9-HA cDNAs were then digested by *NheI* and *Bam*HI Fast-Digest restriction enzymes (Fermentas), cloned in the pcDNA3.1-Hygro⁺ expression vector, and used to transform Library Efficiency DB3.1 competent bacteria (Life Technologies). Plasmid DNA was extracted from amplified positive colonies using the NucleoBond Xtra Maxi kit (Macherey-Nagel) and further analyzed first by genotyping, then by sequencing, as previously described.

Transfection of HCT116 and HeLa cell lines with wild-type SMAP1 and SMAP1-A9 mutant HA-tagged expression vectors

Exponentially growing HeLa cells were transfected with pcDNA3.1-SMAP1-A9-HA and Lipofectamine 2000 (Life Technologies) according to manufacturer instructions. The HCT116 cell line that carries a biallelic SMAP1-A9 mutation was

used to establish clones that stably re-expressed a SMAP1^{wt} allele. Cells were transfected with the pcDNA3.1-SMAP1-HA expression vectors containing the (A)10 wild-type *SMAP1* cDNA tagged with a C-terminal HA epitope, using the Nucleofector kit (Amaxa, Lonza). Cells were plated at 10⁴, 3000, 1000 or 300 cells per well in flat-bottomed 96 well plates. Four days later, cell cultures were put under selection with hygromycin B (100 µg/mL). Hygromycin B-resistant colonies derived from HCT116 cells were screened for the expression of SMAP1-HA by western blotting, as described below.

Proteasome inhibition

To inhibit proteasome, 5 µM MG132 (Sigma-Aldrich) was added to HCT116 cells immediately after transient transfection with pcDNA3.1-SMAP1-HA or pcDNA3.1-SMAP1-A9-HA expression vector. After 8h at 37°C, cells incubated with or without MG132 were lysed to perform western blot analysis, as described below.

Western blot

Total protein extracts were obtained by lysing cells in Laemmli sample buffer (Bio-Rad). Protein extracts were sonicated, denatured at 95°C for 5 min and quantified using the Qubit fluorometer (Life Technologies). Lysates (25 µg) were separated on NuPAGE Novex 4-12% Bis-Tris Mini Gels (Life Technologies). Proteins transfer on 0.2 µm nitrocellulose membranes (Bio-Rad) was performed with the Trans-Blot Semi-Dry Electrophoretic Transfer Cell (Bio-Rad). Membranes were saturated with Phosphate Buffered Saline (PBS) containing 0.1% v/v Tween 20 (PBST) and in 5% w/v non-fat dry milk (PBSTM), before incubation with the purified primary MaxPab mouse polyclonal antibody directed against SMAP1 protein (H00060682-B01P, Abnova) or with a rat monoclonal anti-HA tag conjugated with

horseradish peroxidase (HRP) (3F10, Roche Diagnostics) at 1/1000 dilution in PBST. When necessary, the membranes were incubated with a secondary goat anti-mouse antibody coupled with HRP (Molecular Probes, Invitrogen) diluted at 1/1000 in PBSTM. Loading controls were performed with a mouse monoclonal anti-GAPDH conjugated with HRP at a 1/5000 dilution in PBSTM (ab9482, Abcam). Membrane signals were revealed with the SuperSignal West Femto Maximum Sensitivity Substrate (Pierce, Thermo Fisher Scientific), detected and quantified using the G:BOX chemiluminescence image analyzer and the GeneTools software.

Transferrin internalization assay

Confluent cells were cultured in serum-free DMEM during 30 min at 37°C for transferrin (Tf) starvation. Cells were then harvested in PBS-1% EDTA, centrifuged, and incubated in DMEM-1% BSA-20mM Hepes supplemented with 5 µg/mL A488- or A633-labeled Tf (Life Technologies, France) during 1h at 4°C. Cells were then washed in ice-cold DMEM-1% BSA-20mM Hepes and further incubated in pre-warmed DMEM-1% BSA-20mM Hepes at 37°C. Cell suspension aliquots taken at various time points, were transferred in ice-cold PBS. For quantitative analyses, cells were washed in an acidic medium to eliminate membrane-bound Tf and resuspended in PBS to be analyzed by flow cytometry (LSR II, BD Biosciences, France). The results are expressed as the percentage of the ratio between the internal fluorescence measured at various time points and the surface-bound fluorescence at the beginning of the experiment. For immunofluorescence analysis, cells were fixed in paraformaldehyde 4% during 15 min at room temperature and observed using an immunofluorescence microscope (Leica).

Transferrin recycling assay

Confluent cells were starved from Tf for 30 min at 37°C in serum-free DMEM, harvested and further incubated in pre-warmed DMEM-0.1% BSA supplemented with 5 µg/mL A633-Tf during 30 min at 37°C to allow Tf internalization. Cells were washed in ice-cold DMEM-0.1% BSA, centrifuged and resuspended in pre-warmed DMEM-10% FCS in absence of Tf. At different time points, an aliquot of cell suspension was taken, put in ice-cold PBS-0.1% BSA and further kept on ice. At the end of the kinetics, cells were washed in an acidic medium, centrifuged and resuspended in ice-cold PBS. Samples were analyzed by flow cytometry and results are expressed as the percentage of internal fluorescence at various time points relative to the fluorescence measured at the end of the internalization process.

Statistical analyses

Differences in distributions between the variables examined were assessed with the Student's t-test (age distribution), the Chi-square (tumor location, stage) or the Fisher's exact test (patients under 65-yr, sex, differentiation), as appropriate. The primary outcome was DFS, defined as the time from the date of surgery to the date of relapse or last contact; DFS was considered as censored when the patient was alive without relapse at last contact. DFS curves for the group of stage III patients receiving FL were compared according to the tumor SMAP1 mutation status. Survival curves were estimated by the method of Kaplan and Meier; the hazard ratios (HRs) and 95% confidence intervals (CIs) for recurrence were compared using a two-sided log-rank test (Prism; GraphPad, San Diego, CA).

To compare transcript expression levels in primary tumors and cell lines, we performed a one-tailed Mann-Whitney test. For the analysis of *SMAP1* mutations frequency according to the TNM stage of the tumor, we used a Chi-square test for trend comparing all data and a two-tailed t test for pairwise comparisons. For cell

proliferation assay, we used a one-way ANOVA, Dunnett's test comparing the three clones to the parental cell line defined as the control. For clonogenic assay, we performed a two-tailed t test. For TfR internalization and recycling, we used the Two-way ANOVA test (GraphPad PRISM).

Results

***SMAP1* mutation frequency in primary MSI CRCs.**

Based on our previous data showing that *SMAP1* is frequently mutated in primary MSI CRCs (7), we determined *SMAP1* mutation frequency according to TNM stage in a series of 89 MSI CRCs including stage II to stage IV cases (Figure 2D). *SMAP1* mutation frequency was highest among the 41 stage II MSI CRCs analyzed of which 28 (68%) carried *SMAP1* mutations, decreasing to 57% for stage III MSI CRCs (23/40) and to 25% for stage IV MSI CRCs (2/8), showing that disease extension and *SMAP1* mutation frequency were significantly inversely correlated ($P = .03$, Chi-square test for trend). In the group of stage III MSI CRCs, the clinical features of patients whose tumor harbored *SMAP1* mutations did not significantly differ from those of patients having a tumor with wild-type *SMAP1* (Table 1). Indeed, there were no difference in the age and sex of patients ($P = .24$, t-test and $P = .53$, Fisher's exact test, respectively), on tumor location ($P = .62$, Chi2 Yates correction), on differentiation grade ($P = .32$, Fisher's exact test) and stage ($P = .86$, Chi2 test). As shown by immunohistochemistry, *SMAP1* is expressed in the cytoplasm of normal colonic mucosa (Figure 2A). Immunohistochemical analyses revealed that among tumors in which *SMAP1* mutations have been detected at the molecular level, some showed an heterogeneous *SMAP1* expression pattern, with some areas of the tumor displaying normal expression, while others were negative (Figure 2B), and some showed a homogeneous negative pattern (Figure 2C). This indicates that *SMAP1* mutations are not necessarily arising at an early stage of tumor development.

Effect of *SMAP1* mutations on prognosis or predictive value in stage III CRCs.

We further investigated whether the presence of *SMAP1* mutations may have an impact on the outcome of patients with stage III MSI CRC treated by adjuvant chemotherapy (n = 33) (Figures 2E and 2F). As expected DFS for the group of patients treated with 5FU (n = 21) was significantly shorter compared to those who were given FOLFOX (n = 12) (hazard ratio [HR] 5.55; 95% confidence interval 1.3–25.0; $P = .026$) (Figure 2E). Thus, the prognostic value of *SMAP1* mutations could not be investigated in this series of patients without taking into account the chemotherapy regimen. Moreover, the predictive value of *SMAP1* mutations could be defined only for patients treated with 5FU-based adjuvant chemotherapy since none of those receiving FOLFOX relapsed (Figure 2E and 2F). In our series of patients, the status of *SMAP1* mutations had no significant impact on DFS of patients receiving 5FU (HR 2.54; 95% CI 0.5–13.4; $P = .27$) (Figure 2F).

Characterization of *SMAP1* and *SMAP2* mutations in cancer cell lines.

Indel mutations arising in the (A)₁₀ and (A)₈ coding repeats located in *SMAP1* and *SMAP2* genes, respectively, were screened in 46 cell lines established from 25 MSI and 21 MSS (MicroSatellite Stable) tumors from various tissues (Figures 3A and 3B). Cell lines established from MSI CRCs displayed the highest mutation frequency, with 11 of 14 cell lines (79%) having *SMAP1* mutations, most often biallelic (8 cell lines) (Figure 3B). *SMAP1* (A)₁₀ repeat was also frequently mutated in cell lines derived from gastric and endometrial MSI carcinomas, with 3 of 4 cell lines displaying biallelic mutations (Figures 3A and 3B). Three of the 7 lymphomas displaying MSI analyzed (43%) harbored *SMAP1* mutations in the coding (A)₁₀ repeat, but only monoallelic mutations could be detected in this panel of cell lines (Figures 3A and 3B). Deletions were more frequent than insertion, with only 3 cell lines displaying an (A)₁₁ repeat corresponding to a 1 bp insertion, whereas 25 mutated alleles had lost

1 bp (24 cases) or 2 bp (1 case). None of the cell lines established from MSS tumors displayed mutations in the *SMAP1* (A)₁₀ repeat (Figure 3A). By contrast, no mutation could be found in the coding A8 repeat of *SMAP2* in any of the 46 cell lines analyzed (data not shown).

To investigate whether *SMAP1* mutations were restricted to the (A)₁₀ repeat or could occur elsewhere, we sequenced the entire coding exons and the exon/intron junctions in all human MSS or MSI cell lines devoid of biallelic mutations in the (A)₁₀ repeat. Four cell lines, two MSS (HGT1, MKN28) and two MSI (SNU1, TC71), carried a previously described polymorphism in exon 7 (c.635C>T, rs2273566), resulting in an Ala to Val change at position 185 or 212 depending on the *SMAP1* isoform (Figure 3A). We found another single nucleotide polymorphism located in the (A)₁₀ repeat (c.515A>G, rs138685504), which replaces Lys by Arg at codon 145 or 172, in three cell lines, one MSS (MKN28) and two MSI (LIM2405 and RKO) (Figure 3A). All other *SMAP1* changes were out-of-frame mutations in the (A)₁₀ repeat confirming that they occurred only in MSI cell lines; yet, no *SMAP1* mutation could be found in any of the 21 MSS cell lines analyzed, regardless of the tissue from which they originated.

SMAP1 mRNA levels have been quantified relative to 18S by RT-qPCR in a collection of 13 MSI and 8 MSS cell lines and expressed as a ratio using HeLa cells as a reference. As shown in Figure 3C, cell lines lacking *SMAP1* mutations expressed *SMAP1* mRNA at similar levels regardless of the MSI status, with mean levels of 128 ± 15 vs 105 ± 19 (mean $\pm$ sem) for MSS and MSI cell lines, respectively ($P = .50$, Mann-Whitney two-tailed test). As expected, *SMAP1* transcript levels were significantly lower in MSI cell lines harboring *SMAP1* mutations (40 ± 10) compared to MSS and MSI cell lines lacking *SMAP1* mutation ($P = .006$ and $P = .02$, respectively, Mann-Whitney one-tailed test); the decrease in *SMAP1* transcript levels

was more pronounced for MSI cell lines displaying biallelic mutations (31 ± 13 ; $P < .001$ and $P = .02$, compared to MSS and MSI cell lines) than for those with monoallelic mutations (mean 63; $P = .02$ and $P = .10$, compared to MSS and MSI cell lines, Mann-Whitney one-tailed test) (Figure 3C). However, *SMAP1* transcripts levels in LoVo and SW48 were comparable to cell lines devoid of *SMAP1* mutations, despite these two cell lines carried biallelic *SMAP1* mutations, suggesting that the NMD system was inefficient in these cells (Figure 3C). Thus, although aberrant *SMAP1* transcripts were expressed at lower levels than wild-type transcripts, they remained detectable and possibly translated into a truncated SMAP1 protein.

Consequences of *SMAP1* gene mutations on SMAP1 protein expression.

In order to determine if mutated transcript was translated into a truncated protein that could act as a dominant negative, we used a polyclonal antibody recognizing amino acids 101-115 of human SMAP1 located downstream the ArfGAP domain and upstream the PTC generated by both insertion and deletion SMAP1 mutations. In HeLa and other cells with a *SMAP1*^{wt} allele, SMAP1 was detected using this antibody at the expected size for normal SMAP1 protein (55kDa). By contrast, no band could be detected at the sizes expected for the (A)9 and (A)11 SMAP1 forms (21kDa and 18kDa, respectively) that might be synthesized in the HCT116 and LoVo cells since they carry biallelic A9/A9 and A9/A11 mutations, respectively (Figure 4A and not shown). The absence of signal suggested that residual abnormal *SMAP1* transcript is not translated and/or that truncated proteins are degraded by the proteasome machinery. To determine if truncated SMAP1 is sensitive to proteasome degradation, we constructed an expression vector of the truncated SMAP1-A9 form (pcDNA3.1-SMAP1-A9-HA), corresponding to the mutation most commonly found in MSI tumors, tagged with an HA epitope at the C-

terminal extremity. HCT116 cells were transfected with pcDNA3.1-SMAP1-A10-HA corresponding to the wild-type form or with pcDNA3.1-SMAP1-A9-HA, and treated with the proteasome inhibitor MG132 or vehicle (Figure 4B). The p53 protein whose cellular levels are tightly regulated through ubiquitin-dependent proteasomal degradation was used to check that MG132 efficiently inhibited proteasome in our experiments (Figures 4B and 4C). As expected, compared to untreated cells, p53 levels increased at least twice in cells transfected with SMAP1-expression vectors upon treatment with MG132 (Figures 4B and 4C). By contrast, MG132 treatment did not increase the expression of the HA-tagged wild-type SMAP1 with a 0.8 ratio between treated and untreated cells, indicating that SMAP1-A10-HA was not sensitive to proteasome degradation (Figure 4B); whereas the expression of SMAP1-A9-HA increased upon MG132 treatment (ratio=1.8; Figures 4B and 4C), meaning that the truncated form of SMAP1 is sensitive to degradation by the proteasome. Hence, even if *SMAP1* abnormal transcripts are not completely degraded by the NMD system, the truncated protein translated is not stable in cells being degraded by the proteasome machinery. Altogether these results are in favor of loss-of-function of SMAP1 in cells harboring frameshift *SMAP1* mutations.

Effects of SMAP1 mutations on cell proliferation and cell cycle.

In order to investigate the consequences of SMAP1 inactivation, we transfected an expression vector of wild-type *SMAP1* cDNA in the SMAP1-null HCT116 cell line, and selected three clones that re-expressed SMAP1 at a level similar to that of HeLa cells (*SMAP1*^{wt}) (Figure 5A). Since Arf6 protein is involved in cytokinesis, we analyzed whether the absence of SMAP1 had consequences on cell cycle and proliferation by comparing the SMAP1-null HCT116 cell line with the three HCT116-SMAP1 clones. Restoring SMAP1 decreased cell proliferation of all three

HCT116-SMAP1 clones to the same extent; the number of cells recovered 48 hours after seeding were significantly lower compared to the HCT116 parental cells ($P = .010$, one-way ANOVA, Dunnett's test); cell numbers were lower as soon as 24 hours after seeding but the differences were not statistically significant ($P = .1236$, one-way ANOVA, Dunnett's test) (Figure 5B).

We further analyzed the cell cycle after synchronizing cells with nocodazole, an inhibitor of microtubule polymerization blocking cells during prometaphase. The transition from G2/M to G1 phase was faster for the HCT116 cells compared to the HCT116-SMAP1 clones, which remained longer in the G2/M phase. Indeed, more than 50% HCT116-SMAP1 clones remained in the G2/M phase after 4h of nocodazole chase compared to 20% for HCT116 ($P = .008$) (Figure 5E). Such a slower cell cycle transition may explain the proliferation defect observed for the HCT116-SMAP1 clones. SMAP1 re-expression had no effect on cell death (data not shown).

By performing a clonogenic assay, we observed that restoration of SMAP1 expression not only decreased proliferation of HCT116 cells, but also altered their ability to grow when seeded at a low concentration (Figures 5C and 5D). Indeed, the number of clones established from all three HCT116-SMAP1 cell lines varied from 28 to 41, which was significantly lower compared to the HCT116 parental cells that gave rise to 70 clones in average ($P = .02$ and $P = .04$, respectively, t test). Moreover, the clones derived from HCT116-SMAP1 cell lines were smaller (Figure 5D) in agreement with our results showing that restoring SMAP1 expression decreased cell proliferation (Figure 5B).

SMAP1 mutations consequences on TfR clathrin-mediated endocytosis.

Since SMAP1 is a GTPase-activating protein specifically acting on Arf6, a protein involved in clathrin-dependent endocytosis, we analyzed the impact of SMAP1 inactivation on TfR endocytosis, a classical model for clathrin-dependent endocytosis. To monitor endocytosis, we used fluorescent transferrin (Tf), since Tf remains bound to its receptor during the complete endocytosis process, i.e. internalization and recycling (15). The effects of SMAP1 expression on TfR endocytosis were measured in SMAP1-null HCT116 cells and in the 3 counterpart HCT116-SMAP1 clones. Cells have been exposed to a single initial pulse with labeled Tf at 4°C followed by an incubation at 37°C for various times, and fluorescence has been measured after removal of membrane bound Tf by acid wash. We first analyzed qualitative changes on TfR endocytosis by fluorescence microscopy (Figure 6A). No differences were observed in the initial binding of Tf to the cell surface ($t = 0$ min) or in the internalization of TfR when comparing the HCT116-SMAP1 clones to the HCT116 parental cells. Only the HCT116-SMAP1 C12 clone is shown as it behaved like the two others. Figure 6B represents the ratio between intracellular Tf relative to the amount of Tf bound at the beginning of the pulse measured by flow cytometry, with the ascending phase of the curve corresponding to TfR internalization and the descending phase to TfR recycling. Although TfR internalization tended to decrease upon SMAP1 re-expression, these differences were not significant ($P = .40$, Two-way ANOVA). So, all three HCT116-SMAP1 clones in which wild-type SMAP1 protein was re-expressed behaved similarly to the parental HCT116 cells both for the internalization process and the recycling to the plasma membrane, indicating that *SMAP1* mutations had no substantial consequences on TfR endocytosis (Figure 6B).

To monitor the TfR recycling phase more specifically, experiments were done with cells that were incubated during 30 min with labeled Tf at 37°C to reach a

plateau of maximal TfR internalization. Cells were then washed to remove free Tf and internal TfR was measured at various time points after removal of membrane-bound Tf. The first part of the curve represents the fast recycling step that occurs during the first 10 min of the experiment, while the slow recycling occurs after 10 min until the end of the experiment. The molecular mechanisms and the organelles engaged in both recycling step are different. As shown in Figure 6C, the slow TfR recycling process operating after 10 min did not significantly vary among the HCT116 clones re-expressing wild-type SMAP1 or compared with the parental HCT116 cells. However, the fast TfR recycling step, which operates during the first 10 minutes, was more efficient in the 3 HCT116-SMAP1 clones compared with the SMAP1-null HCT116 cell line (35% vs 14% at 6 minutes, respectively, $P = .03$ Two-way ANOVA).

SMAP1 seems to act as one of the Arf6GAPs involved in the regulation of Arf6 during TfR fast recycling step.

Discussion

In MSI CRCs, the challenge is to distinguish among the numerous mutations arising in genes harboring a coding mononucleotide repeat, the driver mutations playing a role in the oncogenic process from passenger mutations, occurring coincidentally due to an inefficient MMR system. Recently, *SMAP1*, a gene with a coding (A)₁₀ repeat, was shown to be mutated in 73% of primary MSI CRCs (7); such a mutation rate places *SMAP1* among the group of genes most highly frequently mutated in MSI CRCs. Nevertheless, few mutated genes have been clearly demonstrated as playing a role in oncogenesis, which is the case of the

transforming growth factor receptor 2 (6) and the activin receptor 2 (16) both involved in the TGF β pathway. In this study, we report a series of characteristics of *SMAP1* mutations in favor of a role in MSI oncogenesis. Loss-of-function *SMAP1* mutations occur in MSI tumors of various tissue origins, including CRCs, endometrium and gastric carcinomas, as well as lymphomas, not only in humans but also in murine lymphomas arising in *Msh2*^{-/-} mice in which *Smap1* mutations occur in the conserved (A)10 repeat and is correlated with increased mouse survival (data not shown). Interestingly, *SMAP1* has a homologous gene named *SMAP2* that codes an ArfGAP specific for Arf1, acting in the retrograde, early endosome-to-TGN pathway in a clathrin- and AP-1-dependent manner (12). *SMAP2* contains an (A)8 coding repeat that was not found mutated in the tumors and the cell lines analyzed (data not shown). Since the mean mutation frequency for (A)8 repeats has been evaluated to be around 10% in MSI CRCs and over 30% to be considered as possibly selected in these tumors, our observations suggest that *SMAP2* mutations are likely to be neutral in these tumors and may even be counter-selected. Since oncogenic pathways are tissue-specific, we cannot exclude that *SMAP2* mutations drive cancer development in other organs, as illustrated by the *BRCA1* and *BRCA2* genes, which both carry a coding (A8) repeat that is stable in MSI CRCs, but whose mutations are responsible for breast and ovarian cancers (17). Alternatively, *SMAP1* and *SMAP2* mutations may be considered as functionally redundant, as they both target intracellular traffic. Based on this hypothesis, it is not surprising that mutations target predominantly *SMAP1* since its repeat is longer than the one located in *SMAP2*, thus being more prone to polymerase slippage during replication. In MSI CRCs, the average mutation rate is approximately 3.5 times higher for (A)10 repeats than for (A)8 repeats (3). Altogether, our observations led us to propose to assign a role for *SMAP1* mutations in MSI colorectal oncogenesis.

Several cellular mechanisms exist to avoid the expression of abnormal truncated proteins that may act in a dominant negative manner. One is the NMD system, which degrades abnormal transcripts harboring a premature termination codon (5, 18). As expected, the presence of *SMAP1* mutations in cell lines from various tissue origins caused a significant decrease in *SMAP1* transcript levels, that was more marked for biallelic than for monoallelic mutations, with the exception of LoVo and SW48. The fact that *SMAP1* transcript levels remained in the normal range in LoVo and SW48 cells despite the presence of biallelic mutations suggests that NMD may be deficient in these cell lines. Indeed, both cell lines have a profound MMR defect due to inactivation of one of the major MMR component, i.e. MSH2 mutation and MLH1 epigenetic inactivation, respectively, which gives rise to the accumulation of numerous mutations all over the genome, possibly targeting one of the NMD actors. Yet, several genes coding proteins involved in the recognition of mRNAs containing PTC or in their degradation harbor coding mononucleotide repeats that are prone to MSI-driven mutations, notably SMG7 has one (A)₉ repeat and UPF3A has two such repeats. In addition, SMG1, UPF2, UPF3B and PARP1 all carry at least one repeat of 7 mononucleotides that may also be mutated in MSI tumors. In view of this, it would be interesting to evaluate the contribution of NMD defects in MSI oncogenesis. In a series of primary MSI CRCs, immunohistochemical analyses of *SMAP1* allowed us to visualize that, in some tumors, SMAP1 protein expression was heterogeneous, with some regions totally lacking detectable SMAP1, while others were normally labeled. In view of the possible role of *SMAP1* mutations in tumor progression, as suggested by the decrease in mutation frequency observed in the more advanced tumors, it would be interesting to investigate whether the risk for tumor recurrence is lower in tumors that are homogeneously lacking SMAP1 expression and if SMAP1 is less frequently mutated in the recurrences.

In order to get further insights in the molecular mechanisms through which SMAP1 inactivation drives colorectal oncogenesis, we set up an experimental system in which the SMAP1-null HCT116 cell line was compared to several isogenic clones derived by stably transfecting wild-type *SMAP1* cDNA. Because HCT116 is MMR-deficient, this cell line is hypermutable and has accumulated a collection of mutations throughout the genome, among which some have been selected, being present in the majority of the cell population if not all. By contrast, most mutations that do not confer a growth advantage to the cells will remain unselected and will thus be present in only a fraction of cells. It turns out that the HCT116 cell line, as well as all other MSI cell lines, represents a collection of cells with diverse genotypes displaying specific individual underrated characteristics. In order to perform functional experiments, we selected three independent clones that stably expressed SMAP1 at a level comparable to that of normal cells. Interestingly, all HCT116-SMAP1 clones behaved in the same manner, regardless of the functional test performed, indicating that the differences observed when compared with the parental HCT116 cell line was most likely due to SMAP1 rather than to an unrecognized trait that emerged in a particular clone.

SMAP1 has been previously described as being involved in clathrin-mediated endocytosis of TfR (8). In their experiments, HeLa cells that normally express SMAP1 internalized transferrin efficiently while those overexpressing wild-type SMAP1 or a GAP-negative mutant failed to incorporate transferrin, as did HeLa cells in which endogenous SMAP1 expression was silenced by siRNA (8). Our data show that restoring SMAP1 expression in HCT116 cells had no effect on TfR internalization, but accelerated the fast recycling step in all three HCT116-SMAP1 clones to the same extent. There are several possible reasons for these apparent discrepancies. The first one is that our experiments were performed using the

HCT116 cell lines and several counterpart clones expressing normal levels of SMAP1, while most of the experiments reported in the other study were done with HeLa cells either overexpressing wild-type or GAP-defective SMAP1 (8). Nevertheless, these authors also reported that TfR endocytosis was impaired using HeLa cells in which SMAP1 expression was silenced, a model that is closer to ours since MSI-driven *SMAP1* mutations abolish protein expression (8). In our experiments, TfR endocytosis was quantitatively evaluated by measuring labeled Tf intensity by FACS rather than observed microscopically, and the endocytosis process was followed over time by analyzing numerous time points to be able to detect abnormalities during both the early and late steps independently.

Considering that SMAP1 preferentially acts on Arf6, our results showing that SMAP1 inactivation perturbed the fast recycling step of TfR recycling to the plasma membrane fit well with the data recently reported by the group of Chavrier (19, 20). Indeed, they showed that Arf6 does not contribute to receptor-mediated clathrin-dependent endocytosis but that its interaction with JNK-interacting proteins on endocytic vesicles is required for TfR trafficking in the fast, microtubule-dependent endocytic recycling pathway (20). Because GTP loading and GTP hydrolysis are required for Arf6 incorporation and dynamics in clathrin-coated structures (20), it would be interesting to determine whether SMAP1 may be one of the yet-unknown Arf6-specific GAPs acting in this process.

We observed that the SMAP1-null HCT116 cell line grew faster than its isogenic clones in which SMAP1 expression was restored is reminiscent of one of the clinical features of MSI CRCs, as they tend to be larger-sized tumors compared with MSS CRCs. Increased cell proliferation associated with *SMAP1* mutations was due, as least in part, to a shorter G2/M phase. Interestingly, it was recently shown that expression of an Arf6 mutant unable to undergo GTP hydrolysis led to cytokinesis

defects, suggesting that the hydrolysis of the Arf6-GTP is essential for completion of cytokinesis (19). The Arf6GAPs involved in Arf6 regulation during cytokinesis are yet to be characterized. Interestingly, it was recently shown that expression of GTP-locked Arf6-Q67L or dominant inhibitory Arf6-T27N mutants leads to cytokinesis defects, suggesting that the Arf6 activity and the hydrolysis of the Arf6-GTP are both essential for completion of cytokinesis (19). Furthermore, Arf6-GTP interacts with the Sec10 subunit of the exocyst complex for the docking and exocytosis of vesicles (21). In parallel, it was demonstrated that Rab11- and FIP3-positive vesicles from the endosomal recycling compartment are essential for cytokinesis completion (22). Hence, since SMAP1 is unlikely to be involved in the slow recycling from the endosomal recycling compartment, the trafficking of Rab11- FIP3-positive vesicles should not be impaired upon SMAP1 inactivation. Thus, if SMAP1 were one of the Arf6GAPs involved in cytokinesis by controlling the vesicles exocytosis, its inactivation is expected to shorten the G2/M phase, increasing the Arf6-GTP amount at the plasma membrane and thus the interaction with the exocyst complex and the exocytosis, a hypothesis fitting well with our observations.

In conclusion, we report that loss-of-function SMAP1 mutations frequently occur in MSI tumors of various tissue origins, and we provide evidence for their implications in diverse biological functions, unraveling their role in oncogenesis. SMAP1 is an ArfGAP specific for Arf6, a small GTPase involved in several cellular mechanisms such as clathrin-mediated endocytosis of membrane proteins and actin remodeling. Owing to the recently reported role for Arf6, as well as several ArfGAP proteins, in regulating adhesion, migration, and cancer cell invasion (23-25), we propose to assign a new role for SMAP1 in this emerging oncogenic pathway.

References

1. Zaanani A, Meunier K, Sangar F, *et al.* Microsatellite instability in colorectal cancer: from molecular oncogenic mechanisms to clinical implications. *Cellular oncology*. 2011;34(3):155-176.
2. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res*. 2002;62(9):2447-2454.
3. Woerner SM, Yuan YP, Benner A, *et al.* SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Res*. 2010;38(Database issue):D682-689.
4. Loire E, Praz F, Higuier D, *et al.* Hypermethylability of genes in Homo sapiens due to the hosting of long mono-SSR. *Mol Biol Evol*. 2009;26(1):111-121.
5. Isken O, Maquat LE. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Dev*. 2007;21(15):1833-1856.
6. Wang J, Sun L, Myeroff L, *et al.* Demonstration that mutation of the type II transforming growth factor beta receptor inactivates its tumor suppressor activity in replication error-positive colon carcinoma cells. *The Journal of biological chemistry*. 1995;270(37):22044-22049.
7. El-Bchiri J, Guilloux A, Dartigues P, *et al.* Nonsense-mediated mRNA decay impacts MSI-driven carcinogenesis and anti-tumor immunity in colorectal cancers. *PLoS One*. 2008;3(7):e2583.
8. Tanabe K, Torii T, Natsume W, *et al.* A novel GTPase-activating protein for ARF6 directly interacts with clathrin and regulates clathrin-dependent endocytosis. *Mol Biol Cell*. 2005;16(4):1617-1628.

9. D'Souza-Schorey C, Chavrier P. ARF proteins: roles in membrane traffic and beyond. *Nat Rev Mol Cell Biol.* 2006;7(5):347-358.
10. Marcos I, Borrego S, Rodriguez de Cordoba S, *et al.* Cloning, characterization and chromosome mapping of the human SMAP1 gene. *Gene.* 2002;292(1-2):167-171.
11. Sato Y, Hong HN, Yanai N, *et al.* Involvement of stromal membrane-associated protein (SMAP-1) in erythropoietic microenvironment. *J Biochem.* 1998;124(1):209-216.
12. Natsume W, Tanabe K, Kon S, *et al.* SMAP2, a novel ARF GTPase-activating protein, interacts with clathrin and clathrin assembly protein and functions on the AP-1-positive early endosome/trans-Golgi network. *Mol Biol Cell.* 2006;17(6):2592-2603.
13. Kon S, Tanabe K, Watanabe T, *et al.* Clathrin dependent endocytosis of E-cadherin is regulated by the Arf6GAP isoform SMAP1. *Exp Cell Res.* 2008;314(7):1415-1428.
14. Zaanan A, Cuilliere-Dartigues P, Guilloux A, *et al.* Impact of p53 expression and microsatellite instability on stage III colon cancer disease-free survival in patients treated by 5-fluorouracil and leucovorin with or without oxaliplatin. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO.* 2010;21(4):772-780.
15. Qian ZM, Tang PL, Wang Q. Iron crosses the endosomal membrane by a carrier-mediated process. *Prog Biophys Mol Biol.* 1997;67(1):1-15.
16. Hempen PM, Zhang L, Bansal RK, *et al.* Evidence of selection for clones having genetic inactivation of the activin A type II receptor (ACVR2) gene in gastrointestinal cancers. *Cancer Res.* 2003;63(5):994-999.

17. Miquel C, Jacob S, Grandjouan S, *et al.* Frequent alteration of DNA damage signalling and repair pathways in human colorectal cancers with microsatellite instability. *Oncogene*. 2007;26(40):5919-5926.
18. Gardner LB. Nonsense-mediated RNA decay regulation by cellular stress: implications for tumorigenesis. *Molecular cancer research : MCR*. 2010;8(3):295-308.
19. Montagnac G, Sibarita JB, Loubery S, *et al.* ARF6 Interacts with JIP4 to control a motor switch mechanism regulating endosome traffic in cytokinesis. *Current biology : CB*. 2009;19(3):184-195.
20. Montagnac G, de Forges H, Smythe E, *et al.* Decoupling of activation and effector binding underlies ARF6 priming of fast endocytic recycling. *Current biology : CB*. 2011;21(7):574-579.
21. Prigent M, Dubois T, Raposo G, *et al.* ARF6 controls post-endocytic recycling through its downstream exocyst complex effector. *The Journal of cell biology*. 2003;163(5):1111-1121.
22. Fielding AB, Schonteich E, Matheson J, *et al.* Rab11-FIP3 and FIP4 interact with Arf6 and the exocyst to control membrane traffic in cytokinesis. *The EMBO journal*. 2005;24(19):3389-3399.
23. Sabe H. Requirement for Arf6 in cell adhesion, migration, and cancer cell invasion. *J Biochem*. 2003;134(4):485-489.
24. Hashimoto S, Onodera Y, Hashimoto A, *et al.* Requirement for Arf6 in breast cancer invasive activities. *Proc Natl Acad Sci U S A*. 2004;101(17):6647-6652.
25. Muralidharan-Chari V, Hoover H, Clancy J, *et al.* ADP-ribosylation factor 6 regulates tumorigenic and invasive properties in vivo. *Cancer Res*. 2009;69(6):2201-2209.

Abbreviations

Arf, ADP ribosylation factor; **bp**, base pair; **CALM**, clathrin assembly lymphoid myeloid leukemia; **CB**, clathrin box; **CRC**, colorectal cancer; **GAP**, GTPase activating protein; **GTPase**, Guanine triphosphatase; **MMR**, mismatch repair; **MSI**, microsatellite instability; **MSS**, microsatellite stable; **NMD**, nonsense mediated mRNA decay; **PTC**, premature termination codon; **SMAP1**, small ArfGAP 1

Acknowledgements

We thank Drs Alex Duval and Christophe Lamaze for their contribution in the initiation phase of the project, Drs Philippe Chavrier and Guillaume Montagnac for their constant support and helpful discussions. We are further grateful to Dr Peggy Cuilliere-Dartigues for sharing data. This work has been supported by grants from the Association pour la Recherche sur le Cancer (#5091) and from the Ligue Nationale Contre le Cancer (RS10/75-72).

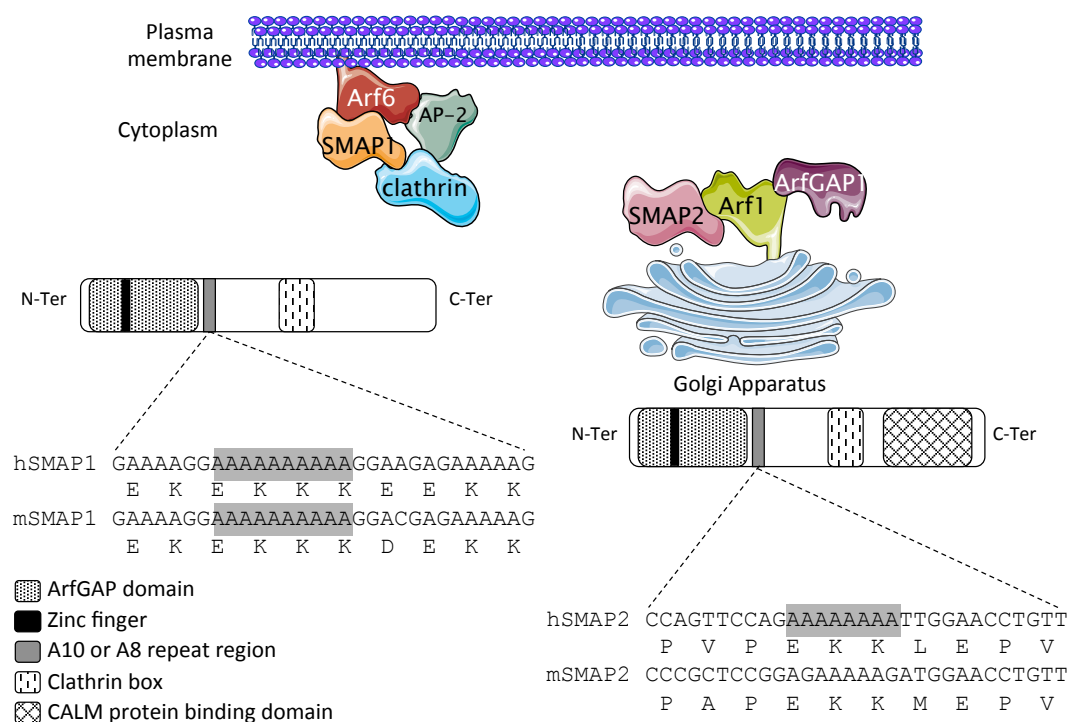


Figure 1

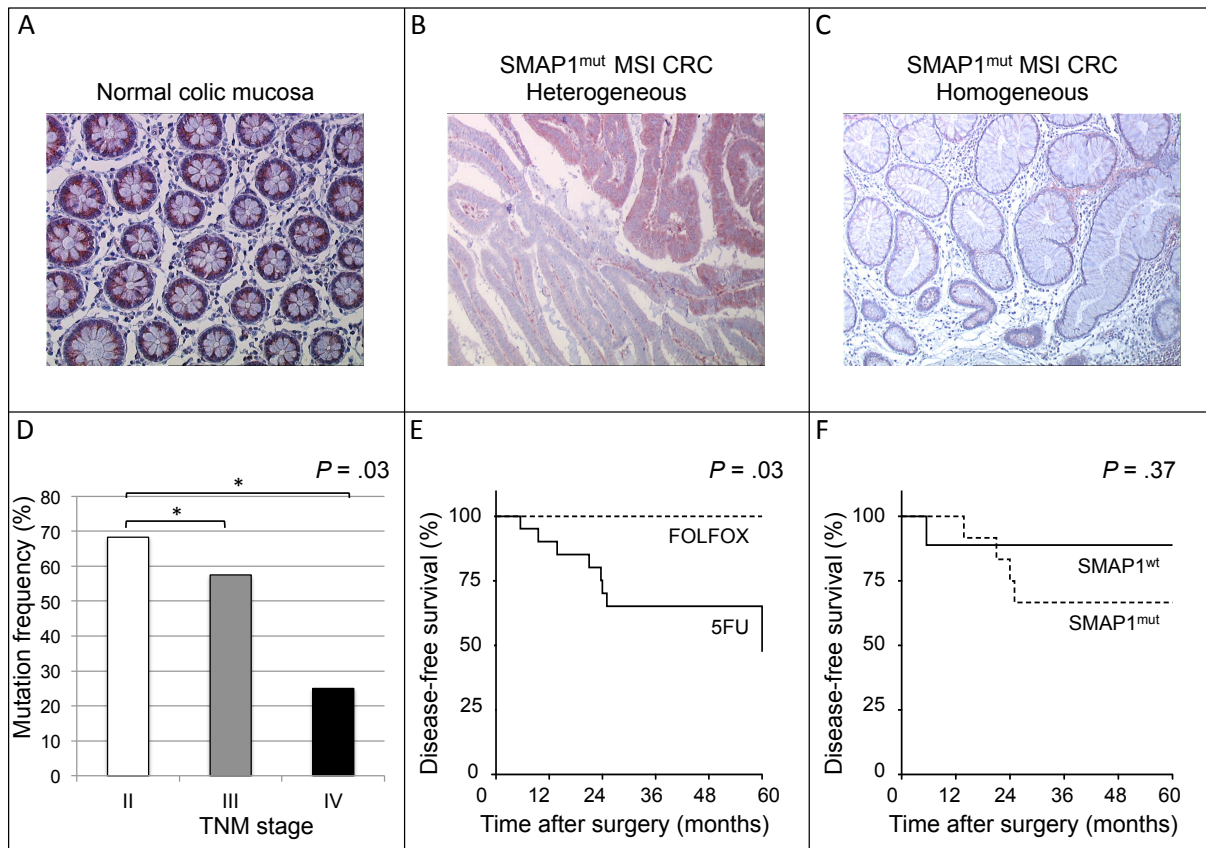


Figure 2

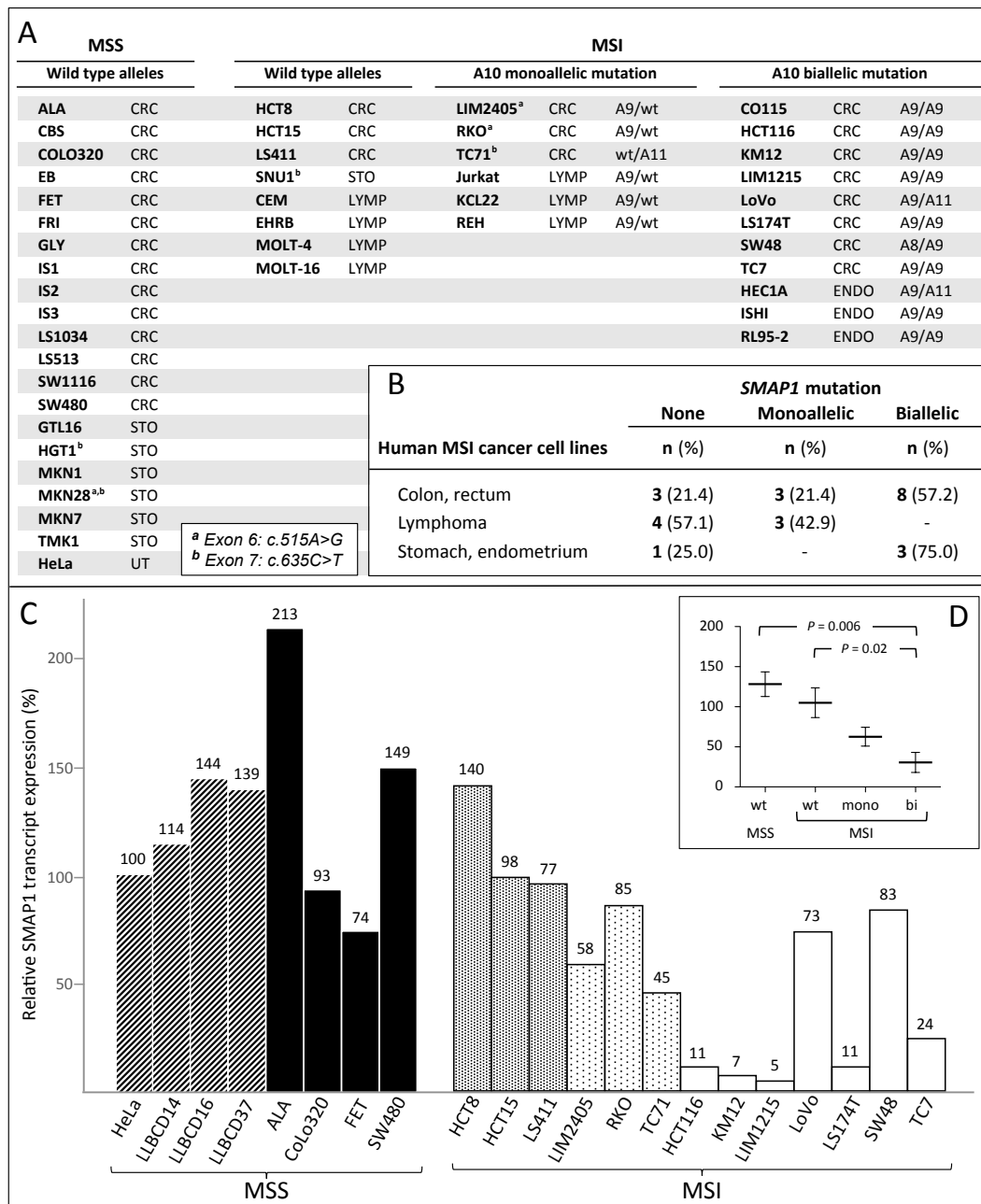


Figure 3

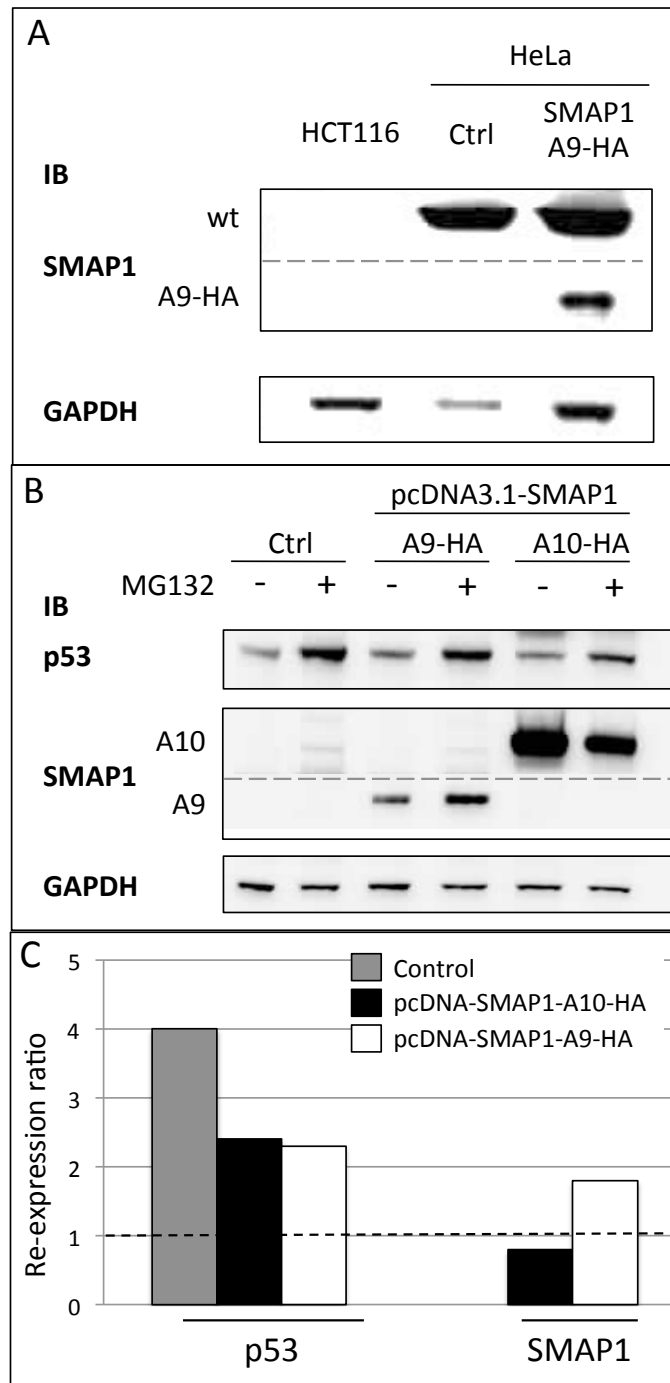


Figure 4

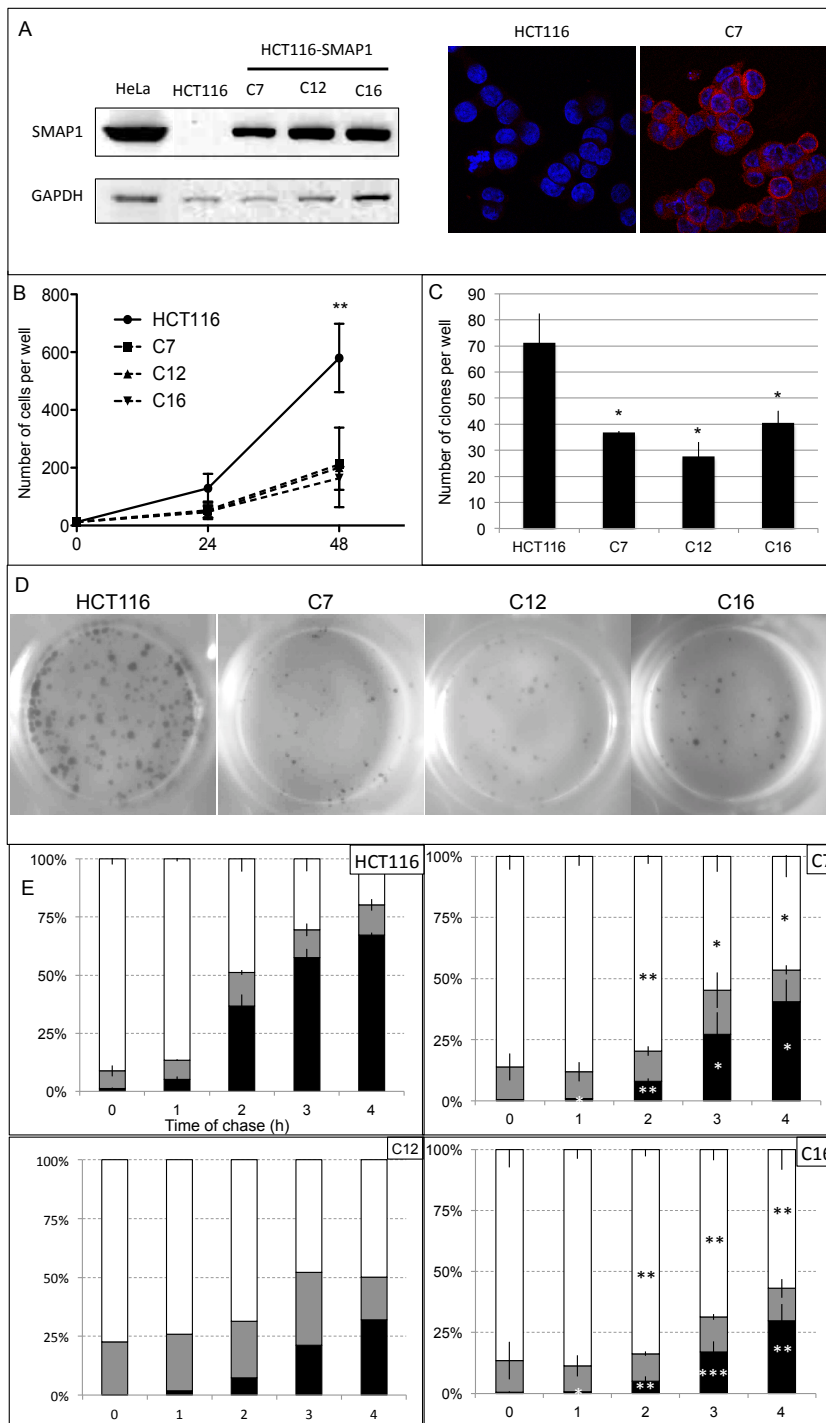


Figure 5

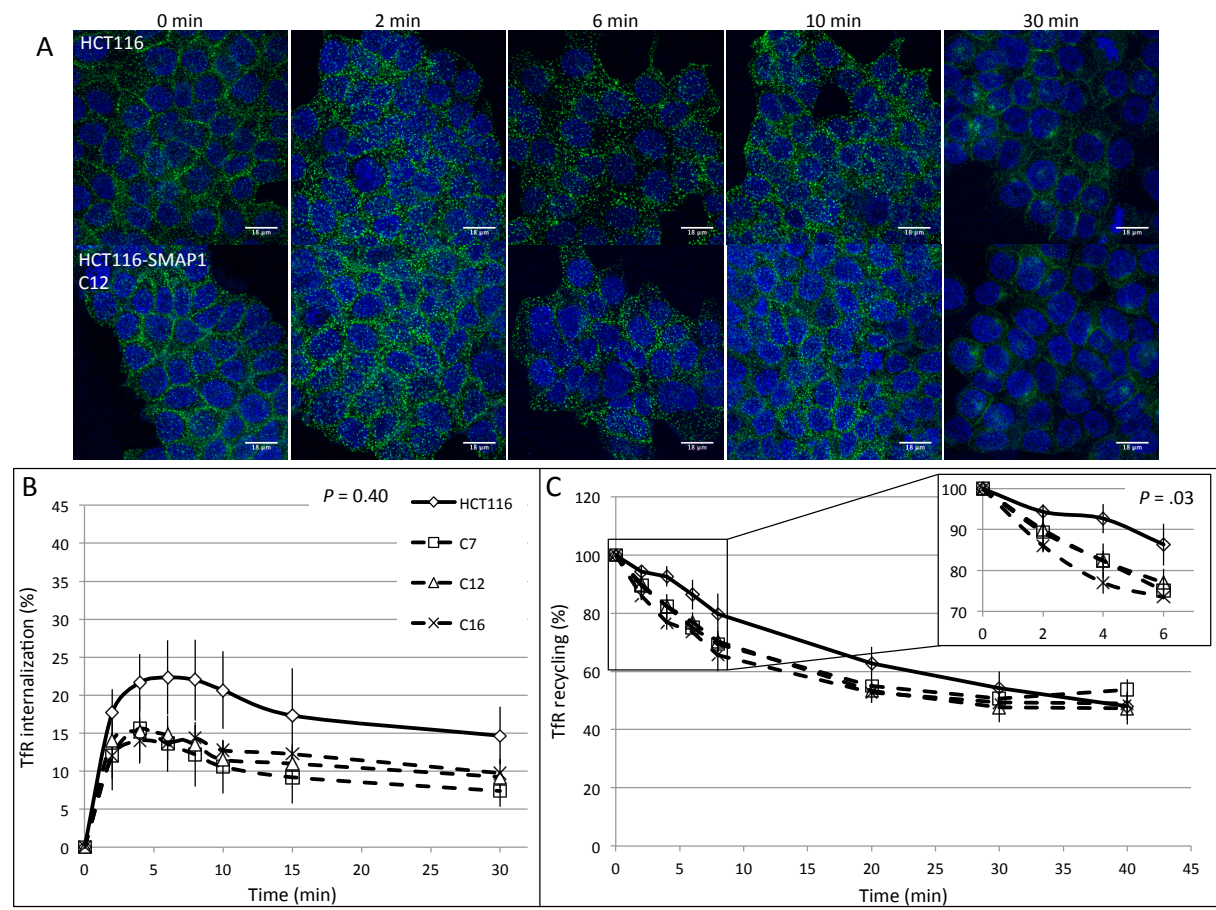


Figure 6

	SMAP1 ^{wt} (n = 17)		SMAP1 ^{mut} (n = 23)		
	No. of patients	%	No. of patients	%	<i>P</i>
Age					
Median	49.8		58.9		0.24
Range	31.1 - 76.0		30.8 - 90.6		
< 65 yr	13	76.5	12	57.1	0.19
Sex					
Male	10	58.8	10	47.6	0.53
Female	7	41.2	11	52.4	
Unknown			2		
Tumour location					
Proximal	14	82.4	15	71.4	0.69
Distal	3	17.6	6	28.6	
Unknown			2		
Differentiation grade					
Well / moderate	10	58.8	6	40.0	0.32
Poor	7	41.2	9	60.0	
Unknown			8		
Stage					
III A (T1-T2, N1)	2	11.8	1	4.8	0.86
III B (T3-T4, N1)	8	47.0	9	42.8	
III C (Tx, N2)	7	41.2	11	52.4	
Unknown			2		
Adjuvant chemotherapy					
Fluorouracil	9	52.9	11		
FOLFOX	8	47.1	5		
None			3		
Unknown			4		

Table 1

SEQUENCING PRIMERS			SEQUENCE (5' -> 3')	EXON SIZE (bp)	PCR PRODUCT SIZE (bp)
HS SMAP1	Exon 1	SMAP1-E1-F SMAP1-E1-R	cgccaggtgcgttcactct gaaggtcaccggtcgcaaga	190	257
HS SMAP1	Exon 2	SMAP1-E2-F SMAP1-E2-R	ccgagaatgcttctaataatgct cgtggcatatgccatttcaa	134	341
HS SMAP1	Exon 3	SMAP1-E3-F SMAP1-E3-R	tgtgtatgtatgtgcaggaatca cacagctcacagcaagtcaa	86	301
HS SMAP1	Exon 4	SMAP1-E4-F SMAP1-E4-R	cagctttgtggagtggcgtgaata acttctgctctgaagagtca	76	326
HS SMAP1	Exon 5	SMAP1-E5-F SMAP1-E5-R	tcttcctgtgcctgtcaattctc tccgaatgacacactgcaaaa	81	235
HS SMAP1	Exon 6	SMAP1-E6-F SMAP1-E6-R	taccgctgtgcctctgtctta gcacacagcattcactcaaca	81	450
HS SMAP1	Exon 7	SMAP1-E7-F SMAP1-E7-R	gtcattgcaacttcaacagtca acaggtacgttttggcatactc	88	380
HS SMAP1	Exon 8	SMAP1-E8-F SMAP1-E8-R	ggggtaggtagaaggagaacaaa tctagcctaagaacaggagact	125	398
HS SMAP1	Exon 9	SMAP1-E9-F SMAP1-E9-R	gctgcgctttactatgcaga acaaaagcactcactacatggaac	172	336
HS SMAP1	Exon 10	SMAP1-E10-F SMAP1-E10-R	gtgcaactacacgtgatagctc ttctttctgccagatcactagac	308	472
HS SMAP1	Exon 11	SMAP1-E11-F SMAP1-E11-R	tgaccaactgtgtggctaaaga cagtacagaccaacacgggtcaa	223	366
GENOTYPING PRIMERS			SEQUENCE (5' -> 3')		
HS SMAP1	A10 repeat	SMAP1-A10-HS-F-HEX SMAP1-A10-HS-R	HEX-ttatcaaaactttgggctgt cagctgtaagtggttttgc		
HS SMAP2	A8 repeat	SMAP2-A8-HS-F-FAM SMAP2-A8-HS-R	6FAM-ttatcaaaactttgggctgt cagctgtaagtggttttgc		
MM SMAP1	A10 repeat	SMAP1-A10-MM-F-FAM SMAP1-A10-MM-R	6FAM-ctgttctgttggacatgtga gtggttttcaggcttttc		

Supplementary Table 1

Figure legends

Figure 1. Protein domain structures and subcellular localizations of SMAP1 and SMAP2 proteins.

The *SMAP1* gene A10 repeat and the corresponding protein regions are conserved between human and murine species, whereas the *SMAP2* gene A8 repeat is not, although the corresponding protein region is conserved. SMAP1 protein is an ArfGAP of Arf6, a GTPase involved in clathrin-mediated endocytosis whereas SMAP2 is an ArfGAP specific for Arf1, involved in trans-Golgi network trafficking. SMAP1 ArfGAP domain: aa 18-136; SMAP1 Zn finger: aa 33-56; SMAP1 clathrin box: 218-222.

Figure 2. SMAP1 mutations in human MSI colorectal cancers.

A. SMAP1 expression pattern by immunohistochemistry in normal colic mucosa, **B.** in MSI CRC harboring an heterogeneous pattern, and **C.** in MSI CRC harboring an homogeneous pattern. **D.** SMAP1 mutation frequency according to tumor stage. Stage II: in situ tumors (n=41); stage III: tumors with lymph node metastasis (n=45); stage IV: tumors with distant metastasis (n=8). **E.** Percent of disease-free survival in patients developing a stage III MSI CRC, according to adjuvant chemotherapy treatment, FOLFOX (n=12) or 5FU (n=22). **F.** Percent of disease-free survival in patients with stage III MSI CRC treated by 5FU, according to SMAP1 mutations. Survival data were analyzed according to the method of Kaplan and Meier, and compared using a two-sided log-rank test.

Figure 3. Characterization of *SMAP1* mutations arising in tumor-derived cell lines and *SMAP1* transcript expression according to tumor *SMAP1* mutations.

A. *SMAP1* gene sequencing in cell lines derived from MSS or MSI tumors of various tissue origins. **B.** Distribution of MSI tumor-derived cell lines from various tissue origins accordingly to *SMAP1* mutation status, wild-type (no mutation), monoallelic or biallelic mutations. **C.** Relative *SMAP1* transcript expression in MSS and MSI tumor-derived cell lines. Hatched box: MSS cell lines derived from non-colorectal tumors; black box: MSS CRC-derived cell lines; closely punctated box: MSI CRC-derived cell lines with wild-type *SMAP1* allele only; largely punctated box: MSI CRC-derived cell lines harboring a monoallelic *SMAP1* mutation; white box: MSI CRC-derived cell lines harboring biallelic *SMAP1* mutations. **D.** Relationship between *SMAP1* transcript level in tumor-derived cell lines according to *SMAP1* mutation status using the Mann and Whitney test. Wt: wild-type; mono: monoallelic mutation; bi: biallelic mutations.

Figure 4: Truncated *SMAP1* protein expression and stability.

A. *SMAP1* expression was analyzed in the HCT116 cell line carrying a homozygous 1bp deletion *SMAP1* mutation (*SMAP1*-A9) and in HeLa cells (*SMAP1*^{wt}) transfected with the expression vector of the HA-tagged truncated *SMAP1* form corresponding to the 1 bp deletion mutation (*SMAP1*-A9-HA) or not (Ctrl). The membrane was revealed with an antibody directed against the *SMAP1* peptide 101-115 aa (Sigma); the 55kDa and 28kDa bands correspond to the endogeneous wild-type *SMAP1* protein and the *SMAP1*-A9 truncated protein, respectively. Loading controls were obtained using an HRP-conjugated anti-GAPDH antibody (ab9385, AbCam). **B.** HCT116 cells were transfected with pcDNA3.1-*SMAP1*-A9-HA or pcDNA3.1-*SMAP1*-HA and treated with a proteasome inhibitor, MG132, at 5μM during 8 hours. The expression of *SMAP1*-HA and *SMAP1*-A9-HA was detected by Western blot with an

anti-HA tag antibody. The p53 protein used as a control for proteasome inhibition was detected with the DO-1 monoclonal mouse antibody. **C.** Quantification of p53 and SMAP1 proteins upon proteasome inhibition. Results are represented as the ratio between expression in MG132-treated cells and untreated cells after quantification normalization using GAPDH as control.

Figure 5. Cell proliferation and cell cycle analysis.

A. Western blot analysis of SMAP1 expression in HeLa (SMAP1wt), HCT116 (SMAP1-null) and three clones derived from HCT116 cells by stably transfecting SMAP1; immunofluorescence detection of SMAP1 is shown for HCT116 cells and a representative HCT116-SMAP1 clone. **B.** To evaluate proliferation, cells were seeded at 12×10^3 per well and counted 24h or 48h after seeding. Results represent the mean of 3 independent experiments $\pm$ sem. **C** and **D.** For clonogenic assays, cells were seeded at 50 per well, and stained with Toluidine blue at day 6 to count clones. Results represent the mean of 3 independent experiments $\pm$ sem each performed in triplicate. **E.** For cell cycle analysis, cells were treated with nocodazole during 16 hours at 37°C, then washed and incubated and fixed at several time points. Cells were labelled with propidium iodide and analyzed by flow cytometry. Results are means of three independent experiments (except for C12, only one experiment) and are expressed as percentage $\pm$ s.e.m. Black box: G0/G1 phase; grey box: S phase; white box: G2/M phase.

Figure 6. Effects of SMAP1 mutations on transferrin receptor (TfR) internalization and recycling.

A. HCT116 cells and the HCT116-SMAP1 C12 clone stably expressing SMAP1 were pre-incubated with Tf-A488. Endocytosis was stopped at different time points and

images were obtained by confocal microscopy. **B.** HCT116 cells and HCT116-SMAP1 clones (C7, C12, and C16) were pre-incubated with Tf-A488, washed and incubated at 37°C to allow internalization. TfR Internalization was stopped at different time points and analyzed by flow cytometry. Results are expressed as the ratio of internalized Tf/surface bound Tf and represent the mean $\pm$ sem of at least four independent experiments. **C.** TfR recycling was analyzed in cells incubated with transferrin-A633 (Tf). After Tf excess was eliminated, cells were incubated in a Tf-A633 free medium and analyzed at various time points by flow cytometry. Results are expressed as the ratio of internalized Tf-A633 at various time points/total internalized Tf-A633 and represent the mean $\pm$ sem of at least three independent experiments.

Table 1 : Clinical features of patients affected with a TNM stage III MSI CRC according to SMAP1 mutations.

II. Résultats complémentaires

1. Matériels et Méthodes

1.1. *Immunofluorescence et cytométrie en flux*

Pour l'immunofluorescence, les cellules ont été cultivées sur lamelles de verre pendant 48 h. Toutes les incubations autres que celles des anticorps ont été réalisées pendant 15 min à température ambiante. Les cellules ont d'abord été fixées au para-formaldéhyde 2% puis, perméabilisées dans un mélange acétone/méthanol (1/1, v/v). Enfin, les cellules ont été incubées avec du PBS-BSA 1% afin de bloquer les sites non spécifiques d'attachement des anticorps. Une fois préparées, les cellules ont été incubées avec les anticorps primaires dilués dans du PBS-BSA 1%, anti-E-cadhérine (Mouse monoclonal, clone HECD-1, AbCam) au 1/400 ; anti- β -caténine (rabbit polyclonal antibody, C2206, Sigma-Aldrich) au 1/500 pendant 1 h à température ambiante. Après lavages avec du PBS, l'anticorps secondaire couplé à l'Alexa Fluor 488 (Life Technologies, France) dilué au 1/200 dans du PBS-BSA 1% a été ajouté aux cellules pendant 1 h à l'obscurité. Les noyaux des cellules ont ensuite été colorés avec une solution de DAPI 300nM pendant 10 min. Après montage sur une lame de microscope, les cellules ont été observées au microscope confocal.

Pour l'analyse en cytométrie en flux, les cellules ont été cultivées dans une plaque 6 puits pendant 48 h, décollées avec une solution de PBS-EDTA 1%, et 10^6 cellules ont été incubées avec l'anticorps anti-E-cadhérine dilué au 1/400 dans du PBS-BSA 1% pendant 1 h à 4°C. Après lavages avec une solution de PBS froide, les cellules en suspension ont été incubées avec l'anticorps secondaire couplé à l'Alexa Fluor 488 dilué au 1/200 dans du PBS-BSA 1% pendant 1 h à 4°C. Les cellules ont été

lavées avec du PBS puis analysées par cytométrie en flux (BD LSRII, BD Biosciences).

1.2. *Migration en deux dimensions*

Les cellules, 5×10^4 dans chaque chambre des inserts IBIDI (Biovalley, France), ont été cultivées dans du milieu sans sérum jusqu'à ce qu'elles soient confluentes (Figure 23A). Une fois les cellules confluentes, les inserts ont été retirés et du milieu avec 10% de sérum a été ajouté aux cellules. Le moment du retrait de l'insert constitue le T0 de la migration. En plus du T0, les cellules ont été prises en photos à différents intervalles de temps afin de suivre la fermeture de la blessure. L'analyse des photos a été effectuée à l'aide du logiciel ImageJ.

1.3. *Invasion à travers une matrice de Matrigel*

Pour ces expériences, nous avons utilisé des inserts pré-coulés avec du Matrigel (BD Matrigel Invasion Chamber, BD Biocoat, BD Biosciences, France) (Figure 24A). Les inserts ont été utilisés selon le protocole du fournisseur. Les cellules ont été incubées sur le Matrigel à raison de $1,5 \times 10^5$ cellules dans du milieu sans sérum. Les inserts sont placés dans un puits de plaque 24 puits contenant du milieu avec 10% de sérum. Les cellules ont été incubées à 37°C pendant 24 h. Le Matrigel est retiré de l'insert avec un coton-tige afin de se débarrasser des cellules incapables d'envahir la matrice. Les cellules qui sont passées sur la face inférieure de l'insert sont fixées au méthanol 100% froid pendant 10 minutes à -20°C puis, colorées au DAPI 300nM afin de marquer les noyaux pendant 10 min à l'obscurité. Les inserts sont ensuite montés entre lame et lamelle dans du milieu de montage (Aqua/Poly Mount, Tebu-Bio) pour permettre la prise de photos au microscope à fluorescence.

Les images ont été analysées à l'aide du plugin « Comptage Boyden » du logiciel ImageJ, gracieusement mis à disposition par l'unité INSERM U903 à Reims, France.

2. Résultats

2.1. Les mutations de SMAP1 stabilisent les jonctions adhérentes

Les protéines SMAP1 et Arf6 ont été décrites comme impliquées dans l'internalisation de la E-cadhérine [86] [79]. Nous avons donc analysé les conséquences des mutations de SMAP1 sur l'endocytose de la E-cadhérine. Pour cela, nous avons utilisé le système cellulaire HCT116 muté pour SMAP1 et les trois clones ré-exprimant SMAP1 (HCT116-SMAP1). La complémentation en SMAP1 des cellules HCT116 résulte en la relocalisation de la E-cadhérine de la membrane plasmique au cytoplasme (Figure 21A), la localisation membranaire diminuant d'au moins 25% (Figure 21B). Cette relocalisation s'accompagne de la diminution de l'expression de la E-cadhérine dans les clones HCT116-SMAP1 (Figure 21C). La perte de l'expression membranaire de la E-cadhérine dans les clones HCT116-SMAP1 entraîne un changement morphologique de cellules. En effet, les cellules HCT116 sont cohésives et forment des amas cellulaires compacts, les clones HCT116-SMAP1 sont plus étalés sur le support et forment des protrusions membranaires (Figure 21D).

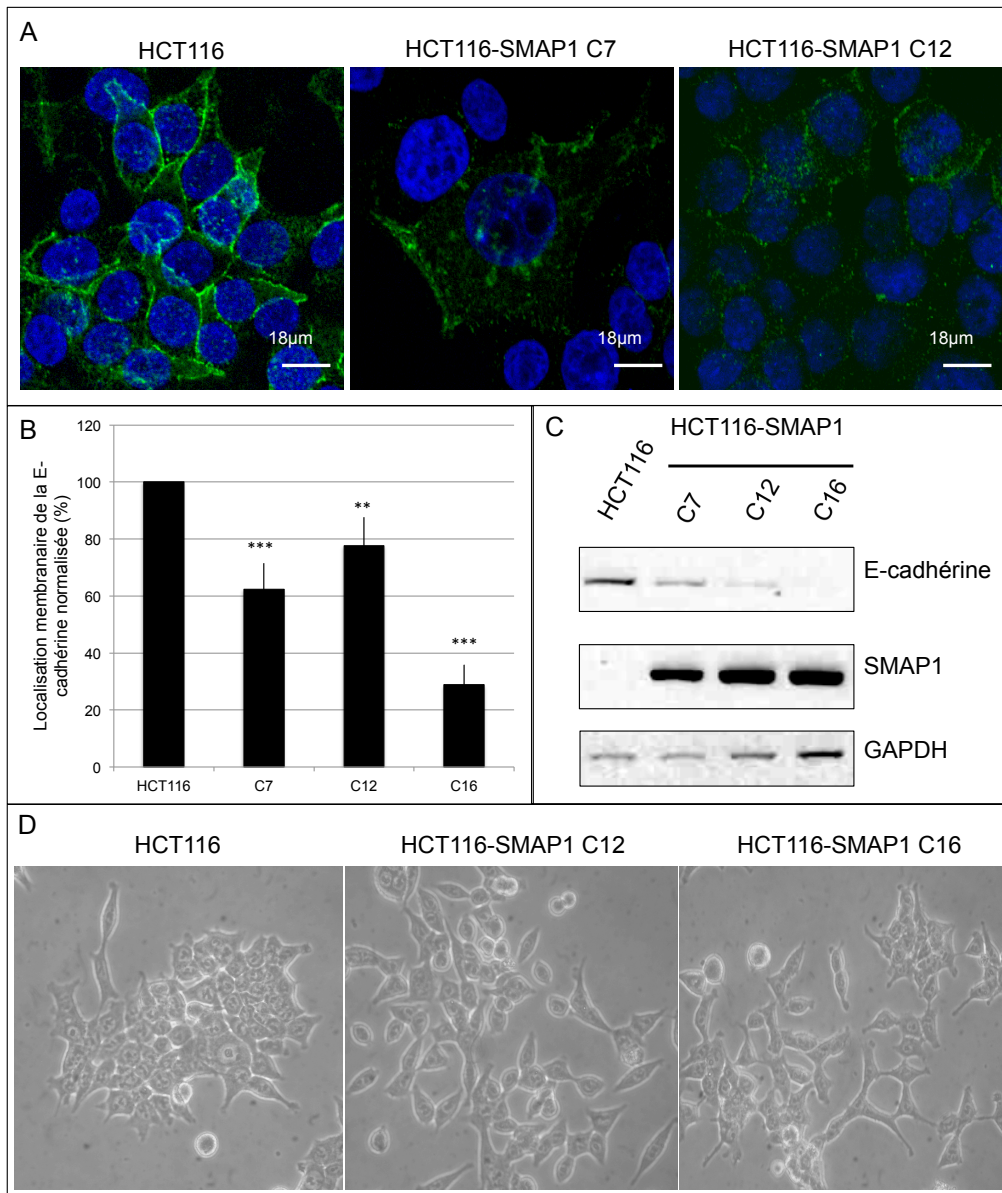


Figure 21: Localisation subcellulaire de la E-cadhérine. A. Immuno-fluorescence anti-E-cadhérine, en vert, sur les cellules HCT116 et 2 clones HCT116-SMAP1, C7 et C12. En bleu, les noyaux. B. Mesure du pourcentage de E-cadhérine à la MP dans les clones par rapport à HCT116. Ces résultats ont été obtenus en cytométrie en flux après immuno-marquage de la E-cadhérine et représentent 3 expériences indépendantes $\pm$ s.e.m. C. Western blot de lysats protéiques totaux. D. Morphologie des cellules en microscopie à contraste de phase.

L'internalisation de la E-cadhérine est couplée à celle de la β -caténine. Nous avons donc étudié la localisation cellulaire de cette protéine, qui une fois internalisée est normalement dégradée par le protéasome ou transférée au noyau. En concordance

avec les résultats obtenus pour la E-cadhérine, les clones HCT116-SMAP1 présentent une localisation cytoplasmique et surtout nucléaire de la β -caténine (Figure 22).

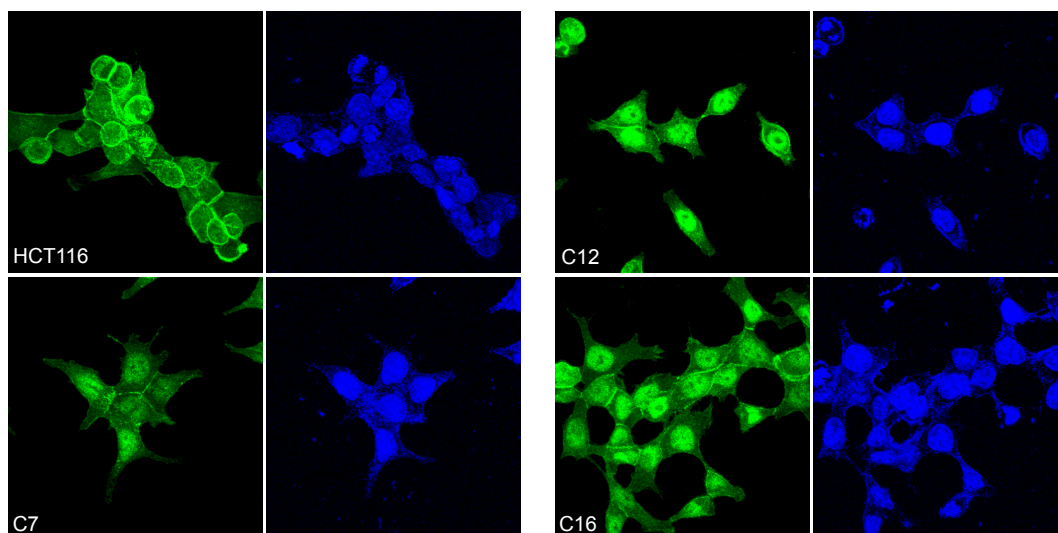
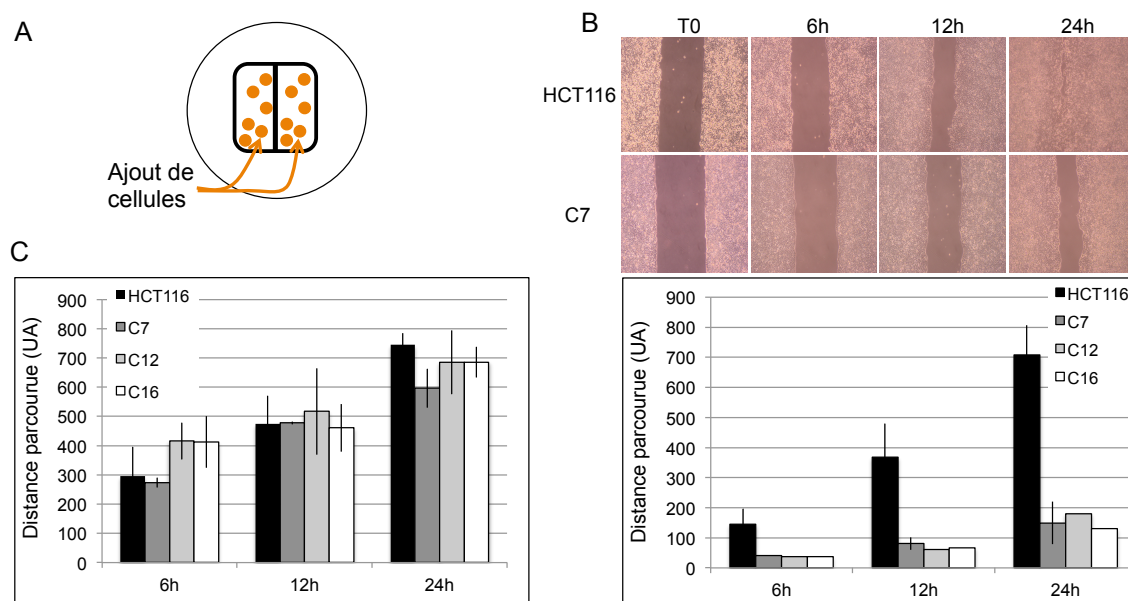


Figure 22: Localisation subcellulaire de la β -caténine. Immunofluorescence anti- β -caténine, en vert. En bleu, les noyaux.

2.2. Les mutations de SMAP1 augmentent la migration cellulaire

Les résultats obtenus sur les localisations cellulaires de la E-cadhérine et de la β -caténine nous permettent d'émettre l'hypothèse d'une transition épithélio-mésenchymateuse initiée par la réexpression de SMAP1 dans les cellules HCT116. Nous avons donc analysé la migration des cellules exprimant ou non SMAP1 par un test de cicatrisation (*wound-healing*). Les cellules sont cultivées jusqu'à confluence dans des inserts présentant deux chambres séparées par une barrière centrale mimant une blessure (Figure 23A). Après retrait de l'insert, la cicatrisation est observée jusqu'à fermeture. Sur un support plastique, les clones HCT116-SMAP1 migrent 4,5 fois moins que les cellules parentales (Figure 23B) ce qui va à l'encontre des résultats obtenus sur la stabilité des jonctions adhérentes. Afin de s'assurer que

ces résultats ne soient pas des artefacts dus aux conditions de culture sur plastique, nous avons reproduit cette expérience sur un substrat de collagène de type I. Dans ces conditions, aucune différence significative n'est observée entre les clones et les



cellules parentales (Figure 23C).

Figure 23: Migration deux dimensions. A. Représentation schématique de l'insert. B. Migration des cellules sur support plastique. Les histogrammes sont la quantification de la cicatrisation visualisée sur les photos et représentent au moins une expérience. C. Migration des cellules sur collagène de type I.

2.3. Les mutations de *SMAP1* diminuent le pouvoir invasif des cellules

Les résultats de migration en deux dimensions n'étant pas en accord avec les résultats de stabilité des jonctions adhérentes, nous avons étudié la capacité des cellules à migrer et à digérer une matrice de Matrigel, mimant la lame basale des épithélia. Pour cela, nous avons ensemencé les cellules sur un gel de Matrigel

déposé au dessus d'une membrane perforée de pores de 8µm dans un milieu sans sérum (Figure 24A). La membrane est placée dans une plaque 24 puits contenant du milieu avec sérum jouant le rôle d'attractant. Les cellules sont incubées pendant 24h afin de leur permettre de migrer et de dégrader le Matrigel et ainsi se retrouver sur la face inférieure de la membrane. Les cellules ayant traversé le Matrigel et la membrane sont comptabilisées. Ce test a permis de mettre en évidence un pouvoir invasif plus important des clones HCT116-SMAP1 par rapport aux cellules parentales en concordance avec les résultats sur la stabilité des jonctions adhérentes (Figure 24B). La différence observée entre les données de migration et d'invasion ne semble pas si étonnante. En effet, les cellules se comportent différemment en deux et en trois dimensions. L'expérience d'invasion se rapproche plus de ce qui peut se passer *in vivo* par la présence du gel de Matrigel qui mime la lame basale.

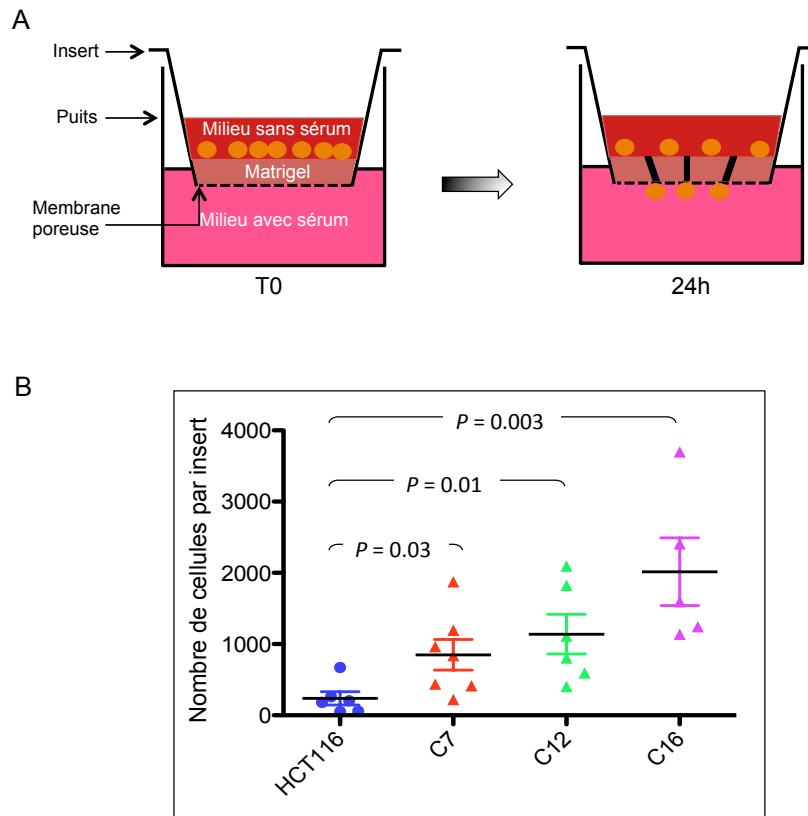


Figure 24: Invasion à travers un gel de Matrigel. A. Représentation schématique de l'expérience. B. Mesure de l'invasion des cellules après 24h d'incubation.

DISCUSSION & PERSPECTIVES

PARTIE III. DISCUSSION & PERSPECTIVES

I. Discussion

Ma Thèse portant sur « **l'étude fonctionnelle de SMAP1 : un nouveau gène à la croisée du trafic vésiculaire et de l'oncogenèse** » avait pour objectifs :

- la caractérisation des mutations de SMAP1 ;
- l'étude de l'impact des mutations de SMAP1 sur ses fonctions cellulaires ;
- l'analyse des conséquences des mutations sur le pouvoir invasif des cellules tumorales

Les travaux réalisés pendant ma Thèse ont mis en évidence un nouveau gène cible de l'oncogenèse MSI, tant d'un point de vue moléculaire que fonctionnel, le gène SMAP1. Il s'agit, à ma connaissance, de la première démonstration de l'implication d'un gène du trafic membranaire dans l'oncogenèse colorectale MSI. Les mutations de SMAP1 dans les CCR MSI ont des conséquences fonctionnelles qui expliquent en partie des caractéristiques cliniques des tumeurs MSI telles que leur volume et leur faible invasivité.

Dans la suite de mon manuscrit, je discuterai les résultats obtenus dans les trois objectifs initiaux et les confronterai avec les données actuelles de la littérature.

1. La caractérisation des mutations de *SMAP1* dans les tumeurs

Mon projet de Thèse a vu le jour à la suite des travaux réalisés au sein du laboratoire [21]. Cette étude a permis de mettre en évidence plusieurs nouveaux gènes potentiellement cibles de l'oncogenèse MSI, dont le gène *SMAP1*. La fréquence de mutations de *SMAP1* dans une collection de 44 CCR MSI était de 73%, plaçant le gène *SMAP1* parmi le groupe de gènes très fréquemment muté. Notre étude porte sur 89 CCR MSI et définit la fréquence de mutations de *SMAP1* à 60%. Cette différence observée n'est pas étonnante. La première explication possible est le stade TNM des tumeurs (Table 3). En effet, la collection utilisée dans l'étude de El-Bchiri et al., comportait au moins 25 tumeurs de stade II, 14 tumeurs de stade III et 1 tumeur de stade IV alors que la répartition au sein de notre collection est plus homogène entre les stades II et III, 41 et 40 respectivement. Nous avons montré que moins la tumeur était de stade avancé et plus la fréquence de *SMAP1* est élevée (68% en stade II, 51% en stade III et 25% en stade IV). Ceci explique la fréquence de mutation plus élevée dans l'étude d'El-Bchiri et al., enrichie en tumeurs de stade II. De plus, bien que nous ayons enrichi notre collection de tumeurs MSI par rapport à l'étude d'El-Bchiri et al. en doublant le nombre de cas, nos effectifs restent faibles, ce qui pourrait expliquer les différences observées. Une autre explication, plus technique, est également possible. En effet, l'extraction d'ADN est réalisée à partir de tissu congelé ou inclus en paraffine qui contient, en plus de la tumeur, du tissu sain. Malgré l'identification de la zone tumorale et la macro-dissection, l'élimination complète du tissu normal n'est pas possible. Ainsi, selon le taux de contamination par le tissu sain, le résultat du génotypage peut varier. C'est également la raison pour laquelle il n'est pas possible de définir le génotype précis d'une tumeur ; nous définissons les tumeurs par leur statut sauvage ou muté. De plus, les mutations par

insertion/délétion dans un contexte MSI sont très dynamiques ; ainsi, au sein même de la tumeur le génotype pour un gène donné peut être hétérogène.

Les mutations de SMAP1 sont spécifiques des tumeurs MSI quelque soit l'origine tissulaire. En effet, notre étude sur 46 lignées cellulaires dérivées de cancers MSI (n=25) ou non (n=21) a permis de mettre en évidence la spécificité phénotypique des mutations de SMAP1 puisqu'aucune lignée non-MSI n'est mutée alors que 68% des lignées MSI le sont. De plus, ces mutations ne sont présentes que dans le microsatellite poly(A10) du gène. La majorité de ces mutations consiste en la délétion d'une paire de bases, transformant le poly(A10) en poly(A9) et décalant ainsi le cadre de lecture. Le transcrit anormal comporte un codon STOP prématuré en position 575. La cellule dispose d'un système de surveillance des ARNm permettant la dégradation des ARN comportant un codon STOP prématuré évitant ainsi l'expression de protéines tronquées anormales : le système NMD (*Non-sense Mediated mRNA Decay*) [75]. Ainsi, le transcrit de SMAP1 dans les tumeurs MSI et dans les lignées cellulaires MSI est diminuée. Dans les tumeurs MSI, la contamination par le tissu sain ne permet pas de quantifier avec précision le niveau d'expression des transcrits qui malgré tout diminue d'un facteur 2. Le système NMD n'étant pas totalement efficace, il est possible que les transcrits anormaux soient traduits en une protéine tronquée présentant un néo-peptide C-terminal (Figure 4), expliquant le caractère immunogène des tumeurs MSI. Ainsi, nous avons recherché la présence éventuelle d'une protéine SMAP1 tronquée, dont la taille théorique est de 21kDa, en western blot dans la lignée cellulaire HCT116 dérivée d'un CCR MSI. Les mutations de SMAP1, en particulier la mutation poly(A9), aboutissent à une absence de protéine détectable en western blot. Cette lignée cellulaire n'exprime que 11% de transcrit résiduel (Figure 2C de l'article) et nous avons également démontré que la protéine tronquée poly(A9) exprimée transitoirement dans les cellules est

sensible à la dégradation par le protéasome (Figure 4B et 4C de l'article). De plus, la lignée de CRC MSI, LoVo, mutée de manière bi-allélique pour *SMAP1*, exprime plus de 70% de transcrit résiduel (Figure 3 C de l'article) mais aucune bande correspondant à la taille théorique n'a jamais été observée en western blot, ce qui concorde avec nos résultats d'inhibition du protéasome. Ces données d'expression nous ont permis de conclure raisonnablement sur une mutation de type perte-de-fonction et ceci, bien qu'il soit possible que l'absence de protéine détectable en western blot ne soit due qu'au seuil de détection de la technique. Les deux études fonctionnelles menées sur les gènes cibles de l'oncogénèse MSI, TGF β -RII [18] et ACVR2 [20], n'ont pas mis en évidence la présence d'une protéine tronquée dans les lignées cellulaires utilisées, HCT116 et HCT15 respectivement. Alors qu'une étude sur le facteur de transcription TCF4 a mis en évidence l'expression d'une protéine TCF4 tronquée dans la lignée LoVo dérivée d'un CRC MSI [177]. Ainsi, l'expression d'une protéine tronquée est dépendante du gène étudié. En effet, les codons STOP prématurés sont reconnus plus ou moins efficacement par le système NMD selon leur position dans le transcrit [75].

2. Le rôle de SMAP1 dans l'endocytose clathrine-dépendante du récepteur de la transferrine

La protéine SMAP1 est une Arf6-GAP qui serait impliquée dans l'endocytose clathrine-dépendante (CME) de protéines membranaires tels que le récepteur à la transferrine (TfR) et plus précisément lors de son internalisation [77]. Nos résultats obtenus dans la lignée cellulaire HCT116 et les trois clones ré-exprimant stablement

SMAP1 (HCT116-SMAP1) ne sont pas cohérents avec les résultats de Tanabe. En effet, nous montrons que les mutations de SMAP1, donc son absence d'expression, aboutissent à un défaut dans le recyclage rapide du TfR mais n'ont aucune conséquence détectable sur son internalisation (Figure 5 de l'article). Nos résultats sont en concordance avec ceux obtenus par l'équipe de Philippe Chavrier [3, 85]. Ces travaux ont permis de mettre en évidence le rôle d'Arf6 dans le recyclage du TfR plutôt que dans son internalisation. La forme active Arf6-GTP s'associe avec la protéine JIP3/4 pour faciliter la liaison avec la dynactine, protéine motrice des microtubules permettant le déplacement de la vésicule cargo vers l'intérieur de la cellule. Ainsi, la diminution de l'hydrolyse d'Arf6-GTP diminue l'association des vésicules cargos avec la protéine motrice des microtubules kinésine-1 qui permet leur déplacement vers la membrane plasmique [3]. L'absence de SMAP1 rompt l'équilibre du cycle GTP/GDP d'Arf6, en faveur de la forme Arf6-GTP, équilibre essentiel à son bon fonctionnement. Cependant, l'absence d'Arf6 induit un défaut de recyclage qui se prolonge jusqu'à 30 minutes après le début de l'expérience [85], alors que l'absence de SMAP1 n'a de répercussions que lors du recyclage rapide, c'est-à-dire pendant les 6 premières minutes de l'expérience (Figure 5 de l'article). Une protéine GEF et une protéine GAP seraient dédiées à chaque fonction cellulaire d'Arf6 [89]. Ainsi, d'après nos résultats, il est tentant d'émettre l'hypothèse selon laquelle SMAP1 serait la Arf6-GAP impliquée dans la régulation de la fonction d'Arf6 lors du recyclage rapide du TfR (Figure 25).

Dans notre système cellulaire, où nous comparons les cellules HCT116 avec les clones HCT116-SMAP1, dans lesquels un niveau physiologique de SMAP1 est ré-exprimé, les taux d'internalisation et de recyclage sont plus bas que ceux obtenus avec les cellules HeLa, modèle cellulaire de l'étude de la CME. Le taux maximal d'internalisation atteint dans les cellules HCT116 est de 22% et celui du recyclage

est de 50% alors que dans les cellules HeLa, ceux-ci avoisinent les 70% [85]. Cette différence élevée peut s'expliquer par le type cellulaire utilisé. En effet, la lignée cellulaire HeLa est dérivée d'une tumeur MSS du col de l'utérus alors que la lignée cellulaire HCT116 est dérivée d'une tumeur colorectale MSI mucineuse. De plus, d'autres systèmes de captation du fer sérique ont été décrits dans les cellules intestinales notamment via la signalisation mucine-mobilferrine-intégrine [178]. La mucine capterait le fer présent dans la lumière intestinale le rendant soluble. Les intégrines présentes à la surface membranaire permettraient le transfert du fer à la mobilferrine, protéine navette cytosolique [178].

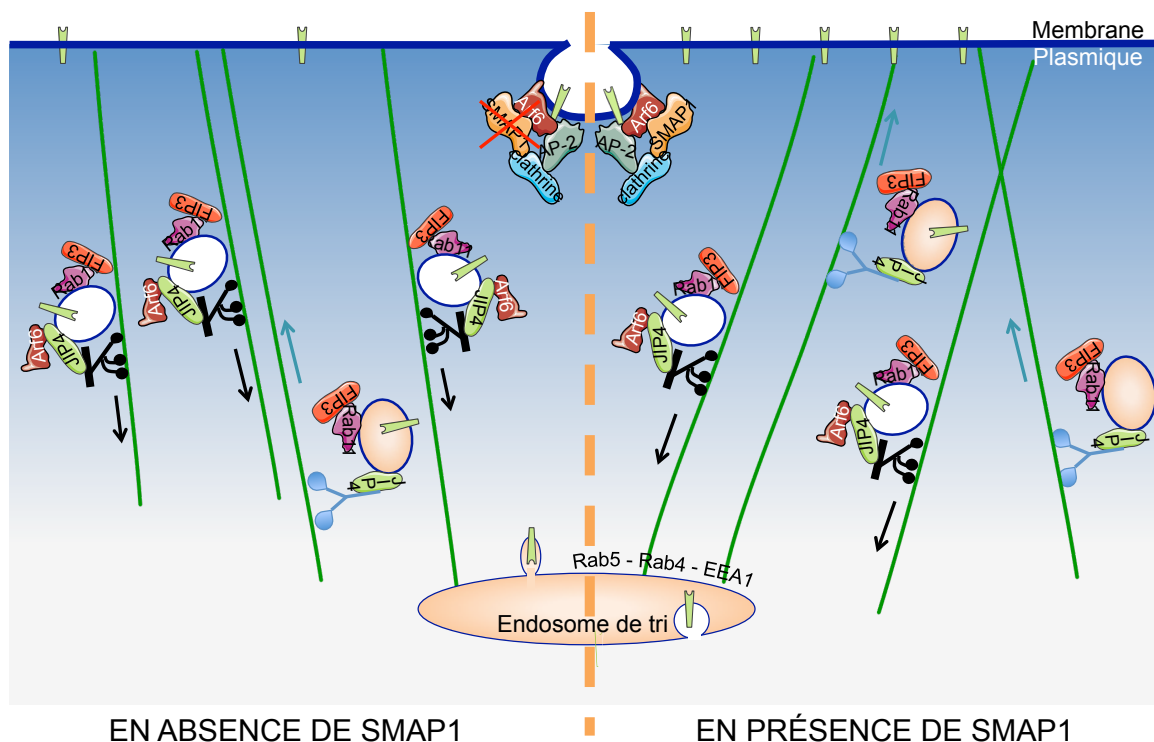


Figure 25 : Modèle du rôle de SMAP1 dans le recyclage rapide du TfR. L'absence de SMAP1 dans les cellules HCT116 (à gauche) n'empêche pas l'internalisation du TfR. Arf6-GTP se lie à JIP3/4 qui interagissent avec la dynactine permettant le transport des vésicules vers les différents organelles intracellulaires. L'absence de SMAP1 ne permet pas l'hydrolyse du GTP d'Arf6 ; ainsi la kinésine 1 n'interagit pas avec JIP3/4 et le recyclage rapide du TfR est diminué. Après réexpression de SMAP1 (à droite), la quantité d'Arf6-GTP diminue, JIP4 est libre d'interagir avec la kinésine 1 permettant de rétablir un recyclage rapide plus élevé.

3. Le rôle de SMAP1 dans la cytokinèse

De manière inédite, j'ai obtenu des résultats indiquant que la protéine SMAP1 pourrait être impliquée dans la complétion du cycle cellulaire (Figure 6 de l'article). Le rôle d'Arf6 dans la phase d'abscission lors de la cytokinèse, dernière étape du cycle cellulaire, est bien documenté dans la littérature [150] [149] [9] [3]. C'est principalement pour son rôle lors du recyclage lent qu'Arf6 a été décrit lors de l'abscission. En effet, les endosomes de recyclage impliqués dans l'étape d'abscission sont en partie issus du compartiment de l'endosome de recyclage Rab11- et Arf6-positif [149] [179]. D'après l'hypothèse émise dans le paragraphe précédent, à savoir que SMAP1 serait la Arf6-GAP impliquée dans la régulation d'Arf6 lors du recyclage rapide en interphase, SMAP1 ne serait pas impliquée dans l'étape d'abscission. Or, la réexpression de SMAP1 dans les cellules HCT116 rallonge la phase G2/M du cycle cellulaire (Figure 6 de l'article), contredisant ainsi notre hypothèse. Les endosomes de recyclage lent ne sont pas les seuls endosomes impliqués dans l'abscission. En effet, les endosomes Rab35-positifs issus de l'endosome de tri ont également un rôle dans la complétion de la cytokinèse [141]. Ainsi, le rétablissement d'un recyclage rapide plus efficace après expression de SMAP1 dans le cas du TfR devrait avoir lieu lors de la cytokinèse et permettre une complétion plus rapide. Cependant, le rétablissement de l'équilibre du cycle GDP/GTP aboutit à la diminution de la quantité d'Arf6-GTP à la MP. Ainsi, une fois les vésicules dans la région péri-membranaire, l'interaction d'Arf6 avec l'exocytose *via* la sous-unité Sec10 est diminuée au même titre que l'interaction avec Rab11-FIP3, aboutissant à une exocytose moindre. De plus, le rétablissement de l'équilibre du cycle GDP/GTP d'Arf6 diminue l'activation de la PIP5K donc la production locale de PIP2, phosphoinositide essentiel à la complétion de la cytokinèse par son rôle

dans le recrutement d'autres protéines à la MP [150] [89]. Ainsi, nos résultats sur la complétion du cycle cellulaire semblent être en accord avec un rééquilibre du cycle GDP/GTP par la réexpression de SMAP1 (Figure 26).

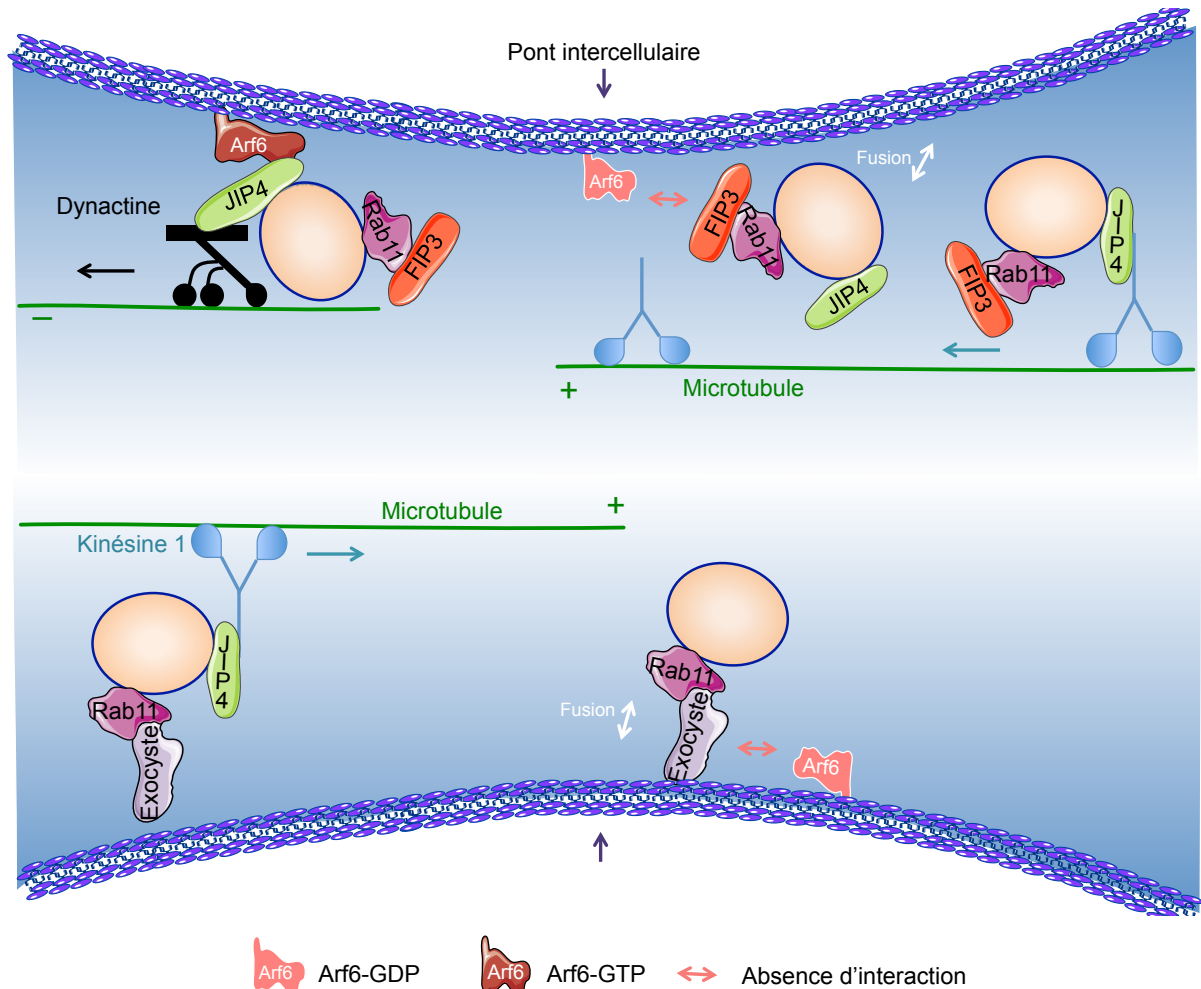


Figure 26: Modèle de l'étape de cytokinèse après réexpression de SMAP1. La protéine Arf6 est inactivée par SMAP1. Ainsi les complexes Rab11/FIP3 et Rab11/Exocyste n'interagissent pas Arf6 diminuant l'exocytose des vésicules cargos avec la membrane plasmique.

Dans le cas de la lignée parentale HCTT16 n'exprimant pas SMAP1, l'augmentation de la forme GTP d'Arf6 n'a pas de conséquence sur le recyclage lent des vésicules cargos ; ainsi, l'apport des endosomes Rab11- et Arf6-positifs à la MP n'est pas perturbé. Arf6-GTP peut se lier à Sec10 et permettre ainsi l'exocytose des vésicules

qui est également facilitée par l'interaction de Rab11 avec une autre sous-unité de l'exocyste, Sec15 [152]. La PIP5K est activée par Arf6-GTP permettant la production du PIP2 nécessaire à l'abscission. Les cellules HCT116 prolifèrent donc plus rapidement que les clones ré-exprimant SMAP1 (Figure 27). Cette observation est cohérente avec la caractéristique clinique des tumeurs MSI qui sont plus volumineuses que les tumeurs MSS.

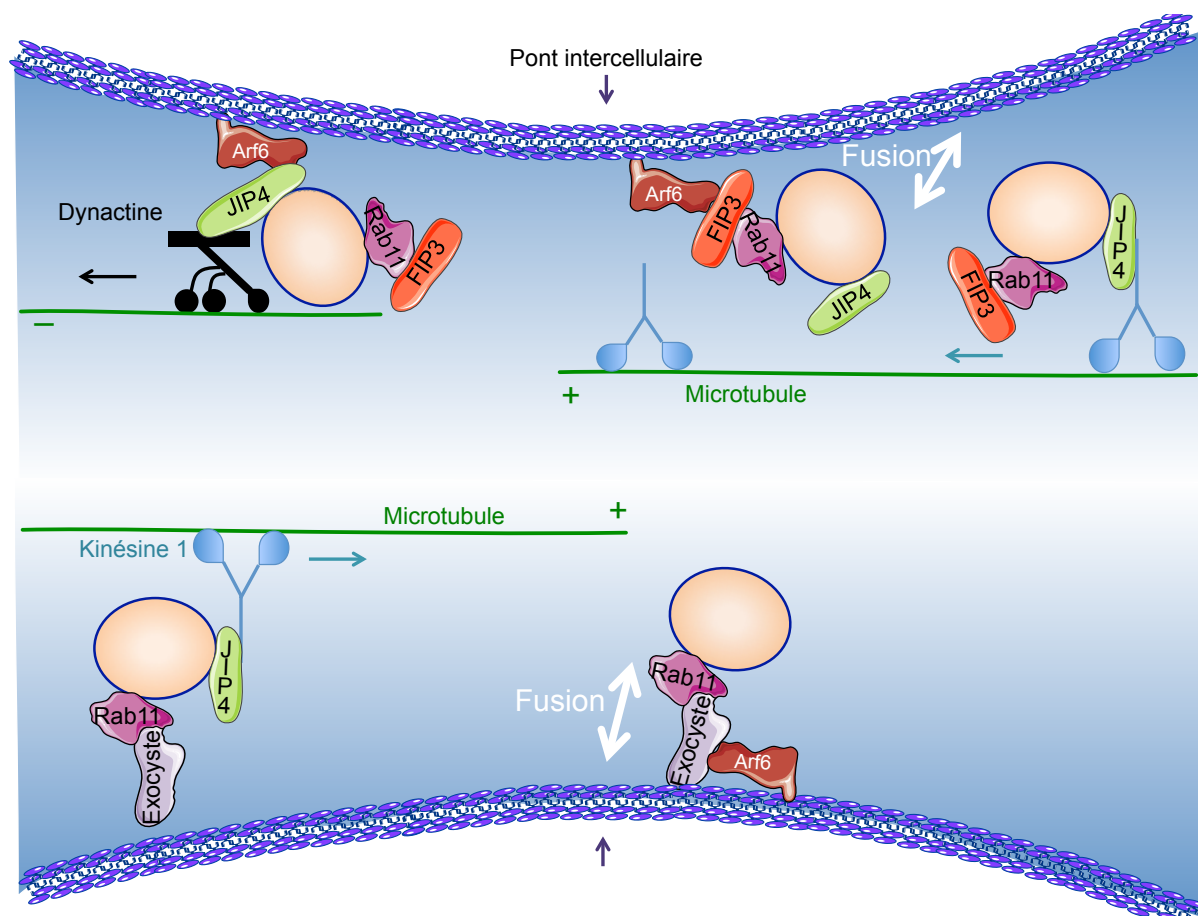


Figure 27: Modèle de la cytokinèse dans les cellules HCT116. Arf6 n'est pas inactivée par l'hydrolyse de son GTP par SMAP1. Ainsi, les interactions des complexes Rab11/FIP3 et Rab11/Exocyste sont augmentées permettant la fusion des vésicules cargos à la membrane plasmique. De plus, l'augmentation du PIP2 membranaire permet également l'accélération de la cytokinèse (non représenté).

4. Le rôle de SMAP1 dans la migration et l'invasion cellulaires

La protéine SMAP1 a été décrite comme étant impliquée dans l'internalisation de la E-cadhérine [79]. Dans cette étude, les auteurs montrent qu'aussi bien la surexpression de SMAP1 que l'inhibition de son activité GAP a pour conséquence un défaut d'internalisation de la E-cadhérine. Nos résultats ne sont pas concordants avec ceux de Kon et al.. En effet, l'expression de SMAP1 dans les cellules HCT116 mène à la relocalisation cytoplasmique de la E-cadhérine et à la diminution de son expression (Figure 21). Cette relocalisation de la E-cadhérine s'accompagne du transfert nucléaire de la β -caténine, la E-cadhérine étant internalisée en complexe avec la β -caténine [165]. La diminution d'expression de la E-cadhérine dans les clones peut être expliquée par l'activation transcriptionnelle des répresseurs SNAIL et SLUG par la β -caténine. Nous montrons que la E-cadhérine semble être normalement internalisée et n'est pas recyclée à la MP. Or, La réexpression de SMAP1 aboutit à la diminution de la quantité d'Arf6-GTP qui est nécessaire à l'internalisation de la E-cadhérine [86] [14]. Par conséquent, dans les clones HCT116-SMAP1, l'internalisation de la E-cadhérine devrait être diminuée, ce que nous n'avons pas observé. Ainsi, comme dans le cas du TfR, la protéine SMAP1 ne semble pas être la Arf6-GAP impliquée dans la régulation du cycle GDP/GTP d'Arf6 lors de l'internalisation de la E-cadhérine. Cependant, dans le cas de la E-cadhérine, la réexpression de SMAP1 aboutit à un défaut de recyclage contrairement au cas du TfR où l'expression de SMAP1 permet de rétablir son recyclage rapide. À la différence du TfR, la E-cadhérine est recyclée à la MP par les endosomes du recyclage lent, Rab11-positifs, et nous avons montré que SMAP1 n'est pas impliqué dans cette phase du recyclage. Ainsi, l'apport des vésicules contenant la E-cadhérine à recycler n'est pas perturbé comme déduit des résultats du cycle

cellulaire. L'effet de l'expression de SMAP1 sur la déstabilisation des jonctions adhérentes peut s'expliquer par la diminution de la quantité d'Arf6-GTP à la MP nécessaire à l'exocytose des vésicules (Figure 28). Ainsi, nos résultats sur la localisation de la E-cadhérine sont en accord avec ceux obtenus sur le recyclage du TfR et sur le cycle cellulaire dans notre système cellulaire et confortent notre hypothèse selon laquelle SMAP1 serait la Arf6-GAP impliquée dans les fonctions de recyclage rapide d'Arf6 et de régulation de l'activation de la PIP5K.

La déstabilisation des jonctions adhérentes dans les clones HCT116-SMAP1 s'accompagne de l'augmentation de leur pouvoir invasif à travers une matrice de Matrigel (Figure 24B). Bien que ce résultat ne soit pas significatif pour les trois clones analysés, la tendance est néanmoins à l'augmentation du pouvoir invasif. Les différences observées entre les clones peuvent être attribuées à la variabilité importante qui peut exister dans les expériences de biologie cellulaire, nécessitant ainsi la répétition des expériences pour pouvoir conclure. Étonnamment, la réexpression de SMAP1 aboutit à la diminution de la migration des cellules (Figure 23) alors que les clones ont des jonctions adhérentes déstabilisées et un pouvoir invasif plus élevé (Figures 21 et 24). Ainsi, il semblerait que les conséquences de l'expression de SMAP1 sur l'invasion cellulaire soient dues à une meilleure digestion de la matrice plutôt qu'à l'augmentation de la migration.

La réexpression de SMAP1 rétablirait l'équilibre du cycle GDP/GTP d'Arf6 mais est responsable de l'entrée en TEM des cellules et de l'augmentation de leur pouvoir invasif. Cette observation est en faveur de l'hypothèse selon laquelle les conséquences des mutations de SMAP1 dans les tumeurs MSI « compenseraient » celles d'un autre défaut, la réexpression de SMAP1 déclenchant alors l'entrée des cellules en TEM et à l'augmentation de leur pouvoir invasif. En effet, les cellules HCT116 présentent une mutation à l'état hétérozygote sur le gène de la β -caténine

qui consiste en la délétion du codon TCT codant la sérine 45, site de phosphorylation de la β -caténine par la GSK-3 β [180]. Cette phosphorylation dirige la β -caténine vers le protéasome afin d'y être dégradée et permet ainsi de réguler l'activité transcriptionnelle de la β -caténine [180]. Le gène du facteur de transcription TCF4 n'étant pas muté dans les cellules HCT116 [177], la mutation de la β -caténine permet d'augmenter l'activité transcriptionnelle du complexe β -caténine/TCF4 [180] et devrait donc augmenter le pouvoir migratoire/invasif de cette lignée cellulaire. Or, les cellules HCT116 sont très peu invasives (Figure 23). Ainsi, il est possible de supposer que la mutation de SMAP1 dans les cellules HCT116 permet de contrecarrer les effets de la mutation de la β -caténine. En effet, nous avons démontré que la mutation de SMAP1 permet de stabiliser les jonctions adhérentes (Figure 21). Ainsi, la majorité de la β -caténine cellulaire est localisée aux jonctions adhérentes, ne pouvant être transférée au noyau et exercer son activité transcriptionnelle. Cette hypothèse est concordante avec le faible pouvoir invasif des tumeurs MSI, une de leurs caractéristiques cliniques.

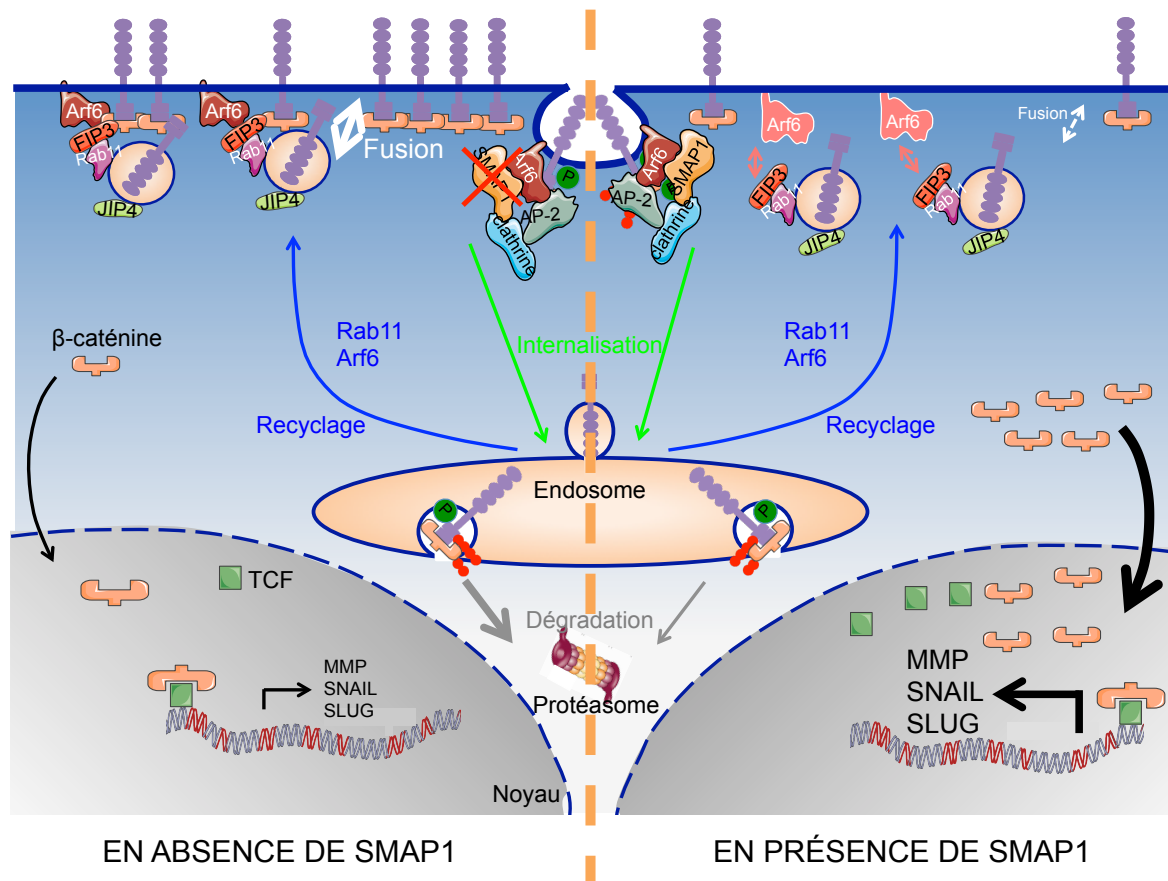


Figure 28: Modèle du rôle de SMAP1 dans la stabilisation des jonctions. Dans les cellules HCT116 (à gauche), la quantité élevée d'Arf6-GTP permet un turnover plus rapide de la E-cadhérine et donc une bonne stabilité des jonctions adhérentes. Arf6-GTP permet également la signalisation de la E-cadhérine et de la β-caténine au protéasome ainsi, une faible quantité de β-caténine est transférée au noyau. Dans les clones HCT116-SMAP1 (à droite), le rééquilibrage du cycle GDP/GTP diminue la quantité d'Arf6-GTP ainsi, l'exocytose de la E-cadhérine est diminuée alors que son internalisation n'est pas modifiée. La signalisation au protéasome est diminuée ainsi, la quantité de β-caténine transférée au noyau augmente aboutissant à l'augmentation de son activité transcriptionnelle.

II. PERSPECTIVES

Mon travail de Thèse a permis de définir le gène SMAP1 comme gène cible de l'oncogenèse MSI dans différents tissus. Les travaux menés *in vitro* dans la lignée de CCR MSI HCT116 ont mis en évidence le rôle de SMAP1 dans la régulation de la fonction d'Arf6 dans le recyclage rapide de protéines membranaires. De plus, les résultats obtenus dans les expériences de cycle cellulaire et de recyclage de la E-cadhérine nous ont permis d'émettre l'hypothèse selon laquelle les mutations de SMAP1 détectées chez les patients atteints d'un CCR MSI ont un effet protecteur. Les différentes hypothèses émises à la suite de mes travaux laissent la porte ouverte à des études plus détaillées afin de les valider.

• **Hypothèse 1 : SMAP1 est impliquée dans la régulation de la fonction d'Arf6 lors du recyclage rapide.**

Le rôle de SMAP1 dans le recyclage rapide sera analysé en étudiant d'autres récepteurs dont le rôle dans l'oncogenèse colorectale a été établi, notamment le récepteur à l'EGF (*Epidermal Growth Factor*, *EGFR*). En effet, selon la nature du ligand, l'EGFR peut être dégradé ou recyclé à la MP par la voie rapide ou lente [181]. De plus, il a été décrit que l'EGFR peut être co-endocyté avec la E-cadhérine. Dans l'optique d'une recherche globale de protéines membranaires dont le recyclage rapide est contrôlé par SMAP1, la technique SILAC (Stable Isotope Labeling by Amino acids in Cell culture) sera utilisée. Cette technique de protéomique en spectrométrie de masse consiste à étudier la différence d'expression de protéines entre deux populations cellulaires dont une est cultivée en présence d'acide aminé

« lourd ». Dans notre système cellulaire, la comparaison du contenu membranaire en protéines sera analysée entre les cellules HCT116 et les clones HCT116-SMAP1.

•Hypothèse 2 : SMAP1 est impliquée dans la régulation de la fonction d'Arf6 lors de l'activation de la PIP5K.

Les effets de la réexpression de SMAP1 sur la cytocinèse et la stabilité des jonctions adhérentes seraient dus au rôle d'Arf6 sur l'activation de la PIP5K. Afin de vérifier cette hypothèse, la quantité de PIP2 aux membranes sera mesurée dans notre système cellulaire. Pour cela, des expériences d'immunofluorescence pourront être réalisées puis les images pourront être quantifiées. Une autre technique est possible, après lyse des cellules, fractionnement des membranes et extraction des lipides. Les lipides seront ensuite déposés sur une membrane de nitrocellulose et le PIP2 sera détecté à l'aide d'un anticorps primaire spécifique et d'un anticorps secondaire couplé à la peroxydase. Ainsi, si l'hypothèse est vérifiée, les clones HCT116-SMAP1 devraient avoir moins de PIP2 à la MP que les cellules parentales

•Hypothèse 3 : SMAP1 est impliquée dans la dégradation de la MEC plutôt que dans la migration.

Cette hypothèse sera testée en étudiant la capacité de digestion de la gélatine fluorescente des cellules exprimant ou non SMAP1. La digestion de la MEC est réalisée par les MMP dont les gènes sont activés par le complexe β -caténine/TCF. Ainsi, la β -caténine étant transférée au noyau après réexpression de SMAP1, le niveau d'expression des MMP2, 7 et 9 secrétées sera analysé ainsi que celui de la MMP14 (ou MT1-MMP) transmembranaire. L'effet transcriptionnel de la β -caténine sera analysé plus largement en utilisant le système TOP-FOP luciférase qui consiste en deux plasmides, le TOP-FLASH qui possède les sites de liaison au TCF dans le

promoteur du gène de la luciférase et le FOP-FLASH, contrôle négatif, qui ne présente aucun site de liaison au TCF dans sa séquence. Ainsi, après transfection transitoire, il est possible de mesurer l'activité de la luciférase et donc indirectement l'activité transcriptionnelle du complexe β -caténine/TCF.

L'invasion des cellules est rendue possible par la formation de protrusions membranaires tels que les invadopodes, riches en filaments d'actine notamment. Le *turnover* de la formation du réseau d'actine est important dans la formation de ces structures. Ainsi, en transfectant les cellules exprimant ou non SMAP1 avec le LifeAct, un plasmide exprimant un peptide de 17 acides aminés fluorescent capable d'interagir avec la F-actine, il sera possible de visualiser en temps réel la formation des filaments d'actine. Il sera également possible d'étudier la formation des invadopodes à travers une matrice de Matrigel en étudiant le réseau d'actine.

Ainsi, une fois le rôle de SMAP1 établi dans l'invasion cellulaire *in cellulo*, une étude *in vivo* pourra être mise en place afin d'analyser le pouvoir invasif des cellules $\pm$ SMAP1 après injection sous-cutanée d'un agglomérat de cellules contenu dans du Matrigel, permettant ainsi de mimer une tumeur solide. L'analyse des éventuelles métastases ou micro-métastases sera d'abord effectuée dans le foie et les poumons, sites privilégiés des cellules tumorales du côlon après xénogreffe. Cependant, une étude systématique de chaque organe devra être réalisée afin de s'assurer qu'il n'y ait pas de métastases ou micro-métastases ailleurs que dans le foie ou les poumons. Ceci pourra être réalisé par PCR en amplifiant les séquences nucléotidiques Alu, absentes du génome murin.

- **Hypothèse 4 : Les mutations de SMAP1 seraient protectrices chez les patients atteints d'un CCR MSI**

Cette hypothèse sera étudiée en analysant un plus grand nombre de CCR MSI afin d'arriver à une puissance statistique plus importante. De plus, nous pourrions étudié l'éventuel rôle prédictif de récives des mutations de SMAP1 chez des patients atteints d'un CCR MSI de stade II, tumeur localisée sans envahissement ganglionnaire, qui ne sont généralement pas traités par chimiothérapie, certaines tumeurs présentant une hétérogénéité de marquage en immunohistochimie.

CONCLUSION GÉNÉRALE

PARTIE IV. CONCLUSION GÉNÉRALE

La compréhension des mécanismes moléculaires menant au cancer colorectal est le meilleur moyen de développer de nouvelles stratégies thérapeutiques qui permettront de diminuer le taux de mortalité lié à ce cancer, troisième cause de mortalité par cancer en France. Lors de ma Thèse, je me suis particulièrement intéressée au CCR MSI qui, comme vous avez pu le constater, représente un phénotype particulier de CCR dont les mécanismes moléculaires néoplasiques, les caractéristiques cliniques et la réponse aux traitements diffèrent des CCR MSS.

De manière originale, nous avons mis en évidence le caractère protecteur des mutations d'un gène cible de l'oncogenèse MSI, le gène SMAP1. Les conséquences fonctionnelles des mutations de SMAP1 *in cellulo* expliquent quelques unes des caractéristiques cliniques des tumeurs MSI, le volume et le faible potentiel invasif des CCR, expliquant en partie le meilleur pronostic de ces tumeurs.

Cette étude a permis d'initier l'analyse fonctionnelle des mutations de SMAP1 dans les CCR MSI et ouvre la porte à de multiples perspectives afin de définir plus précisément les mécanismes moléculaires dans lesquels SMAP1 est impliquée.

ANNEXES

REVIEW

Microsatellite instability in colorectal cancer : from molecular oncogenic mechanisms to clinical implications

Aziz Zaanani, Katy Meunier, Fatiha Sangar, Jean-François Fléjou and Françoise Praz

Cell Oncology

2011

Microsatellite instability in colorectal cancer: from molecular oncogenic mechanisms to clinical implications

Aziz Zaanan · Katy Meunier · Fatiha Sangar ·
Jean-François Fléjou · Françoise Praz

Accepted: 13 March 2011
© International Society for Cellular Oncology 2011

Abstract

Background Microsatellite instability (MSI) constitutes an important oncogenic molecular pathway in colorectal cancer (CRC), representing approximately 15% of all colorectal malignant tumours. In roughly one third of the cases, the underlying DNA mismatch repair (MMR) defect is inherited through the transmission of a mutation in one of the genes involved in MMR, predominantly MSH2 and MLH1, or less frequently, MSH6 or PMS2. In the overwhelming number of sporadic cases, MSI results from epigenetic MLH1 silencing through hypermethylation of its promoter. MMR deficiency

promotes colorectal oncogenesis through the accumulation of numerous mutations in crucial target genes harbouring mononucleotide repeats, notably in those involved in the control of cell proliferation and differentiation, as well as DNA damage signalling and repair.

Design In this review, we describe the molecular aspects of the MMR system and the biological consequences of its defect on the oncogenic process, and we discuss the various experimental systems used to evaluate the efficacy of cytotoxic drugs on MSI colorectal cells lines. There is increasing evidence showing that MSI CRCs differ from all CRCs in terms of prognosis and response to the treatment. We report the clinical studies that have evaluated the prognostic and predictive value of MSI status on clinical outcome in patients treated with various chemotherapy regimens used in the adjuvant setting or for advanced CRCs. **Conclusion** In view of this, the opportunity of a systematic MSI phenotyping in the clinical management of patients with CRC is further discussed.

A. Zaanan · K. Meunier · F. Sangar · J.-F. Fléjou · F. Praz
INSERM, UMR_S 938, Centre de Recherche Saint-Antoine,
F-75012 Paris, France

A. Zaanan · K. Meunier · F. Sangar · J.-F. Fléjou · F. Praz
UPMC Univ Paris 06, UMR_S 938,
Centre de Recherche Saint-Antoine,
F-75012 Paris, France

A. Zaanan
Service d'Oncologie Médicale, Hôpital Saint-Antoine,
Assistance Publique - Hôpitaux de Paris,
75012 Paris, France

K. Meunier
Service de Chirurgie Oncologique, Centre Paul Papin,
Angers, France

J.-F. Fléjou
Service d'Anatomie et Cytologie Pathologique,
Hôpital Saint-Antoine, Assistance Publique - Hôpitaux de Paris,
75012 Paris, France

F. Praz (✉)
Centre de Recherche Saint-Antoine, INSERM UMRS_938-UPMC,
184 rue du Faubourg Saint-Antoine,
75571 Paris Cedex 12, France
e-mail: francoise.praz@upmc.fr

Keywords Biomarker · Chemotherapy · Colorectal cancer ·
Microsatellite instability · Mismatch repair · Prognosis

Abbreviations

5FU	5-fluorouracil
ACVR2	Activin type II receptor
CAPOX	Capecitabine and oxaliplatin
CIMP	CpG island methylator phenotype
CRC	Colorectal cancer
DFS	Disease-free survival
EGFR	Epidermal growth factor receptor
FOLFOX	5-fluorouracil and oxaliplatin
HNPCC	Hereditary non-polyposis colorectal cancer
HR	Hazard ratio
IDL	Insertion-deletion loop

IHC	Immunohistochemistry
mCRC	Metastatic colorectal cancer
MSI	Microsatellite instability
MMR	Mismatch repair
MSS	Microsatellite stable
NER	Nucleotide excision repair
OR	Odds ratio
OS	Overall survival
PFS	Progression-free survival
RFS	Relapse-free survival
RR	Response rate
TGFR2	Transforming growth factor β receptor 2
TIL	Tumour-infiltrating lymphocytes
VEGF	Vascular endothelial growth factor

1 Introduction

The importance of DNA mismatch repair (MMR) system in human oncogenesis clearly emerged in the early 90's by the demonstration that its inactivation owing to germline mutations was the primary cause of hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch syndrome [17, 39, 89, 153, 189]. Colorectal cancers (CRCs) arising in the context of a Lynch syndrome, are mainly due to germline mutations in MSH2 or MLH1, more rarely MSH6 or PMS2, and account for 2–5% of all CRCs. These tumours are characterized by microsatellite instability (MSI) due to the accumulation of insertion/deletion mutations in tandem repeats of short DNA motifs (1–6 bp) called microsatellites. MSI also arises in approximately 12–15% sporadic cases of CRCs through somatic mutations, epigenetic inactivation of MMR, mainly by silencing of MLH1 through promoter hypermethylation. Identifying patients with Lynch syndrome, the most common inherited cause of CRC, is important not only for clinical management, notably because of the higher risk of developing a second primary cancer, but also for their family members some of whom have inherited the MMR gene defect. As MSI CRCs have a better prognosis than microsatellite stable (MSS) CRCs and behave differently in response to various therapeutic strategies including treatment with 5-fluorouracil (5FU), it may now be time to evaluate the relevance of systematic MSI testing rather than confining it to genetic screening [17, 39, 166, 171, 189]. In this review, we will summarize the molecular basis of MMR, the biological and clinical characteristics of MSI tumours, and the methods used to detect MSI. We will further describe the functions of MMR and the consequences of its defect on the response of cell lines to various drugs, notably those given to patients with CRC. Finally, we will report data regarding the impact of MSI on prognosis and clinical outcome of patients with CRC either treated or not by chemotherapy.

2 Post-replicative DNA mismatch repair

The DNA MMR system was originally identified in bacteria, in which its inactivation results in an increased rate of spontaneous mutations [134, 154]. The DNA MMR system has first been described for its role in post-replicative repair of the DNA polymerases errors that have escaped proofreading activity. Replicative polymerases may incorporate the wrong nucleotide creating Watson-Crick base-base mismatches or slip on short repeated sequences such as microsatellites, creating insertion-deletion loops (IDL). If left unrepaired, the daughter strand synthesized during the next round of replication carries a single nucleotide point mutation or a frameshift mutation that usually leads to a premature STOP codon resulting in the synthesis of a truncated protein (Fig. 1).

2.1 Bacterial MMR

In procaryotes, MMR system involves three specialized proteins: MutS, MutL, and MutH [107]. MutS is an ATPase acting as a homodimer that translocates along DNA, recognizes base-base mismatches, as well as small IDL, and promotes DNA loop formation [7, 134]. After loading onto the DNA, MutS undergoes a conformational change, allowing its interaction with MutL in an ATP-dependent manner [62]. MutL, another ATPase acting as a homodimer, couples mismatch recognition to further steps resulting in the removal of a large newly replicated DNA fragment containing the mismatch [12]. MutH is an endonuclease that can nick the newly synthesized nascent strands, which are tagged by the transient absence of methylation at GATC sequences [197].

2.2 Yeast and mammalian MMR

Yeast has six homologs of MutS (MSH1-MSH6), but mammals lack MSH1 homolog, which is involved in the repair of mitochondrial DNA [33]. With the exception of MSH4 and MSH5, which promote crossing-over and chromosome synapsis during meiosis [50, 82, 160], all other MutS homologs participate in mutation avoidance. Four homologs of MutL have been identified both in yeast (MLH1, MLH2, MLH3, PMS1) and mammals (MLH1, MLH3, PMS1, PMS2). These multiple homologs of MutS and MutL play specialized but partly redundant functions in yeast and mammals. MSH2 forms a heterodimer with either MSH6 or MSH3, giving rise to the formation of MutS α and MutS β complexes, respectively, with different relative abilities to bind to base-base mismatches and to IDL [4, 63, 127, 128] (Fig. 1). Studies performed in yeast have shown that MutS α complexes are able to efficiently bind to most base-base mismatches (with the exception of CC mismatches), and to loops of one or two nucleotides,

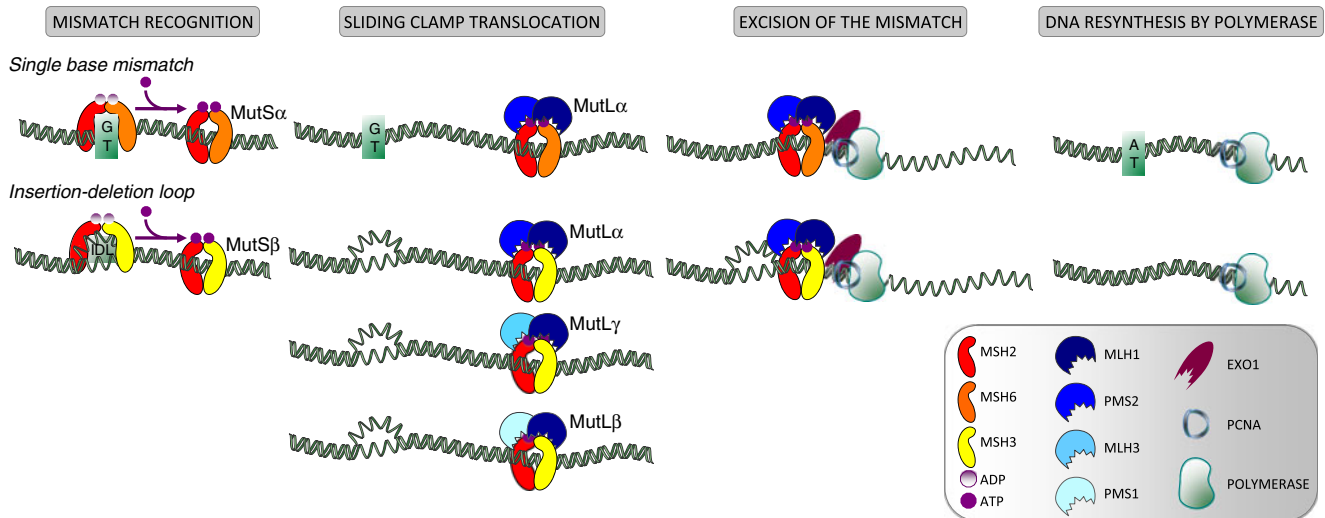


Fig. 1 Eukaryotic DNA mismatch repair. Single nucleotide mismatches, as well as 1 nucleotide loops typically occurring at microsatellite mononucleotide repeated sequences, are recognized by the MutS α heterodimer formed by MSH2 and MSH6; this complex binds to DNA as a sliding clamp after MSH2 ADP has been exchanged for ATP (*upper panel*). The heterodimer MutS β , formed by MLH1 and PMS2, then binds to MutS α . The MutS α -MutS β complex moves along the DNA until it meets the DNA polymerase complex, PCNA and Exo I. The exonuclease then removes up to several hundreds bases from the newly synthesized DNA strand, to allow correct

resynthesis of DNA by the replicative polymerase. IDL of two or more nucleotides are preferentially recognized by the MutS β complex, formed by MSH2 and MSH3 (*lower panel*). After binding to DNA, MutS β exchanges ADP for ATP and recruits the MutL α complex, a heterodimer of MLH1 and PMS2 heterodimer. Subsequently, excision and resynthesis are performed as described for the repair of single base mismatches. MLH1 may also form heterodimers with MLH3 and PMS1 to form respectively the MutL γ and MutL β complexes, but their roles in human MMR function are not entirely clear

whereas MutS β preferentially repairs heteroduplexes with larger IDL [6, 127]. In human cells, the mismatch recognition specificities of MutS α and MutS β are very similar to those of their yeast counterparts. MutS α is present at around ten times higher levels than MutS β and plays the major role, while MutS β is dedicated to the repair of IDL larger than two repeated motifs [4, 63, 126, 144]. Cells deficient in both MSH3 and MSH6 have a high mutation rate, similar to that observed in MSH2-deficient cells, whereas disruption of either gene leads to a weak mutator phenotype, owing to the partially redundant functions of MSH3 and MSH6 [41, 51]. MSH2 is the most abundantly expressed MutS homolog, and its presence is necessary to prevent degradation of MSH6 and MSH3 [31].

Yeast and mammals have four MutL homologs; yeast PMS1, MLH1, MLH2 and MLH3 correspond to human PMS2, MLH1, PMS1 and MLH3, respectively [137]. In humans, MLH1, the major MutL homolog, interacts with its three partners, PMS2, PMS1 and MLH3 through a unique region, to form MutL α , MutL β and MutL γ , respectively [104]. Thanks to its latent endonuclease activity that is activated in the presence of a mismatch, MutS α , RFC (replication factor C), PCNA (proliferating cell nuclear antigen) and ATP, MutL α introduces nicks into DNA to help degradation of the mismatch-containing strand by EXO1 [96]. MutL α is the major component in MMR, operating in conjunction with either MutS α or

MutS β to repair both base-base mismatches and IDL. In vitro studies failed to identify a role for human MutL β formed by MLH1 and PMS1, in postreplicative MMR [155]. Yeast MutL γ , the heterodimer of MLH3 and MLH1, acts in concert with MutS β to repair a proportion of the IDL, decreasing the rate of frameshift mutations [60]. In mammals, MutL γ seems to recognize MutS α substrates, such as G/T mismatch and small IDL, rather than larger loops recognized by MutS β [27]. MLH1 is required to stabilize PMS2 expression in human cells [31], but is dispensable for MLH3 stability, which is generally much less abundant than the two other partners of MLH1 [27]. Thus, MutL γ is unlikely to play a major role in MMR in vivo, but may act as a backup to MutL α in the absence of PMS2 [27].

2.3 Other functions of MMR

In addition to correcting misincorporated nucleotides, MMR proteins have further been implicated in double-strand break repair and recombination [76]. MMR also recognizes mismatches in heteroduplex recombination intermediates preventing completion of recombination between diverged sequences, thus contributing to genetic stability [156, 176]. MMR components are involved in DNA damage response, cell cycle regulation and in the p53-dependent apoptotic DNA damage response to a variety of compounds [19, 93, 141, 175].

3 Colorectal oncogenesis associated with MMR deficiency

3.1 MMR defects involved in hereditary and sporadic cases

The predisposition of some families to gastrointestinal and endometrial cancers was first reported in 1913 by Dr Whartin [194]. In 1966, Dr Lynch studied two large families who developed multiple primary cancers, notably adenocarcinomas of the colon, endometrium and stomach, and demonstrated an autosomal dominant pattern of inheritance [122]. In the early 90's, linkage analysis of large kindreds with CRC, performed to identify regions of loss of heterozygosity in the tumour led to the mapping of two susceptibility loci to chromosome 2p21-22 [1, 139, 148] and 3p21 [139, 140]. Based on the similarities between microsatellite genetic instability and the mutator phenotype displayed by MMR-deficient bacteria and yeast, MSH2 and MLH1, the human homologs of *E. coli* MutS and MutL genes, located on chromosome 2p [59, 110] and 3p [21], respectively, were identified as candidate genes for Lynch syndrome. Mutations of two additional MutL homologs, PMS1 and PMS2, were further reported in patients with HNPCC [138]. The role of PMS1 was later challenged since the mutations occurring in this family was associated with a mutation in MSH2 which co-segregated with MSI CRC, while the PMS1 mutation did not [117]. Germline mutations of MMR genes are predominantly found in MSH2 and MLH1, accounting for approximately 40% and 50% of the families, in whom 200 and 250 pathogenic mutations have been identified, respectively [21, 59, 110, 138, 147].

Lynch syndrome is characterized by an autosomal dominantly inherited predisposition to early onset multiple tumours, most frequently in the colon, rectum, endometrium, stomach, accounting for approximately 2–5% of the total CRCs [74]. Cancer of the stomach, ureter, renal pelvis, small bowel, bile ducts and tumours of the brain are also associated with Lynch syndrome and are preferentially associated with MSH2 inactivation [195]; the lifetime risk of developing one of these extracolonic epithelial-derived tumours remains relatively low (<10%) [185]. The risk of developing endometrial cancer is highest in families with MSH6 mutations (64–71%) compared to patients carrying MSH2 or MLH1 mutation carriers (40–50%) [25, 80, 167]. The mean age at diagnosis for CRC is 43–46 years for MSH2 or MLH1 mutations carriers [15, 80, 147], and 51–57 years for MSH6 mutation carriers [15, 80]. Regarding MLH1 and MSH2 mutation carriers, the lifetime risk for CRC is estimated to be 60–80% [2, 69, 75, 174].

Approximately 10 to 15% of all CRCs display MSI in absence of germline mutations in one of the genes involved in MMR. In the overwhelming majority of cases, MMR

defect results from inactivation of MLH1, due to biallelic methylation of its promoter [38, 81, 98]. Similarly to Lynch syndrome, sporadic MSI tumours are most frequently diploid and patients have a better prognosis than those with MSS tumours. However, whereas patients with Lynch syndrome are younger than those with sporadic CRC, sporadic MSI cases are older. Some studies have suggested that MSI phenotype could be associated with mutations in EGFR signalling pathway, such as PI3K/PTEN or BRAF. Mutations in PIK3CA, which encodes the catalytic subunit of PI3K, have been reported to be more frequent in MSI than in MSS CRCs [3], but these results have not been observed in other series of patients [142, 188]. Nevertheless, PTEN, which is the only tumour suppressor in this pathway, is not only mutated but also epigenetically silenced with higher frequency in MSI than MSS tumours [66]. Regarding the alternative EGFR signalling pathway, BRAF mutations, occurring usually at the V600E hot spot, are frequent in sporadic MLH1-deficient MSI tumours in contrast with hereditary MLH1-deficient tumours [42, 129, 192]. Finding the common V600E BRAF mutation, together with MLH1 promoter methylation, can serve as excluding a diagnosis of Lynch syndrome since BRAF mutation are virtually absent in hereditary cases [47, 121]. BRAF mutations are commonly associated with a group of sporadic carcinomas characterized by the CpG island methylator phenotype (CIMP), including sporadic MSI CRCs due to methylation of the of MLH1 promoter [84, 97, 196]. Although the concept of CIMP is not yet clearly established, CIMP is strongly associated with sporadic MSI CRCs where MLH1 and p16 promoter methylation occurs, and is less frequent with MSS CRCs, in which promoter methylation targets other tumour suppressor genes [43, 67]. The very high frequency of BRAF mutations in colorectal polyps with serrated morphology showing extensive DNA methylation further indicates that sporadic MSI CRCs may originate from sessile serrated adenomas, rather than from conventional adenomas [97].

3.2 Inactivation of target genes through MSI-driven mutations

Microsatellites are ubiquitous in all genomes explored so far and are especially abundant in eukaryote genomes. Simple sequence repeats are not confined to intergenic non-coding repeated sequences but are also present in exonic sequences, most often mononucleotide repeats, in which an insertion or a deletion of a repeated motif leads to frameshift mutations. Accumulation of such mutations in genes harbouring a coding repeat is considered as the major molecular mechanism driving MMR-deficient cells oncogenic transformation. Inactivation of the transforming growth factor β receptor 2 (*TGFBR2*) and the activin type

II receptor (*ACVR2*) genes have been established as playing a key role in MSI-driven colorectal oncogenesis (Fig. 2) [79, 94, 145]. These frameshift mutations cause loss of protein expression and tumour suppressor growth mediated through SMAD signalling [95, 125]. In absence of functional studies, discriminating between the genes that are true targets for colorectal MSI-driven mutagenesis from those whose mutations occur randomly and are left unrepaired due to the absence of MMR is not self-evident. High mutation prevalence in a given gene is generally regarded as strongly indicating its oncogenic role, although mutation rates of microsatellites are extremely variable,

being influenced by flanking sequences, sequence composition, motif unit size, and recombination rate [26, 36, 37]. A bi-allelic inactivation and a role in a growth suppressor pathway are among other widely accepted criteria, but no consensus has been achieved, especially no mutation incidence cut-off value could be defined [18, 48]. To overcome this problem, several statistical analyses based on gene mutation frequency independently of functional considerations have been proposed [49, 199]. A database of human mononucleotide microsatellite mutations and their potential impact to tumorigenesis called SelTarBase for “Selective Targets in Human MSI-H Tumorigenesis

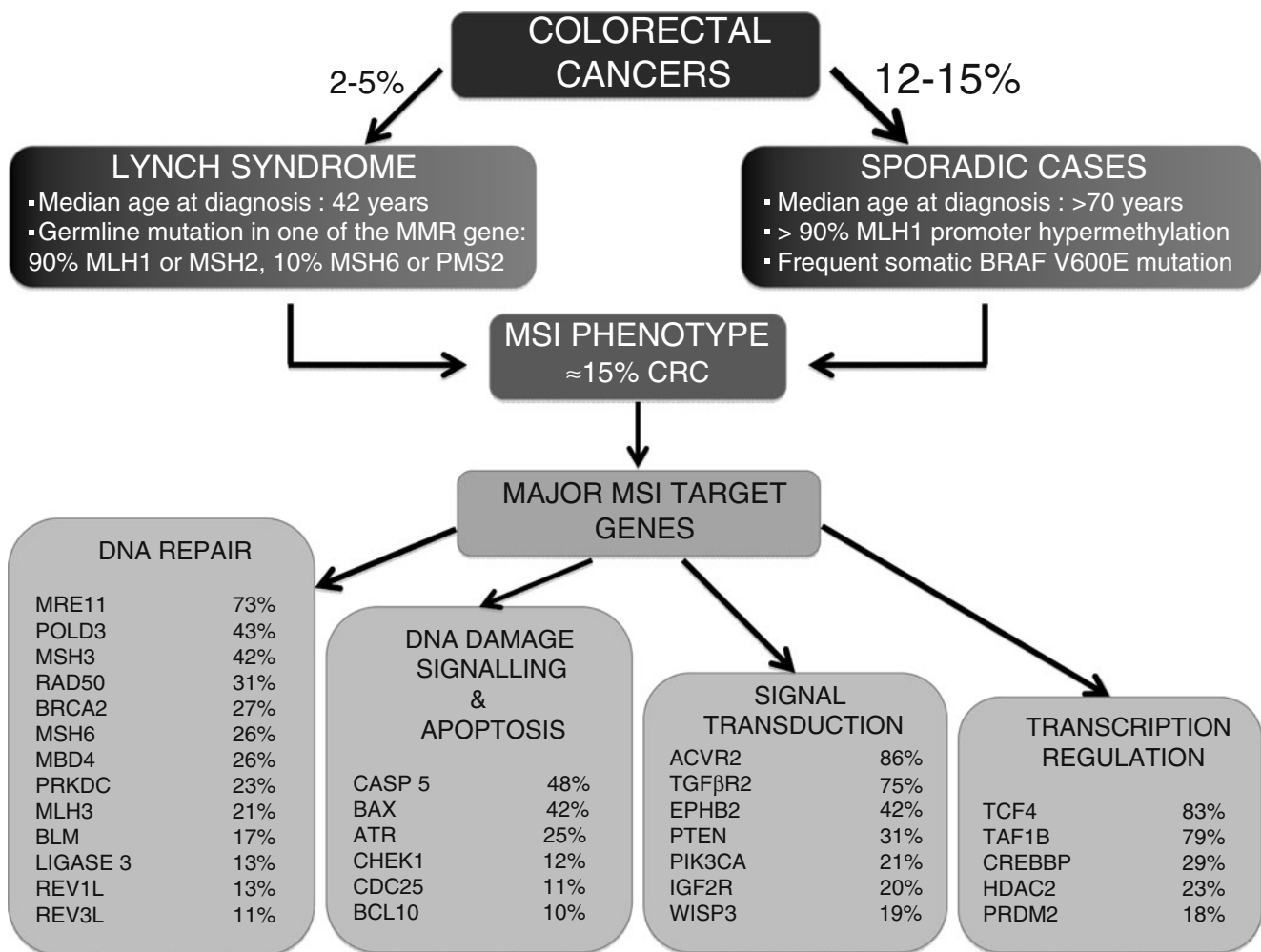


Fig. 2 Molecular pathways involved in the development of MSI CRC. MSI CRCs arise either in a family context known as Lynch syndrome, which accounts for 2–5% of all CRCs or in sporadic cases, representing approximately 12–15% of all CRCs. Hereditary tumours develop at a young age due to the presence of a germline mutation in one of the MMR genes, followed by inactivation of the wild-type allele through gene conversion, loss of chromosome, mutation or methylation (especially for MLH1). Most sporadic cases of CRC displaying MSI are due to methylation-induced silencing of MLH1. Sporadic MSI CRCs are generally diagnosed at an age older than for all CRCs, and frequently carry the common BRAF V600E mutation.

Defects in MMR results in the accumulation of somatic mutations in genes containing microsatellite repeated sequences, referred to as target genes. The identification of the relevant genes among the multitude that carry mutations in a proportion of MSI CRCs remains a challenge for understanding this oncogenic process. Several target genes involved in key cellular functions such as DNA damage signalling and repair, apoptosis, signal transduction and transcription regulation are suspected to play a critical role in tumour initiation and/or progression. Only genes carrying mutations in more than 10% MSI CRCs are shown; mutation frequency data are from the 201006.25, June 2010 SelTarbase release (<http://www.seltarbase.org/>) [200]

Database” was recently built [200]. Today, 74 coding mononucleotide repeats located within the nearly 600 genes analyzed have been reported as predicted targets genes [200]. Classes of genes mutated in MSI tumours include DNA damage signalling and repair, apoptosis, cell growth and differentiation, as well as transcription factors [169, 189, 200]. The 1,291 potentially hypermutable human genes that host a coding mononucleotide repeat longer than seven base pair are implicated in a restricted number of functions, such as cell cycle and maintenance of DNA integrity [119]. In keeping, DNA damage signalling and repair genes are frequently mutated in MSI CRCs [64, 133], endometrial cancers [187], as well as urothelial cell carcinomas of upper urinary tract [135].

4 Experimental models used to evaluate the role of MMR on the cytotoxicity of chemotherapeutic agents

Considering that most cancer chemotherapy agents target DNA and create various types of adducts, the question of whether MMR-deficient tumours behaved differently or not has been largely investigated. Studies have been conducted either with matched, isogenic pair of cell lines, where the cell line differences are supposed to be restricted to the mutated MMR gene or by comparing a panel of MMR-deficient cell lines with a panel of MMR-proficient cell lines [85, 112, 158]. Both approaches have their pitfalls due to intrinsic characteristics of MMR-deficient cells: (1) cells accumulate a great number of mutations that may be selected during oncogenesis or during cell culture, (2) since these mutations occur preferentially in repeated sequences through insertion or deletion of the repeated motif, they are highly dynamic and may revert in a proportion of the tumour cells [83, 132]. Although MMR-deficient cell lines derived from MSI cancers may share a common defect in an MMR gene, they are highly variable with respect to the target gene mutation spectra; MSI cell lines may be seen as “mutation factories”, a characteristic that contributes to some of the discrepancies observed between laboratories. In order to provide useful information regarding the response of MSI tumours to treatments, this approach is reliable only if a reasonable number of cell lines are investigated. This approach is interesting as it recapitulates the MSI pathogenesis rather than investigating the impact of MMR defect *per se*, a situation that is seen only at the very beginning of the disease. This is particularly relevant since many genes involved in DNA damage signalling and/or repair carry microsatellite repeats that are frequently mutated in an MSI background [11, 64, 65, 101, 111, 112, 119, 133]. Furthermore, cell lines selected after introduction of a chromosome or transfection of a cDNA-expressing vector may not be truly isogenic as they undoubtedly differ in the

mutation status of several target genes, a situation that may complicate the interpretation of the data (unpublished personal observations). In order to circumvent the problem caused by the heterogeneity in target gene mutations, an isogenic system, in which MLH1 expression is regulated by doxycycline at the level of the whole cell population has been established, giving the opportunity to properly investigate the role of the MMR defect itself [30]. Finally, experimental models to test the *in vivo* sensitivity have been achieved by xenografting human cancer cells either subcutaneously or orthotopically [20, 151, 158].

4.1 5-fluorouracil

5FU has been the gold standard adjuvant chemotherapy for CRC for many years, and is still used in association with other compounds. Most studies indicate that MSI CRC cell lines grown *in vitro* are more resistant to 5FU and its analogues than MSS cell lines [28, 130]. Interestingly, a direct role for MLH1 could be established in a cell line in which MLH1 deficiency was due to its promoter methylation, by reexpressing MLH1 after addition of the demethylating agent 5-azacytidine [10]. In keeping, the resistance to 5FU of orthotopically xenografted colon cancers was higher for tumours displaying MSI [151].

4.2 Platinum compounds

Cisplatin is a chemotherapeutic agent with widespread-spectrum of action on human malignancies that predominantly forms intrastrand crosslinks between adjacent purine. As these DNA adducts are repaired by the nucleotide excision repair (NER) system, cells deficient in this repair pathway are highly sensitive to cisplatin. In contrast to NER, a deficiency in MMR appears to contribute to increased resistance to cisplatin [5, 57, 58]. The 1,2 diguanyl crosslinks are poorly recognized by MutS α unless a mispaired thymine is incorporated opposite the 3' guanine; thus, crosslinks that have undergone error-prone replication are efficiently recognized by MMR [202]. Cell death results from inefficient cycles of translesion synthesis through cisplatin DNA adducts followed by MMR-mediated removal of the newly synthesized mispaired DNA strand, explaining how loss of MMR confers resistance to cisplatin. Oxaliplatin is a third generation analogue of platinum compound that has shown clinical activity against colon cancer. Oxaliplatin and cisplatin form similar DNA adducts, and modelling studies suggest that both platinum compounds induce similar distortion of secondary DNA structure. However, oxaliplatin-induced adducts differ from those created by cisplatin in that they are poorly recognized by MutS complexes [204]. As expected, loss of MMR activity does not significantly affect oxaliplatin cytotoxicity [181].

4.3 Topoisomerase inhibitors

Topoisomerases are nuclear enzymes that modify the topological state of DNA and participate in replication, transcription, repair, recombination, and chromosome segregation. Topoisomerases catalyze the relaxation of supercoiled DNA through the transient cleavage of DNA strands, creating single or double strand breaks upon action of topoisomerase I or topoisomerase II, respectively. Camptothecin and etoposide are among the many anticancer drugs targeting topoisomerases I and II. Irinotecan, a water-soluble camptothecin derivative has proven its clinical efficacy in CRC. Although not all studies agree [56], MSI cell lines are generally more sensitive to topoisomerase I inhibitors compared with MMR-proficient cell lines [85, 123]. Xenografted human MSI CRCs are also more sensitive to irinotecan but the difference was more modest than observed in vitro [20]. MSI-driven mutations occurring in genes involved double strand break signalling and repair, such as *ATR*, *MRE11*, *RAD50* and *DNA-PKcs* alter efficiency and specificity of the repair process and are most likely responsible for the increased sensitivity of MSI CRCs to drugs generating DNA double strand break [86, 133, 158, 177, 190].

5 MSI phenotyping: who, how and what for?

Lynch syndrome is the most common hereditary colon cancer syndrome. However, whether large-scale screening should be performed or not remains controversial. A number of strategies have been defined to select high-risk subgroups that should benefit MSI screening. Clinical criteria to select families with high-risk CRC were first established in Amsterdam in 1991 by the International Collaborative Group on HNPCC, referred as Amsterdam I criteria, but they were too restrictive to allow effective diagnosis of HNPCC in small families, whereas large families with random clustering of CRCs could be falsely diagnosed as HNPCC [183]. Several guidelines for testing of colorectal tumours for MSI were further developed to increase sensitivity. The Amsterdam II Criteria proposed in 1998 were derived from the Amsterdam I criteria except that not only CRCs, but also extracolonic tumours are considered [186]. However, among families meeting the most stringent Amsterdam criteria I, 30–50% do not have pathogenic germline mutations in MMR genes and lack tumour MSI [116, 118, 149]. The Bethesda Guidelines developed in 1996 take into account the occurrence of extracolonic tumours as well as some characteristic pathology features such as tumour right-sided location and poor differentiation [180].

The methods and criteria used to define MSI in CRCs are still evolving and no consensus has been reached yet.

MSI testing was originally performed on paraffin-embedded tumour tissues using a polymerase chain reaction-based assay to examine genetic alterations at four selected microsatellite loci [179]. A panel of five microsatellites, two mononucleotide and three dinucleotide repeats, was later validated and recommended as a reference panel at the National Cancer Institute workshop held in 1997 [18]. Tumours were classified as high-frequency MSI (MSI-H), if at least two of the five microsatellite markers showed instability, low-frequency MSI (MSI-L) if only one of the five markers showed instability, and microsatellite stable (MSS) if none of the markers showed instability [18]. Since dinucleotide markers are highly polymorphic, the interpretation of the profiles required the comparison with normal DNA from each patient. In order to avoid the need for matching normal DNA, an alternative molecular method based exclusively on quasi-monomorphic mononucleotide markers was developed and proved to be more specific and sensitive than the original National Cancer Institute panel [24]. Detecting MMR protein expression by immunohistochemistry (IHC) represents an alternative method that is widely used. IHC has the advantage of identifying the affected gene by detecting loss of its protein product and to be less expensive and available in numerous pathology departments. MSI testing and IHC are complementary, and loss of MMR protein expression by IHC has been shown to be highly concordant with DNA based MSI testing, especially when MSH6 and PMS2 expression are investigated together with MLH1 and MSH2 [115, 178, 193]. MSI-L and MSS CRCs are generally grouped together because they have similar pathological features and clinical outcome [146].

Because IHC is almost equally sensitive (83%) as the latest molecular MSI technique for tumours with mutations in MLH1 (91%), and MSH2 (87%) or MSH6 (77%), is as specific (90% versus 89%), has the advantage of directing gene mutation screening, and is easy to perform routinely, IHC tends to replace molecular MSI as a screening method for MMR-deficient tumours [39, 184, 185]. MSI tumours tend to share various clinicopathologic characteristics including preferential location in the proximal colon, poor differentiation with a mucinous aspect, a Crohn-like inflammatory reaction and an excess of infiltrating lymphocytes within the tumour that have been used to refine the screening of MSI tumours among familial [74, 87] and/or sporadic cases [68, 159]. Guidelines to detect Lynch syndrome or MSI tumours keep changing as our knowledge improves and should not be seen as definitely established, but evidence accumulates showing that it is time to diagnose MSI tumours in the general population [185]. Indeed, despite their overall better prognosis, MSI tumours do not benefit from 5FU-based adjuvant chemotherapy (see chapter 7.1). Systematic MSI phenotyping would better

guide the choice of chemotherapy, increasing its efficiency while reducing its toxicity. Given that the cost of molecular MSI analysis is 10 times higher than by running IHC and is more time-consuming, IHC staining represents an attractive alternative method provided that it is performed in trained laboratories [143]. IHC screening using a 2-antibody panel, MSH6 and PMS2, has recently been shown to be as specific and efficient as a 4-antibody panel also comprising MSH2 and MLH1 [72, 170]. Loss of MSH2 or MLH1 results in concomitant loss of their respective partner, MSH6 or PMS2, while the reverse is not true. On the basis of these biochemical properties and recent observations, it may be attractive to investigate tumour expression of MSH6 and PMS2 as a first-line screening method, and to restrict MSH2 and MLH1 immunostaining to those displaying a defect in MSH6 or PMS2, in order to define the primary gene defect [72, 170].

6 Impact of microsatellite instability on the prognosis of patients with colorectal cancer

Most studies report that CRCs exhibiting MSI-H phenotype have improved survival compared to MSI-L or MSS tumour, but the prognostic value of MSI status varies significantly between studies. This may be because relatively few studies were population-based or controlled for tumour stage as well as other known confounding factors such as the age. Indeed, when patients with Lynch syndrome, who tend to be younger than average, are excluded from a series of patients, MSI tumours are mainly those arising in sporadic cases; the fact that they are older than the average may contribute to their shorter overall survival (OS). Moreover, patients receiving chemotherapy were most often analysed together with those who were treated by surgery alone. The favourable clinical outcome of patients with MSI-H tumours was already observed when this form of genetic instability was first discovered in CRCs [120, 179]. Subsequently, several retrospective studies either restricted to patients with stages II or III CRCs [34, 73, 109, 201] or including CRCs at all stages [14, 61, 165] found that MSI-H CRCs had a significantly better outcome compared with MSI-L or MSS tumours. In a large study that focused on patients developing CRC before 50 years of age at diagnosis, MSI was found in 17% of the 607 young patients examined and was associated, in a multivariate analysis, with a significant advantage in survival independently of all standard prognostic factors, including tumour stage [70]. In additional multivariate analyses, the survival advantage conferred by the MSI-H phenotype was generally shown to be independent of tumour stage and other clinical and pathologic variables [14, 61, 165]. Most of this risk reduction occurred in

patients with stage III tumours and to a lesser extent to those with stage II tumours [14, 165], while MSI, which is less frequent in stage IV CRCs, appeared to confer little [165] or no [14] prognostic advantage. The better prognosis of patients with stages II and III tumours was observed in series of patients not receiving chemotherapy [14, 34, 109, 201]. Improved prognosis of patients with MSI tumours was further confirmed in a study based on five randomized trials of 5FU-based adjuvant chemotherapy trials in which 570 patients with stages II and III colon cancers were enrolled [157]. Indeed, among patients not receiving chemotherapy, those with MSI-H tumours had a significantly higher rate of five-year disease-free survival (DFS) and longer OS ($P=0.004$) than those with MSI-L or MSS tumours [157]. These observations have recently been confirmed in another study based on five completed randomized clinical trials including patients with pathologically confirmed stage II ($n=218$) or III ($n=239$) colon cancers [166]. In univariate models in patients treated with surgery alone, MMR deficient status was associated with improved DFS ($P=0.03$) and a trend towards improved OS ($P=0.06$) [166]. These trends remained in multivariate models for both DFS and OS but were not significant ($P=0.17$ and $P=0.24$; respectively) [166].

However, some studies failed to demonstrate significant differences in recurrence-free survival (RFS), DFS or OS between patients with MSI-H and MSS/MSI-L, whether analysis was restricted to stages II and III CRCs [99] or to tumours at all stages [13, 108, 162]. One was based on approximately one third of patients participating in a large cooperative group trial conducted from 4 randomized colon cancer adjuvant treatment studies ($n=542$), among whom 173 had not received chemotherapy [99]. The hazard ratio (HR) for RFS when comparing MSI-H to MSS/MSI-L patients was 0.77 (95% CI, 0.40 to 1.48) in the untreated group, indicating that there was a trend towards improved RFS in patients with MSI-H CRC although the difference was not statistically significant [99]. Another study was performed on 416 consecutive patients with stages I to IV CRCs, among whom 202 did not receive chemotherapy. DFS of patients with stage III as well as pooled data of patients with stages II and III was better in patients exhibiting an MSI-H phenotype ($n=35$) than in patients with MSS tumours ($n=218$), but the differences did not reach statistical significance ($P=0.28$) [108]. There was no difference in OS between the subgroups, neither in the study cohort as a whole nor in early stage or locally advanced cancers (stages I to III) [108]. These results further confirmed a previous study based on a series of 181 unselected stages I to IV CRCs, in which survival analysis revealed no statistically significant difference in prognosis between MSI and MSS cases, although a trend towards better survival for MSI cases was observed [162]. The

marginal influence of the MSI status on prognosis observed in these latter studies may be due to unidentified confounding factors in retrospective studies or insufficient sample size to detect moderate effect of MSI [108, 162]. Yet, the prognostic impact of MSI-H on survival has clearly been established in a meta-analysis of 32 studies comprising a total of 7,642 patients with stages I to IV CRCs, including 1,277 MSI cases [152]. The analysis included untreated patients, as well as patients receiving 5FU-based adjuvant chemotherapy. The combined HR estimate for OS associated with MSI-H was 0.65 (95% CI, 0.59 to 0.71); this benefit persisted when restricting analyses to patients with stage II or III disease participating in clinical trial [152]. The association between MSI and favourable prognosis was recently confirmed in another meta-analysis based on pooled data from 31 published studies [71].

The possible molecular mechanism underlying the better outcome of patient with MSI CRC remains largely unknown. A number of MSI tumour characteristics may play a role in the evolution of the disease. MSI tumours generally present with a greater depth of invasion but with a lower stage, as defined by the American Joint Committee on Cancer. In patients having developed MSI CRC, distant metastases were found less frequently in studies focusing on young patients, [70], sporadic CRCs [88] or in a large series of consecutive and unselected patients undergoing surgery for CRC [124]. In addition, aneuploidy, a hallmark of tumours with chromosomal instability, is an independent marker of poor prognosis in patients with CRC [191]. Given that alterations in tumour cell DNA content are inversely associated with MSI-H phenotype, whether both MSI and DNA ploidy contributed to improved clinical outcome was investigated using data obtained from 6 randomized 5FU-based phase III adjuvant studies [173]. MSI-H and DNA ploidy status were found to be independent prognostic variables, with ploidy being the strongest marker; diploidy was associated in better survival in both MSI and MSS patients groups [173]. Finally, tumour-infiltrating lymphocytes (TIL) may also contribute to the survival advantage of patients with MSI CRC by eliciting a protective anti-tumour immune response against tumour-specific carboxy-terminal frameshift neopeptides accumulating as a result of frameshift mutation hotspots in MMR-deficient cells [23, 168, 172]. MSI tumours are characterized by strong tumour lymphocyte infiltration [131], a criterion that has recently been proposed as a marker to more accurately predict MSI in CRCs [68]. TIL in MSI CRCs are activated and cytotoxic containing higher ratios of CD8/CD3 mRNA copy numbers than TIL infiltrating MSS tumours [150]. Interestingly, a recent study reported that CD8⁺ and CD45RO⁺ T cell densities were higher in tumours showing a lymphocytic response or MSI, while FOXP3⁺ was not associated with either of these features

[163]. These observations are in keeping with a recent study reporting that CD8⁺FOXP3⁺ cells display strong immunosuppressive properties in vitro [32].

7 Predictive impact of microsatellite instability for adjuvant chemotherapy

Patients with MSI-H tumours are generally treated according to the same strategy as those with MSI-L or MSS phenotype, with the stage of the tumour being the major determinant for choosing patients who will receive adjuvant chemotherapy. However, evidence accumulates showing that MSI status may also be important in predicting clinical outcome and response to chemotherapy.

7.1 5-fluorouracil-based adjuvant chemotherapy

The predictive role for MMR status has been evaluated in numerous studies investigating the outcome of patients receiving adjuvant 5FU-based therapy or undergoing surgery alone (Table 1). These studies led to contradictory results, with some reporting a benefit from 5FU for patients with MSI tumours, while others concluded that 5FU had no benefit or even a deleterious effect. Various reasons may account for these discrepancies, (1) some studies focused on stage III patients only, while in others, patients with disease at various stages were mixed, (2) in some, patients having rectal carcinomas were excluded, restricting investigations to patients with colon cancer exclusively, (3) multivariate analysis were not always performed. Above all, the patients were generally not allocated to a 5FU-based treatment at random, allowing for selection bias or any of the many other traps associated with nonrandomized comparisons.

Numerous reports concluded that survival of patients with MSI CRC tumours were not improved by 5FU-based chemotherapy [14, 29, 40, 92, 109] (Table 1). Fluorouracil-based adjuvant chemotherapy has even been suspected to be detrimental for patients with MSI colon cancer [157, 166] (Table 1). In a retrospective study of 204 consecutive stages II and III CRC patients, multivariate analysis showed no difference in survival between MSI-H and MSS groups when the analysis did not take into account whether patients were treated or not by 5FU ($P=0.88$) [29]. There was a significant OS benefit among stages II and III patients who were treated with 5FU-based chemotherapy compared to those who were not ($P=0.04$) [29]. When categorized by the MSI status, there was no difference in survival among patients with MSI tumour, irrespective of whether they received 5FU-based chemotherapy ($P=0.52$) [29]. By contrast, there was a significant survival benefit among patients who had MSS or MSI-L tumours and who

Table 1 Impact of microsatellite instability on clinical outcome of adjuvant CRC patients

Reference	Type of study	Number of patients	Tumour stage	Number of MSI tumours (%)	Number of patients receiving adjuvant CT	Comparison	Survival analysis criteria	Results
5-fluorouracil-based adjuvant chemotherapy								
Elsaleh et al., 2000 [52]	R	656	III	56 (8.5)	272	Adj CT: MSS vs MSI	5 year-OS	Longer survival for MSI
Hemminki et al., 2000 [78]	P/NR	95	III	11 (12)	95	Adj CT: MSS vs MSI	3 year-RFS	Longer survival for MSI
Ribic et al., 2003 [157]	RCT	570	II + III	95 (16.7)	283	Adj CT: MSS vs MSI MSI: surgery alone vs adj CT	5 year-DFS 5 year-OS	No benefit or detrimental ^a
Carethers et al., 2004 [29]	R	204	II + III	36 (17.6)	66	MSI: surgery alone vs adj CT	OS	No benefit
de Vos Tot Nederveen Cappel et al., 2004 [40]	R	92	III	92 (100) ^b	28	MSI: surgery alone vs adj CT	5 year-OS	No benefit
Benatti et al., 2005 [14]	R	1263	All stages	256 (20.3)	304	MSI: surgery alone vs adj CT	5 year-OS	No benefit
Westra et al., 2005 [198]	RCT	273	III	44 (16)	273	Adj CT: MSS vs MSI	5 year-DFS	Longer survival for MSI ^c
Jover et al., 2006 [92]	P/NR	754	All stages	66 (8.8)	260	MSI: surgery alone vs adj CT	OS	No benefit
Lanza et al., 2006 [109]	R	718	II + III	114 (15.9)	193	Adj CT: MSS vs MSI MSI: surgery alone vs adj CT	6 year-OS	No benefit
Kim et al., 2007 [99]	RCT	542	II + III	98 (18)	369	Adj CT: MSS vs MSI	5 year-RFS/OS	No significant difference ^d
Lamberti et al., 2007 [108]	P/NR	416	All stages	52 (13)	89 ^e	Adj CT: MSS vs MSI	OS	No significant difference
Sargent et al., 2010 [166]	RCT	457	II + III	70 (15)	229	Adj CT: MSS vs MSI	5 year-DFS	No benefit
	RCT	1027 ^f	II + III	165 (16)	512	MSI: surgery alone vs adj CT	5 year-OS	No benefit or detrimental
Oxaliplatin-containing adjuvant chemotherapy								
Kim et al., 2010 [103]	R	135	All stages	12 (8.8)	121	Adj CT: MSS vs MSI	3 year-DFS 3 year-OS	No significant difference
Zaanan et al., 2010 [203]	R	233	III	32 (14)	5FU, <i>n</i> =124 FOLFOX, <i>n</i> =109	MSI: 5FU vs FOLFOX	3 year-DFS	Longer survival with FOLFOX
Irinotecan-containing adjuvant chemotherapy								
Bertagnolli et al., 2009 [16]	RCT	702	III	96 (13)	5FU, <i>n</i> =348 IFL, <i>n</i> =354	IFL: MSI vs MSS MSI: 5FU vs IFL	5 year-DFS	Longer survival for MSI treated with IFL

Adj adjuvant; CT chemotherapy; FOLFOX 5FU/leucovorin/oxaliplatin; IFL irinotecan/5FU/leucovorin; NR non-randomized; P prospective; R retrospective; RCT randomized controlled trial

^a Trend towards deleterious effect on MSI tumours: 5-yr DFS 70.7 vs 88.0%, *P*=0.07. ^b All patients are from HNPCC families and display MSI tumours. ^c Not significant in multivariate analysis (*P*=0.06).

^d Trend towards benefit for MSI in terms of RFS but results are not significant; identical OS. ^e 89 stage III patients received CT, no information for stage II patients. ^f Pooled data set from patients included in references [157] and [166]. Deleterious effect of 5FU on OS for stage II patients with MSI tumours (*P*=0.04)

received 5FU ($P \leq 0.05$) [29]. Another study including 718 consecutive patients with adjuvant CRC (393 stage II and 325 stage III) showed a significantly better clinical outcome for patients whose tumour lacked expression of either MLH1 or MSH2 [109]. In stage III CRCs, the survival advantage conferred by absence of detectable expression of MMR protein in tumours was more evident among patients undergoing surgery alone than among patients who further received adjuvant 5FU chemotherapy [109]. Similar conclusions were drawn in a prospective study investigating 505 patients with stage II ($n=296$) or stage III ($n=209$) CRCs [90]. Consistent with previous reports, patients with stage II or III CRCs treated with 5FU chemotherapy had better DFS and OS, but the survival benefit of 5FU treatment was again restricted to the group of patients with MMR proficient tumours [91, 92]. In HNPCC families, the 5-year survival of subjects with stage III colon cancer that were treated with ($n=28$) and without ($n=64$) adjuvant 5FU did not differ [40]. Benatti et al. reported that the use of 5FU-based treatment was not a significant factor for survival in a cohort of patients with stage II ($n=491$) or III ($n=461$) CRCs [14]. There was no difference in survival among patients with MSI tumour (22%), irrespective of whether they received 5FU-based chemotherapy, even when adjusting for stage [14]. This also held true for patients with MSS stage II tumours, while those with MSS stage III CRC had a significant survival advantage when receiving 5FU therapy [14].

The impact of MSI status to predict the benefit of 5FU-based adjuvant chemotherapy was further evaluated in a study based on five randomized phase III trials in which 312 patients with stage II and 258 with stage III CRC were enrolled [157]. Among patients who received chemotherapy, those with MSI-H tumours were associated with a slightly lower 5-year survival rate compared to the MSI-L or MSS tumours (70.7% vs 75.5%; $P=0.66$). Chemotherapy with 5FU in patients with MSI-H tumours was associated with a worse outcome (HR for death, 2.14; $P=0.11$). Altogether, these results suggest that 5FU-based chemotherapy does not improve survival of patients with MSI-H CRC and may even be detrimental with a nonsignificant, 2-fold excess in mortality among stage III MSI CRC patients and a 3-fold increase in mortality among stage II MSI CRC patients [157].

In order to provide an independent validation of the findings of Ribic et al., an international collaboration was established to investigate whether patients with MSI tumour benefit from 5FU-based chemotherapy or not. This study was based on the data obtained from 457 patients with stage II or III colon cancer previously enrolled onto five completed, randomized clinical trials of 5FU-based treatment versus no-treatment control [166]. Since the findings obtained with these patients were highly consistent with

those previously obtained by Ribic [157], the two data sets were pooled. In patients not treated with FU-based therapy, deficient MMR status was associated with improved DFS ($P=0.009$) and OS ($P=0.004$). Only patients with stage III disease and MMR-proficient tumours benefited from chemotherapy [166]. There was a statistically significant decreased OS in patients with stage II disease and MMR-deficient tumours who were treated compared with patients who underwent surgery alone ($P=0.04$); in keeping, there was a trend for worse DFS in this group, but the difference was not statistically significant [166]. Neither DFS nor OS was improved by chemotherapy in patients with stage III MMR-deficient tumour [166].

On the other hand, several studies failed to demonstrate a survival difference among patients treated by 5FU between MSI and MSS tumours [99, 108], whereas others reported a greater benefit of adjuvant 5FU chemotherapy for MMR-deficient tumours [52, 53, 78, 198]. One study used data from 391 stage III colon cancer patients who participated in a prospective randomized trial evaluating adjuvant 5FU-based chemotherapy, in which MSI phenotype was determined in 273 patients [198]. In univariate analysis, MSI-H tumours were associated with a longer DFS ($P=0.04$), but in a multivariate model adjusting for nodal involvement, histology, invasion, and grade of tumour, the association of MSI-H status with DFS no longer reached statistical significance ($P=0.06$) [198]. The retrospective series of patients with stage III CRC published by Elsaleh et al. reported that among patients who received 5FU-based chemotherapy, those whose tumour displayed MSI had longer survival [52, 54]. In the earliest report, patients receiving 5FU were on average 13 years younger than those who did not, which may represent an additional bias [52]. Likewise, another study conducted in a group of 95 stage III CRC patients receiving adjuvant 5FU-based chemotherapy after surgery showed that the 3-year RFS was 90% and 43% for MSI and non-MSI tumours ($P=0.02$), respectively [78]. A better clinical outcome of patients with MSI tumour was also reported in a large cooperative group trial conducted from 4 randomized colon cancer treatment studies where 369 patients were treated by 5FU chemotherapy, among whom the 61 patients whose tumour exhibited an MSI-H phenotype had a trend for longer RFS compared to MSS/MSI-L patients, but the differences were not statistically significant [99]. In keeping, Lamberti et al. found no significant difference in OS among patients with stage III CRC treated with 5FU adjuvant chemotherapy when comparing MSI ($n=5$) and MSS ($n=81$) tumours ($P=0.55$) [108].

Finally, a recently published meta-analysis was performed for treated and non-treated MSI population on seven studies representing 3,690 CRC patients (810 stage II and 2,444 stage III), including 454 MSI cases [45]. A total

of 1,444 patients received 5FU-based chemotherapy, whereas 1,518 patients did not. There was no survival difference among MSI patients whether or not they received chemotherapy, whereas MSS patients had a better response to chemotherapy, suggesting that MSI could be considered as a predictive marker of chemoresistance to 5FU [45]. Another recent meta-analysis, based on seven studies stratified by 5FU-based adjuvant chemotherapy of 2,863 stages II and III CRC patients of whom 396 were MSI tumours, reported similar results [71]. The MSS 5FU-treated patients showed a significant better OS as compared to untreated MSS patients (OR=0.52; 95% CI, 0.4 to 0.6; $P<0.0001$). The pooling data of the three randomised studies confirmed the benefit of 5FU treatment for MSS patients (OR=0.62; 95% CI, 0.47 to 0.82; $P=0.003$), while the beneficial effect of 5FU was not obvious for the 214 MSI patients (OR=1.03; 95% CI 0.25 to 4.3; with evidence heterogeneity: $P=0.03$) [71]. Such clinical outcomes are consistent with findings of preclinical data, which showed that 5FU mediated cytotoxicity may be dependent on intact DNA MMR gene function [28].

The lack of benefit from 5FU-based chemotherapy in patients with MMR-deficient tumours justifies to routinely assess MMR status at least for patients being considered for 5FU therapy.

7.2 Oxaliplatin-containing adjuvant chemotherapy

The addition of oxaliplatin to 5FU improves the adjuvant treatment of stage II and III colon cancer by reducing the risk of recurrence and increasing OS [8, 9, 106]. However, adding oxaliplatin did not improve OS in unselected patients with stage II disease [9]. Based on these trials, infusional fluorouracil, leucovorin, and oxaliplatin, known as FOLFOX regimen, is now the current standard adjuvant therapy for stage III colon cancer. Up to now, the predictive impact of MSI status upon the chemosensitivity to FOLFOX remains poorly explored. A recent retrospective study was performed on 135 patients treated by FOLFOX chemotherapy after curative resection for colon adenocarcinoma (13 stage II and 108 stage III: 12 MSI) or with metastases to liver only (14 stage IV: no MSI) [103]. There was no significant difference in terms of DFS and OS according to the MMR status [103] (Table 1). In order to determine the value of MSI phenotype to predict benefit from adjuvant chemotherapy, we recently performed a retrospective study including 233 unselected patients with stage III colon cancer treated by 5FU and leucovorin alone ($n=124$) or with oxaliplatin ($n=109$). In this series, among the patients with stage III colon cancer displaying MSI, those treated by FOLFOX had a significantly longer 3-year DFS than those receiving 5FU chemotherapy ($P=0.01$) [203]. However, because no disease recurrence was observed among patients with MSI tumours treated by FOLFOX,

multivariate analysis to test whether MSI was an independent factor could not be performed [203]. Interestingly, in the subset of patients with MSS tumours, adding oxaliplatin to 5FU remained beneficial, but the improvement in DFS was slight and was not statistically significant ($P=0.15$). Thus, MSI phenotype seems to have a favourable impact on the efficacy of adjuvant treatment by FOLFOX in patients with stage III colon cancer although our results require confirmation in samples from randomized trials. Since future trials comparing FOLFOX to 5FU in adjuvant treatment will probably never be done, prospective validation will not be feasible. In view of this, the effort currently made to analyze materials of previously completed trials such as the MOSAIC study represents a unique chance to translate these observations into clinical application [8].

7.3 Irinotecan-containing adjuvant chemotherapy

Unlike oxaliplatin, irinotecan added to 5FU as adjuvant therapy in patients with stage III colon cancer did not confer a statistically significant improvement in DFS or OS compared with 5FU alone [164, 182]. Bertagnoli et al. evaluated the impact of MSI phenotype in 723 of the 1,264 stage III CRC patients randomly assigned to receive either 5FU-based chemotherapy alone or with irinotecan [16]. Patients with MMR-deficient tumours ($n=96$) treated by 5FU and irinotecan showed an improved 5-year DFS as compared with those whose tumour normally expressed MMR proteins ($P=0.03$) [16]. This relationship was not observed among patients treated with 5FU-based chemotherapy alone. A trend towards longer DFS was observed in patients with MMR-deficient tumours treated by 5FU and irinotecan as compared with those receiving 5FU alone ($P=0.07$) [16]. Loss of tumour MMR function may predict improved outcome in patients treated with the 5FU and irinotecan regimen as compared with those who do not receive irinotecan.

In summary, adding oxaliplatin to 5FU improves clinical outcome for stage III colon cancer; results for stage II are controversial, though high risk stage II colon cancer most probably benefit from the addition of oxaliplatin to 5FU [8, 9, 161]. MSI is an important molecular marker for prognosis and lack of benefit from 5FU adjuvant chemotherapy. Preliminary data indicate that adding either oxaliplatin or irinotecan to 5FU may overcome the poor response observed in patients with stage III MSI colon cancer. Based on these data, patients with stage III colon cancer should receive FOLFOX, whatever the MMR status. For patients having high risk stage II colon cancer, those with MSS tumours could be treated with 5FU or FOLFOX. Besides, in the present state of our knowledge, patients having high risk stage II MSI colon cancer tumours should be given FOLFOX rather than 5FU, if treatment is needed.

8 Predictive impact of microsatellite instability for chemotherapy in metastatic colorectal cancer

Significant developments have been made in chemotherapy regimens for the treatment of metastatic CRC (mCRC) over the past decade with introduction of new cytotoxic drugs including fluoropyrimidine, oxaliplatin, irinotecan and biological agents. Several studies have investigated novel predictive factors that could allow to choose the best individualised therapy but none of these markers has been introduced into routine clinical practice except for *KRAS* mutations that are considered as a predictive marker associated with resistance to anti-EGFR treatment [114, 182]. The predictive value of MSI status on the efficacy of chemotherapy in mCRCs has been investigated in only few studies and one meta-analysis, representing a difficult issue because of the reduced likelihood of metastases at diagnosis (Table 2). Moreover, the chemotherapeutic protocols given to patients with advanced CRC are highly variable, and the parameters to evaluate the impact of MSI differed among the studies being the response rate (RR), progression-free survival (PFS) or OS.

8.1 5-fluorouracil-based chemotherapy

The data regarding the predictive value of MSI-H in the treatment of mCRCs based in 5FU chemotherapy are very

limited. The results from two studies suggested that patients with advanced MSI CRC benefit from 5FU-based chemotherapy [22, 113]. Liang et al. observed in 244 mCRC patients a significantly better outcome among MSI-H mCRCs treated with 5FU/leucovorin in terms of RR (65.7% vs 35.1%; $P=0.001$), and OS (24 vs 13 months; $P=0.0001$), compared to patients with non MSI-H tumours [113]. In their series in which an unusually high proportion of metastatic tumours displayed MSI (>20%), the authors concluded from a multivariate analysis that the better prognosis of stage IV sporadic CRCs with MSI might be associated with better chemosensitivity, rather than lower aggressiveness in biologic behaviour. Similarly, another study including 43 mCRC patients treated in first-line chemotherapy with 5FU and folinic acid showed slightly better RR (72% vs 41%, $P=0.07$) and a significant longer median survival (33 vs 19 months; $P=0.02$) for MSI ($n=7$) patients compared to patients with MSS tumours [22].

8.2 Oxaliplatin-containing chemotherapy

A retrospective study reported the predictive value of MSI in patients treated with palliative first-line combination of 5FU and oxaliplatin (FOLFOX) in 40 patients with mCRC, among whom nine exhibited an MSI phenotype. In this

Table 2 Impact of microsatellite instability in chemotherapy metastatic colorectal cancer

Reference	Type of study	Number of patients	Number of MSI tumours (%)	Protocols of chemotherapy	Outcome analysis criteria	Results
5-fluorouracil-based chemotherapy						
Liang et al., 2002 [113]	NR	244	52 (21)	5FU + leucovorin	RR/OS	Better for MSI
Brueckl et al., 2003 [22]	RCT	43	7 (16)	5FU + folinic acid	RR/OS	Better for MSI
Oxaliplatin-containing chemotherapy						
des Guetz et al., 2007 [44]	R	40	9 (22.5)	FOLFOX	RR/PFS/OS	No significant difference
Müller et al., 2008 [136]	RCT	108	4 (4)	FUFOX or CAPOX	RR/PFS/OS	No significant difference ^a
Chua et al., 2009 [35]	P	118	2 (2)	FOLFOX	RR/PFS/OS	No significant difference
Kim et al., 2009 [102]	R	171	10 (6)	FOLFOX or CAPOX	RR/PFS/OS	No significant difference
Irinotecan-containing chemotherapy						
Fallik et al., 2003 [55]	R	72	7 (9.7)	5FU + irinotecan	RR	Better for MSI
Koopman et al., 2009 [105]	RCT	515	18 (3.5)	Capecitabine or CAPIRI ^b	RR/PFS/OS	No significant difference
Kim et al., 2010 [100]	R	200	23 (11.5)	Irinotecan-based regimen	RR/PFS	No significant difference

CAPIRI capecitabine + irinotecan; CAPOX capecitabine + oxaliplatin; FOLFOX 5FU + leucovorin + oxaliplatin; FUFOX 5FU + oxaliplatin; NR non-randomized; P prospective; R retrospective; RCT randomized controlled trial

^a Only disease control rate (complete/partial remission + stable disease vs progression disease) is significantly better for MSI ($P=0.02$). ^b Patients were randomized between first-line capecitabine, second-line irinotecan, third-line CAPOX (sequential arm) versus first-line CAPIRI, second-line CAPOX (combination arm); RR and survival were analyzed for first-line treatments only

study, in which a frequency of MSI was unexpectedly high (22.5%), there was no significant difference between MSI-H and MSS cases in terms of RR, PFS and OS [44]. Similar conclusions were drawn from a series of 118 mCRC patients who underwent FOLFOX treatment from three successive phase II trials [35]. Of note, the incidence of tumour with MSI-H status was especially low in this study, representing only 2% of patients [35]. Likewise, Kim et al. also reported similar outcomes of patients with MSI and MSS tumours in a study including 171 patients who received capecitabine and oxaliplatin (CAPOX) or FOLFOX in first-line combination chemotherapy for advanced CRC [102]. MMR defect displayed by 6% cases did not significantly influence the RR, PFS or OS in patients with mCRC [102]. Another study investigated the value of MSI-H in 108 patients with mCRC who were enrolled in a prospective, randomised trial comparing two 5FU/oxaliplatin-based first-line chemotherapy (FUFOX vs CAPOX) [136]. MSI-H was found in four cases and was correlated with a lower rate of disease control compared to non-MSI-H patients ($P=0.02$) [136]. However, there was no correlation between MSI-H and RR, PFS or OS [136].

8.3 Irinotecan-containing chemotherapy

Fallik et al. have evaluated the impact of MSI-H phenotype on the tumour response in 72 mCRC patients treated with 5FU and irinotecan after disease had progressed under first-line 5FU-based therapy [55]. In this study, among the seven MSI-H tumours, four responded to chemotherapy, whereas only seven of the 65 MSI-L and MSS tumours did ($P=0.009$) [55]. Thus, MSI appeared to predict the tumour sensitivity to 5FU + irinotecan, but the impact on survival was not investigated [55]. Later, Koopman et al. evaluated the impact of MSI status from 515 mCRC patients who were enrolled in a randomised phase III study (CAIRO study) between first-line capecitabine, second-line irinotecan and third-line capecitabine + oxaliplatin (sequential treatment arm) versus first-line capecitabine + irinotecan and second-line capecitabine + oxaliplatin (combination treatment arm), of whom only 3.5% displayed MSI [105]. For first-line treatments, this study showed no significant difference in terms of objective RR (25% vs 31%; $P=0.63$), median PFS (4 vs 6.9 months; $P=0.28$) and median OS (10.2 vs 17.9 months; $P=0.41$) when comparing patients with MSI versus MSS tumours [105]. At last, Kim et al. recently reported the predictive value of MSI in a series of 200 mCRC patients receiving irinotecan-containing regimen as their first-line chemotherapy [100]. MMR deficiency was found in 23 patients (11.5%). The objective RR was 47% in MSS and 56.5% in MSI tumours ($P=0.11$). The median

PFS tended to be longer with 8.8 months in patients with MSI tumours compared to 6.8 months in patients with MSS tumours, but the difference was not statistically significant ($P=0.09$) [100].

In summary, these studies suggest that MSI phenotype may have a predictive value of better clinical outcomes for patients with mCRC treated by 5FU alone [22, 113], but not for patients treated by combination chemotherapy as 5FU with oxaliplatin or irinotecan. These observations are surprising in light of the lack of benefit survival from adjuvant 5FU chemotherapy in patients with MSI tumours. Nevertheless, these results should be interpreted with caution because of the absence of randomized studies, small number of patients in these studies and the highly variable incidence of MSI in mCRCs. In 2009, des Guetz et al. have investigated the predictive value of MSI status on the effect of chemotherapy given to patients with a mCRC by a meta-analysis study. Statistical calculations were performed on six studies representing 964 patients, with 287 patients receiving 5FU-based chemotherapy, and 678 patients receiving combinations of 5FU or capecitabine with oxaliplatin and/or irinotecan [46]. In this meta-analysis, MSI status did not predict the effect of chemotherapy in terms of RR, which was similar in MSI and MSS mCRC tumours, but the impact on survival was not investigated [46]. To our knowledge, there is no study evaluating the predictive value of MSI phenotype in patients treated by chemotherapy combined with targeted therapies such as monoclonal antibodies against vascular endothelial growth factor (VEGF) or epidermal growth factor receptor (EGFR).

9 General conclusions and perspectives

This review summarizes the increasing evidence showing that MSI CRCs differ from all CRCs in terms of prognosis and response to various treatments. Based on these data, we believe that it is now time to consider performing large scale MSI phenotyping to improve the clinical management of patients with CRC. In this perspective, it becomes essential to define a consensus method for the screening of MSI tumours that is reliable, inexpensive and robust enough to be used in routine laboratories, a goal that has not been achieved yet. In keeping, MSI phenotyping should be systematically investigated in future clinical trials as it represents a powerful biomarker to identify the appropriate personalized treatment regimen based on tumour molecular characteristics. Identifying genes and pathways that are altered preferentially in MSI CRCs should be pursued as it represents a promising opportunity for designing targeted treatments.

References

1. L.A. Aaltonen, P. Peltomäki, F.S. Leach, P. Sistonen, L. Pylkkanen, J.P. Mecklin, H. Jarvinen, S.M. Powell, J. Jen, S.R. Hamilton et al., Clues to the pathogenesis of familial colorectal cancer. *Science* **260**, 812–816 (1993)
2. M. Aarnio, J.P. Mecklin, L.A. Aaltonen, M. Nystrom-Lahti, H.J. Jarvinen, Life-time risk of different cancers in hereditary non-polyposis colorectal cancer (HNPCC) syndrome. *Int. J. Cancer* **64**, 430–433 (1995)
3. J. Abubaker, P. Bavi, S. Al-Harbi, M. Ibrahim, A.K. Siraj, N. Al-Sanea, A. Abduljabbar, L.H. Ashari, S. Alhomoud, F. Al-Dayel, S. Uddin, K.S. Al-Kuraya, Clinicopathological analysis of colorectal cancers with PIK3CA mutations in Middle Eastern population. *Oncogene* **27**, 3539–3545 (2008)
4. S. Acharya, T. Wilson, S. Gradia, M.F. Kane, S. Guerrette, G.T. Marsischky, R. Kolodner, R. Fishel, hMSH2 forms specific mismatch-binding complexes with hMSH3 and hMSH6. *Proc. Natl Acad. Sci. USA* **93**, 13629–13634 (1996)
5. S. Aebi, B. Kurdi-Haidar, R. Gordon, B. Cenni, H. Zheng, D. Fink, R.D. Christen, C.R. Boland, M. Koi, R. Fishel, S.B. Howell, Loss of DNA mismatch repair in acquired resistance to cisplatin. *Cancer Res.* **56**, 3087–3090 (1996)
6. E. Alani, The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol. Cell. Biol.* **16**, 5604–5615 (1996)
7. D.J. Allen, A. Makhov, M. Grilley, J. Taylor, R. Thresher, P. Modrich, J.D. Griffith, MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.* **16**, 4467–4476 (1997)
8. T. Andre, C. Boni, L. Mounedji-Boudiaf, M. Navarro, J. Tabernero, T. Hickish, C. Topham, M. Zaninelli, P. Clingan, J. Bridgewater, I. Tabah-Fisch, A. de Gramont, Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N. Engl. J. Med.* **350**, 2343–2351 (2004)
9. T. Andre, C. Boni, M. Navarro, J. Tabernero, T. Hickish, C. Topham, A. Bonetti, P. Clingan, J. Bridgewater, F. Rivera, A. de Gramont, Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J. Clin. Oncol.* **27**, 3109–3116 (2009)
10. C.N. Arnold, A. Goel, C.R. Boland, Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines. *Int. J. Cancer* **106**, 66–73 (2003)
11. S. Bader, M. Walker, B. Hendrich, A. Bird, C. Bird, M. Hooper, A. Wyllie, Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. *Oncogene* **18**, 8044–8047 (1999)
12. C. Ban, M. Junop, W. Yang, Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell* **97**, 85–97 (1999)
13. R.A. Barnetson, A. Tenesa, S.M. Farrington, I.D. Nicholl, R. Cetnarskyj, M.E. Porteous, H. Campbell, M.G. Dunlop, Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer. *N. Engl. J. Med.* **354**, 2751–2763 (2006)
14. P. Benatti, R. Gafa, D. Barana, M. Marino, A. Scarselli, M. Pedroni, I. Maestri, L. Guerzoni, L. Roncucci, M. Menigatti, B. Roncari, S. Maffei, G. Rossi, G. Ponti, A. Santini, L. Losi, C. Di Gregorio, C. Oliani, M. Ponz de Leon, G. Lanza, Microsatellite instability and colorectal cancer prognosis. *Clin. Cancer Res.* **11**, 8332–8340 (2005)
15. M.J. Berends, Y. Wu, R.H. Sijmons, R.G. Mensink, T. van der Sluis, J.M. Hordijk-Hos, E.G. de Vries, H. Hollema, A. Karrenbeld, C.H. Buys, A.G. van der Zee, R.M. Hofstra, J.H. Kleibeuker, Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. *Am. J. Hum. Genet.* **70**, 26–37 (2002)
16. M.M. Bertagnolli, D. Niedzwiecki, C.C. Compton, H.P. Hahn, M. Hall, B. Damas, S.D. Jewell, R.J. Mayer, R.M. Goldberg, L. B. Saltz, R.S. Warren, M. Redston, Microsatellite instability predicts improved response to adjuvant therapy with irinotecan, fluorouracil, and leucovorin in stage III colon cancer: Cancer and Leukemia Group B Protocol 89803. *J. Clin. Oncol.* **27**, 1814–1821 (2009)
17. C.R. Boland, A. Goel, Microsatellite instability in colorectal cancer. *Gastroenterology* **138**, 2073–2087 (2010). e2073
18. C.R. Boland, S.N. Thibodeau, S.R. Hamilton, D. Sidransky, J.R. Eshleman, R.W. Burt, S.J. Meltzer, M.A. Rodriguez-Bigas, R. Fodde, G.N. Ranzani, S. Srivastava, 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.* **58**, 5248–5257 (1998)
19. P. Branch, G. Aquilina, M. Bignami, P. Karran, Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage. *Nature* **362**, 652–654 (1993)
20. R.A. Bras-Goncalves, C. Rosty, P. Laurent-Puig, P. Soulie, B. Dutrillaux, M.F. Poupon, Sensitivity to CPT-11 of xenografted human colorectal cancers as a function of microsatellite instability and p53 status. *Br. J. Cancer* **82**, 913–923 (2000)
21. C.E. Bronner, S.M. Baker, P.T. Morrison, G. Warren, L.G. Smith, M.K. Lescoe, M. Kane, C. Earabino, J. Lipford, A. Lindblom et al., Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* **368**, 258–261 (1994)
22. W.M. Brueckl, C. Moesch, T. Brabletz, C. Koebeck, C. Riedel, A. Jung, S. Merkel, S. Schaber, F. Boxberger, T. Kirchner, W. Hohenberger, E.G. Hahn, A. Wein, Relationship between microsatellite instability, response and survival in palliative patients with colorectal cancer undergoing first-line chemotherapy. *Anticancer Res.* **23**, 1773–1777 (2003)
23. A. Buckowitz, H.P. Knaebel, A. Benner, H. Blaker, J. Gebert, P. Kienle, M. von Knebel Doeberitz, M. Kloor, Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. *Br. J. Cancer* **92**, 1746–1753 (2005)
24. O. Buhard, F. Cattaneo, Y.F. Wong, S.F. Yim, E. Friedman, J.F. Flejou, A. Duval, R. Hamelin, Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J. Clin. Oncol.* **24**, 241–251 (2006)
25. B.M. Buttin, M.A. Powell, D.G. Mutch, S.A. Babb, P.C. Huettner, T.B. Edmonston, T.J. Herzog, J.S. Rader, R.K. Gibb, A.J. Whelan, P.J. Goodfellow, Penetrance and expressivity of MSH6 germline mutations in seven kindreds not ascertained by family history. *Am. J. Hum. Genet.* **74**, 1262–1269 (2004)
26. C. Campregher, T. Scharl, M. Nemeth, C. Honeder, T. Jascur, C. R. Boland, C. Gasche, The nucleotide composition of microsatellites impacts both replication fidelity and mismatch repair in human colorectal cells. *Hum. Mol. Genet.* **19**, 2648–2657 (2010)
27. E. Cannavo, G. Marra, J. Sabates-Bellver, M. Menigatti, S.M. Lipkin, F. Fischer, P. Cejka, J. Jiricny, Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res.* **65**, 10759–10766 (2005)
28. J.M. Carethers, D.P. Chauhan, D. Fink, S. Nebel, R.S. Bresalier, S.B. Howell, C.R. Boland, Mismatch repair proficiency and in vitro response to 5-fluorouracil. *Gastroenterology* **117**, 123–131 (1999)

29. J.M. Carethers, E.J. Smith, C.A. Behling, L. Nguyen, A. Tajima, R.T. Doctolero, B.L. Cabrera, A. Goel, C.A. Arnold, K. Miyai, C.R. Boland, Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer. *Gastroenterology* **126**, 394–401 (2004)
30. P. Cejka, L. Stojic, N. Mojas, A.M. Russell, K. Heinimann, E. Cannavo, M. di Pietro, G. Marra, J. Jiricny, Methylation-induced G(2)/M arrest requires a full complement of the mismatch repair protein hMLH1. *EMBO J.* **22**, 2245–2254 (2003)
31. D.K. Chang, L. Ricciardiello, A. Goel, C.L. Chang, C.R. Boland, Steady-state regulation of the human DNA mismatch repair system. *J. Biol. Chem.* **275**, 18424–18431 (2000)
32. N. Chaput, S. Louafi, A. Bardier, F. Charlotte, J.C. Vaillant, F. Menegaux, M. Rosenzweig, F. Lemoine, D. Klatzmann, J. Taieb, Identification of CD8+CD25+Foxp3+ suppressive T cells in colorectal cancer tissue. *Gut* **58**, 520–529 (2009)
33. N.W. Chi, R.D. Kolodner, Purification and characterization of MSH1, a yeast mitochondrial protein that binds to DNA mismatches. *J. Biol. Chem.* **269**, 29984–29992 (1994)
34. S.W. Choi, K.J. Lee, Y.A. Bae, K.O. Min, M.S. Kwon, K.M. Kim, M.G. Rhyu, Genetic classification of colorectal cancer based on chromosomal loss and microsatellite instability predicts survival. *Clin. Cancer Res.* **8**, 2311–2322 (2002)
35. W. Chua, D. Goldstein, C.K. Lee, H. Dhillon, M. Michael, P. Mitchell, S.J. Clarke, B. Iacopetta, Molecular markers of response and toxicity to FOLFOX chemotherapy in metastatic colorectal cancer. *Br. J. Cancer* **101**, 998–1004 (2009)
36. H. Chung, C.G. Lopez, J. Holmstrom, D.J. Young, J.F. Lai, D. Ream-Robinson, J.M. Carethers, Both microsatellite length and sequence context determine frameshift mutation rates in defective DNA mismatch repair. *Hum. Mol. Genet.* **19**, 2638–2647 (2010)
37. H. Chung, C.G. Lopez, D.J. Young, J.F. Lai, J. Holmstrom, D. Ream-Robinson, B.L. Cabrera, J.M. Carethers, Flanking sequence specificity determines coding microsatellite heteroduplex and mutation rates with defective DNA mismatch repair (MMR). *Oncogene* **29**, 2172–2180 (2010)
38. J.M. Cunningham, E.R. Christensen, D.J. Tester, C.Y. Kim, P.C. Roche, L.J. Burgart, S.N. Thibodeau, Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res.* **58**, 3455–3460 (1998)
39. A. de la Chapelle, H. Hampel, Clinical relevance of microsatellite instability in colorectal cancer. *J. Clin. Oncol.* **28**, 3380–3387 (2010)
40. W.H. de Vos Test, Survival after adjuvant 5-FU treatment for stage III colon cancer in hereditary nonpolyposis colorectal cancer. *Int. J. Cancer* **109**, 468–471 (2004)
41. N. de Wind, M. Dekker, N. Claij, L. Jansen, Y. van Klink, M. Radman, G. Riggins, M. van der Valk, K. van't Wout, H. te Riele, HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nat. Genet.* **23**, 359–362 (1999)
42. G. Deng, I. Bell, S. Crawley, J. Gum, J.P. Terdiman, B.A. Allen, B. Truta, M.H. Sleisenger, Y.S. Kim, BRAF mutation is frequently present in sporadic colorectal cancer with methylated hMLH1, but not in hereditary nonpolyposis colorectal cancer. *Clin. Cancer Res.* **10**, 191–195 (2004)
43. S. Derks, C. Postma, B. Carvalho, S.M. van den Bosch, P.T. Moerkerk, J.G. Herman, M.P. Weijenberg, A.P. de Bruine, G.A. Meijer, M. van Engeland, Integrated analysis of chromosomal, microsatellite and epigenetic instability in colorectal cancer identifies specific associations between promoter methylation of pivotal tumour suppressor and DNA repair genes and specific chromosomal alterations. *Carcinogenesis* **29**, 434–439 (2008)
44. G. des Guetz, P. Mariani, J. Cucherousset, M. Benamoun, C. Lagorce, X. Sastre, P. Le Toulmeline, B. Uzzan, G.Y. Perret, J.F. Morere, J.L. Breau, R. Fagard, P.O. Schischmanoff, Microsatellite instability and sensitivity to FOLFOX treatment in metastatic colorectal cancer. *Anticancer Res.* **27**(2007), 2715–2719 (2007)
45. G. des Guetz, O. Schischmanoff, P. Nicolas, G.Y. Perret, J.F. Morere, B. Uzzan, Does microsatellite instability predict the efficacy of adjuvant chemotherapy in colorectal cancer? A systematic review with meta-analysis. *Eur. J. Cancer* **45**, 1890–1896 (2009)
46. G. des Guetz, B. Uzzan, P. Nicolas, O. Schischmanoff, G.Y. Perret, J.F. Morere, Microsatellite instability does not predict the efficacy of chemotherapy in metastatic colorectal cancer. A systematic review and meta-analysis. *Anticancer Res.* **29**, 1615–1620 (2009)
47. E. Domingo, P. Laiho, M. Ollikainen, M. Pinto, L. Wang, A.J. French, J. Westra, T. Frebourg, E. Espin, M. Armengol, R. Hamelin, H. Yamamoto, R.M. Hofstra, R. Seruca, A. Lindblom, P. Peltomaki, S.N. Thibodeau, L.A. Aaltonen, S. Schwartz Jr., BRAF screening as a low-cost effective strategy for simplifying HNPCC genetic testing. *J. Med. Genet.* **41**, 664–668 (2004)
48. A. Duval, R. Hamelin, Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res.* **62**, 2447–2454 (2002)
49. A. Duval, S. Rolland, A. Compoin, E. Tubacher, B. Iacopetta, G. Thomas, R. Hamelin, Evolution of instability at coding and non-coding repeat sequences in human MSI-H colorectal cancers. *Hum. Mol. Genet.* **10**, 513–518 (2001)
50. W. Edelmann, P.E. Cohen, B. Kneitz, N. Winand, M. Lia, J. Heyer, R. Kolodner, J.W. Pollard, R. Kucherlapati, Mammalian MutS homologue 5 is required for chromosome pairing in meiosis. *Nat. Genet.* **21**, 123–127 (1999)
51. W. Edelmann, A. Umar, K. Yang, J. Heyer, M. Kucherlapati, M. Lia, B. Kneitz, E. Avdievich, K. Fan, E. Wong, G. Crouse, T. Kunkel, M. Lipkin, R.D. Kolodner, R. Kucherlapati, The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression. *Cancer Res.* **60**, 803–807 (2000)
52. H. Elsaleh, D. Joseph, F. Grieu, N. Zeps, N. Spry, B. Iacopetta, Association of tumour site and sex with survival benefit from adjuvant chemotherapy in colorectal cancer. *Lancet* **355**, 1745–1750 (2000)
53. H. Elsaleh, B. Powell, K. McCaul, F. Grieu, R. Grant, D. Joseph, B. Iacopetta, P53 alteration and microsatellite instability have predictive value for survival benefit from chemotherapy in stage III colorectal carcinoma. *Clin. Cancer Res.* **7**, 1343–1349 (2001)
54. H. Elsaleh, B. Shannon, B. Iacopetta, Microsatellite instability as a molecular marker for very good survival in colorectal cancer patients receiving adjuvant chemotherapy. *Gastroenterology* **120**, 1309–1310 (2001)
55. D. Fallik, F. Borriani, V. Boige, J. Viguier, S. Jacob, C. Miquel, J. C. Sabourin, M. Ducreux, F. Praz, Microsatellite instability is a predictive factor of the tumor response to irinotecan in patients with advanced colorectal cancer. *Cancer Res.* **63**, 5738–5744 (2003)
56. A. Fedier, V.A. Schwarz, H. Walt, R.D. Carpini, U. Haller, D. Fink, Resistance to topoisomerase poisons due to loss of DNA mismatch repair. *Int. J. Cancer* **93**, 571–576 (2001)
57. D. Fink, S. Nebel, S. Aebi, H. Zheng, B. Cenni, A. Nehme, R.D. Christen, S.B. Howell, The role of DNA mismatch repair in platinum drug resistance. *Cancer Res.* **56**, 4881–4886 (1996)
58. D. Fink, H. Zheng, S. Nebel, P.S. Norris, S. Aebi, T.P. Lin, A. Nehme, R.D. Christen, M. Haas, C.L. MacLeod, S.B. Howell, In vitro and in vivo resistance to cisplatin in cells that have lost DNA mismatch repair. *Cancer Res.* **57**, 1841–1845 (1997)
59. R. Fishel, M.K. Lescoe, M.R. Rao, N.G. Copeland, N.A. Jenkins, J. Garber, M. Kane, R. Kolodner, The human mutator

- gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027–1038 (1993)
60. H. Flores-Rozas, R.D. Kolodner, The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proc. Natl Acad. Sci. USA* **95**, 12404–12409 (1998)
61. R. Gafa, I. Maestri, M. Matteuzzi, A. Santini, S. Ferretti, L. Cavazzini, G. Lanza, Sporadic colorectal adenocarcinomas with high-frequency microsatellite instability. *Cancer* **89**, 2025–2037 (2000)
62. L. Galio, C. Bouquet, P. Brooks, ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. *Nucleic Acids Res.* **27**, 2325–2331 (1999)
63. J. Genschel, S.J. Littman, J.T. Drummond, P. Modrich, Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J. Biol. Chem.* **273**, 19895–19901 (1998)
64. G. Giannini, C. Rinaldi, E. Ristori, M.I. Ambrosini, F. Cerignoli, A. Viel, E. Bidoli, S. Berni, G. D'Amati, G. Scambia, L. Frati, I. Screpanti, A. Gulino, Mutations of an intronic repeat induce impaired MRE11 expression in primary human cancer with microsatellite instability. *Oncogene* **23**, 2640–2647 (2004)
65. G. Giannini, E. Ristori, F. Cerignoli, C. Rinaldi, M. Zani, A. Viel, L. Ottini, M. Crescenzi, S. Martinotti, M. Bignami, L. Frati, I. Screpanti, A. Gulino, Human MRE11 is inactivated in mismatch repair-deficient cancers. *EMBO Rep.* **3**, 248–254 (2002)
66. A. Goel, C.N. Arnold, D. Niedzwiecki, J.M. Carethers, J.M. Dowell, L. Wasserman, C. Compton, R.J. Mayer, M.M. Bertagnolli, C.R. Boland, Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers. *Cancer Res.* **64**, 3014–3021 (2004)
67. A. Goel, T. Nagasaka, C.N. Arnold, T. Inoue, C. Hamilton, D. Niedzwiecki, C. Compton, R.J. Mayer, R. Goldberg, M.M. Bertagnolli, C.R. Boland, The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer. *Gastroenterology* **132**, 127–138 (2007)
68. J.K. Greenon, S.C. Huang, C. Herron, V. Moreno, J.D. Bonner, L.P. Tomsho, O. Ben-Izhak, H.I. Cohen, P. Trougouboff, J. Bejhar, Y. Sova, M. Pinchev, G. Rennert, S.B. Gruber, Pathologic predictors of microsatellite instability in colorectal cancer. *Am. J. Surg. Pathol.* **33**, 126–133 (2009)
69. S. Grover, E.M. Stoffel, R.C. Mercado, B.M. Ford, W.K. Kohlman, K.M. Shannon, P.G. Conrad, A.M. Blanco, J.P. Terdiman, S.B. Gruber, D.C. Chung, S. Syngal, Colorectal cancer risk perception on the basis of genetic test results in individuals at risk for Lynch syndrome. *J. Clin. Oncol.* **27**, 3981–3986 (2009)
70. R. Gryfe, H. Kim, E.T. Hsieh, M.D. Aronson, E.J. Holowaty, S. B. Bull, M. Redston, S. Gallinger, Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *N. Engl. J. Med.* **342**, 69–77 (2000)
71. C. Guastadisegni, M. Colafranceschi, L. Ottini, E. Dogliotti, Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. *Eur. J. Cancer* **46**, 2788–2798 (2010)
72. G. Hall, A. Clarkson, A. Shi, E. Langford, H. Leung, R.P. Eckstein, A.J. Gill, Immunohistochemistry for PMS2 and MSH6 alone can replace a four antibody panel for mismatch repair deficiency screening in colorectal adenocarcinoma. *Pathology* **42**, 409–413 (2010)
73. K.C. Halling, A.J. French, S.K. McDonnell, L.J. Burgart, D.J. Schaid, B.J. Peterson, L. Moon-Tasson, M.R. Mahoney, D.J. Sargent, M.J. O'Connell, T.E. Witzig, G.H. Farr Jr., R.M. Goldberg, S.N. Thibodeau, Microsatellite instability and 8p allelic imbalance in stage B2 and C colorectal cancers. *J. Natl Cancer Inst.* **91**, 1295–1303 (1999)
74. H. Hampel, W.L. Frankel, E. Martin, M. Arnold, K. Khanduja, P. Kuebler, M. Clendenning, K. Sotamaa, T. Prior, J.A. Westman, J. Panescu, D. Fix, J. Lockman, J. LaJeunesse, I. Comeras, A. de la Chapelle, Feasibility of screening for Lynch syndrome among patients with colorectal cancer. *J. Clin. Oncol.* **26**, 5783–5788 (2008)
75. H. Hampel, J.A. Stephens, E. Pukkala, R. Sankila, L.A. Aaltonen, J.P. Mecklin, A. de la Chapelle, Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. *Gastroenterology* **129**, 415–421 (2005)
76. B.D. Harfe, S. Jinks-Robertson, DNA mismatch repair and genetic instability. *Annu. Rev. Genet.* **34**, 359–399 (2000)
77. N.J. Hawkins, R.L. Ward, Sporadic colorectal cancers with microsatellite instability and their possible origin in hyperplastic polyps and serrated adenomas. *J. Natl Cancer Inst.* **93**, 1307–1313 (2001)
78. A. Hemminki, J.P. Mecklin, H. Jarvinen, L.A. Aaltonen, H. Joensuu, Microsatellite instability is a favorable prognostic indicator in patients with colorectal cancer receiving chemotherapy. *Gastroenterology* **119**, 921–928 (2000)
79. P.M. Hempen, L. Zhang, R.K. Bansal, C.A. Iacobuzio-Donahue, K.M. Murphy, A. Maitra, B. Vogelstein, R.H. Whitehead, S.D. Markowitz, J.K. Willson, C.J. Yeo, R.H. Hruban, S.E. Kern, Evidence of selection for clones having genetic inactivation of the activin A type II receptor (ACVR2) gene in gastrointestinal cancers. *Cancer Res.* **63**, 994–999 (2003)
80. Y.M. Hendriks, A. Wagner, H. Morreau, F. Menko, A. Stormorken, F. Quehenberger, L. Sandkuijl, P. Moller, M. Genuardi, H. Van Houwelingen, C. Tops, M. Van Puijenbroek, P. Verkuiljen, G. Kenter, A. Van Mil, H. Meijers-Heijboer, G.B. Tan, M.H. Breuning, R. Fodde, J.T. Wijnen, A.H. Brocker-Vriends, H. Vasen, Cancer risk in hereditary nonpolyposis colorectal cancer due to MSH6 mutations: impact on counseling and surveillance. *Gastroenterology* **127**, 17–25 (2004)
81. J.G. Herman, A. Umar, K. Polyak, J.R. Graff, N. Ahuja, J.P. Issa, S. Markowitz, J.K. Willson, S.R. Hamilton, K.W. Kinzler, M.F. Kane, R.D. Kolodner, B. Vogelstein, T.A. Kunkel, S.B. Baylin, Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc. Natl Acad. Sci. USA* **95**, 6870–6875 (1998)
82. N.M. Hollingsworth, L. Ponte, C. Halsey, MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes Dev.* **9**, 1728–1739 (1995)
83. Y. Ionov, H. Yamamoto, S. Krajewski, J.C. Reed, M. Perucho, Mutational inactivation of the proapoptotic gene BAX confers selective advantage during tumor clonal evolution. *Proc. Natl Acad. Sci. USA* **97**, 10872–10877 (2000)
84. J.P. Issa, CpG island methylator phenotype in cancer. *Nat. Rev. Cancer* **4**, 988–993 (2004)
85. S. Jacob, M. Aguado, D. Fallik, F. Praz, The role of the DNA mismatch repair system in the cytotoxicity of the topoisomerase inhibitors camptothecin and etoposide to human colorectal cancer cells. *Cancer Res.* **61**, 6555–6562 (2001)
86. S. Jacob, C. Miquel, A. Sarasin, F. Praz, Effects of camptothecin on double-strand break repair by non-homologous end-joining in DNA mismatch repair-deficient human colorectal cancer cell lines. *Nucleic Acids Res.* **33**, 106–113 (2005)
87. M.A. Jenkins, S. Hayashi, A.M. O'Shea, L.J. Burgart, T.C. Smyrk, D. Shimizu, P.M. Waring, A.R. Ruszkiewicz, A.F. Pollett, M. Redston, M.A. Barker, J.A. Baron, G.R. Casey, J.G. Dowty, G.G. Giles, P. Limburg, P. Newcomb, J.P. Young, M.D. Walsh, S.N. Thibodeau, N.M. Lindor, L. Lemarchand, S. Gallinger, R.W. Haile, J.D. Potter, J.L. Hopper, J.R. Jass, Pathology features in Bethesda

- guidelines predict colorectal cancer microsatellite instability: a population-based study. *Gastroenterology* **133**, 48–56 (2007)
88. S.Y. Jeong, K.H. Shin, J.H. Shin, J.L. Ku, Y.K. Shin, S.Y. Park, W.H. Kim, J.G. Park, Microsatellite instability and mutations in DNA mismatch repair genes in sporadic colorectal cancers. *Dis. Colon Rectum* **46**, 1069–1077 (2003)
 89. J. Jiricny, M. Nystrom-Lahti, Mismatch repair defects in cancer. *Curr. Opin. Genet. Dev.* **10**, 157–161 (2000)
 90. R. Jover, A. Castells, X. Llor, M. Andreu, Predictive value of microsatellite instability for benefit from adjuvant fluorouracil chemotherapy in colorectal cancer. *Gut* **55**, 1819–1820 (2006)
 91. R. Jover, P. Zapater, A. Castells, X. Llor, M. Andreu, J. Cubiella, F. Balaguer, L. Sempere, R.M. Xicola, L. Bujanda, J.M. Rene, J. Clofent, X. Bessa, J.D. Morillas, D. Nicolas-Perez, E. Pons, A. Paya, C. Alenda, The efficacy of adjuvant chemotherapy with 5-fluorouracil in colorectal cancer depends on the mismatch repair status. *Eur. J. Cancer* **45**, 365–373 (2009)
 92. R. Jover, P. Zapater, A. Castells, X. Llor, M. Andreu, J. Cubiella, V. Pinol, R.M. Xicola, L. Bujanda, J.M. Rene, J. Clofent, X. Bessa, J.D. Morillas, D. Nicolas-Perez, A. Paya, C. Alenda, Mismatch repair status in the prediction of benefit from adjuvant fluorouracil chemotherapy in colorectal cancer. *Gut* **55**, 848–855 (2006)
 93. S.H. Jun, T.G. Kim, C. Ban, DNA mismatch repair system. Classical and fresh roles. *FEBS J* **273**, 1609–1619 (2006)
 94. B. Jung, R.T. Doctolero, A. Tajima, A.K. Nguyen, T. Keku, R.S. Sandler, J.M. Carethers, Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers. *Gastroenterology* **126**, 654–659 (2004)
 95. B.H. Jung, S.E. Beck, J. Cabral, E. Chau, B.L. Cabrera, A. Fiorino, E.J. Smith, M. Bocanegra, J.M. Carethers, Activin type 2 receptor restoration in MSI-H colon cancer suppresses growth and enhances migration with activin. *Gastroenterology* **132**, 633–644 (2007)
 96. F.A. Kadyrov, L. Dzantiev, N. Constantin, P. Modrich, Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297–308 (2006)
 97. T. Kambara, L.A. Simms, V.L. Whitehall, K.J. Spring, C.V. Wynter, M.D. Walsh, M.A. Barker, S. Arnold, A. McGivern, N. Matsubara, N. Tanaka, T. Higuchi, J. Young, J.R. Jass, B.A. Leggett, BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. *Gut* **53**, 1137–1144 (2004)
 98. M.F. Kane, M. Loda, G.M. Gaida, J. Lipman, R. Mishra, H. Goldman, J.M. Jessup, R. Kolodner, 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**, 808–811 (1997)
 99. G.P. Kim, L.H. Colangelo, H.S. Wieand, S. Paik, I.R. Kirsch, N. Wolmark, C.J. Allegra, Prognostic and predictive roles of high-degree microsatellite instability in colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project Collaborative Study. *J. Clin. Oncol.* **25**, 767–772 (2007)
 100. J. Kim, Y. Hong, J. Lee, M. Ryu, H. Chang, S. Jang, M. Kim, C. Yu and D.S. Kim, Association between deficient mismatch repair system and efficacy to irinotecan containing first-line chemotherapy in patients with sporadic metastatic colorectal cancer. 2010 ASCO Annual Meeting Abstract **3579** (2010)
 101. N.G. Kim, Y.R. Choi, M.J. Baek, Y.H. Kim, H. Kang, N.K. Kim, J.S. Min, H. Kim, Frameshift mutations at coding mononucleotide repeats of the hRAD50 gene in gastrointestinal carcinomas with microsatellite instability. *Cancer Res.* **61**, 36–38 (2001)
 102. S.T. Kim, J. Lee, S.H. Park, J.O. Park, H.Y. Lim, W.K. Kang, J. Y. Kim, Y.H. Kim, D.K. Chang, P.L. Rhee, D.S. Kim, H. Yun, Y. B. Cho, H.C. Kim, S.H. Yun, H.K. Chun, W.Y. Lee and Y.S. Park, The effect of DNA mismatch repair (MMR) status on oxaliplatin-based first-line chemotherapy as in recurrent or metastatic colon cancer. *Med. Oncol.* (2009)
 103. S.T. Kim, J. Lee, S.H. Park, J.O. Park, H.Y. Lim, W.K. Kang, J. Y. Kim, Y.H. Kim, D.K. Chang, P.L. Rhee, D.S. Kim, H. Yun, Y. B. Cho, H.C. Kim, S.H. Yun, W.Y. Lee, H.K. Chun, Y.S. Park, Clinical impact of microsatellite instability in colon cancer following adjuvant FOLFOX therapy. *Cancer Chemother. Pharmacol.* **66**, 659–667 (2010)
 104. E. Kondo, A. Horii, S. Fukushige, The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2. *Nucleic Acids Res.* **29**, 1695–1702 (2001)
 105. M. Koopman, G.A. Kortman, L. Mekenkamp, M.J. Ligtenberg, N. Hoogerbrugge, N.F. Antonini, C.J. Punt, J.H. van Krieken, Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br. J. Cancer* **100**, 266–273 (2009)
 106. J.P. Kuebler, H.S. Wieand, M.J. O'Connell, R.E. Smith, L.H. Colangelo, G. Yothers, N.J. Petrelli, M.P. Findlay, T.E. Seay, J.N. Atkins, J.L. Zapas, J.W. Goodwin, L. Fehrenbacher, R.K. Ramanathan, B.A. Conley, P.J. Flynn, G. Soori, L.K. Colman, E.A. Levine, K.S. Lanier, N. Wolmark, Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: results from NSABP C-07. *J. Clin. Oncol.* **25**, 2198–2204 (2007)
 107. R.S. Lahue, K.G. Au, P. Modrich, DNA mismatch correction in a defined system. *Science* **245**, 160–164 (1989)
 108. C. Lamberti, S. Lundin, M. Bogdanow, C. Pagenstecher, N. Friedrichs, R. Buttner, T. Sauerbruch, Microsatellite instability did not predict individual survival of unselected patients with colorectal cancer. *Int. J. Colorectal Dis.* **22**, 145–152 (2007)
 109. G. Lanza, R. Gafa, A. Santini, I. Maestri, L. Guerzoni, L. Cavazzini, Immunohistochemical test for MLH1 and MSH2 expression predicts clinical outcome in stage II and III colorectal cancer patients. *J. Clin. Oncol.* **24**, 2359–2367 (2006)
 110. F.S. Leach, N.C. Nicolaides, N. Papadopoulos, B. Liu, J. Jen, R. Parsons, P. Peltomaki, P. Sistonen, L.A. Aaltonen, M. Nystrom-Lahti et al., Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* **75**, 1215–1225 (1993)
 111. K.A. Lewis, J. Bakkum-Gamez, R. Loewen, A.J. French, S.N. Thibodeau, W.A. Cliby, Mutations in the ataxia telangiectasia and rad3-related-checkpoint kinase 1 DNA damage response axis in colon cancers. *Genes Chromosom. Cancer* **46**, 1061–1068 (2007)
 112. H.R. Li, E.I. Shagisultanova, K. Yamashita, Z. Piao, M. Peruchio, S.R. Malkhosyan, Hypersensitivity of tumor cell lines with microsatellite instability to DNA double strand break producing chemotherapeutic agent bleomycin. *Cancer Res.* **64**, 4760–4767 (2004)
 113. J.T. Liang, K.C. Huang, H.S. Lai, P.H. Lee, Y.M. Cheng, H.C. Hsu, A.L. Cheng, C.H. Hsu, K.H. Yeh, S.M. Wang, C. Tang, K. J. Chang, High-frequency microsatellite instability predicts better chemosensitivity to high-dose 5-fluorouracil plus leucovorin chemotherapy for stage IV sporadic colorectal cancer after palliative bowel resection. *Int. J. Cancer* **101**, 519–525 (2002)
 114. A. Lievre, J.B. Bachet, D. Le Corre, V. Boige, B. Landi, J.F. Emile, J.F. Cote, G. Tomasic, C. Penna, M. Ducreux, P. Rougier, F. Penault-Llorca, P. Laurent-Puig, KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. *Cancer Res.* **66**, 3992–3995 (2006)
 115. N.M. Lindor, L.J. Burgart, O. Leontovich, R.M. Goldberg, J.M. Cunningham, D.J. Sargent, C. Walsh-Vockley, G.M. Petersen, M. D. Walsh, B.A. Leggett, J.P. Young, M.A. Barker, J.R. Jass, J. Hopper, S. Gallinger, B. Bapat, M. Redston, S.N. Thibodeau, Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J. Clin. Oncol.* **20**, 1043–1048 (2002)

116. N.M. Lindor, K. Rabe, G.M. Petersen, R. Haile, G. Casey, J. Baron, S. Gallinger, B. Bapat, M. Aronson, J. Hopper, J. Jass, L. LeMarchand, J. Grove, J. Potter, P. Newcomb, J.P. Terdiman, P. Conrad, G. Moslein, R. Goldberg, A. Ziogas, H. Anton-Culver, M. de Andrade, K. Siegmund, S.N. Thibodeau, L.A. Boardman, D. Seminara, Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. *JAMA* **293**, 1979–1985 (2005)
117. T. Liu, H. Yan, S. Kuusimanen, A. Percesepe, M.L. Bisgaard, M. Pedroni, P. Benatti, K.W. Kinzler, B. Vogelstein, M. Ponz de Leon, P. Peltomaki, A. Lindblom, The role of hPMS1 and hPMS2 in predisposing to colorectal cancer. *Cancer Res.* **61**, 7798–7802 (2001)
118. X. Llor, E. Pons, R.M. Xicola, A. Castells, C. Alenda, V. Pinol, M. Andreu, S. Castellvi-Bel, A. Paya, R. Jover, X. Bessa, A. Giros, A. Roca, M.A. Gassull, Differential features of colorectal cancers fulfilling Amsterdam criteria without involvement of the mutator pathway. *Clin. Cancer Res.* **11**, 7304–7310 (2005)
119. E. Loire, F. Praz, D. Higuier, P. Netter, G. Achaz, Hypermutability of genes in *Homo sapiens* due to the hosting of long mononucleotide repeats. *Mol. Biol. Evol.* **26**, 111–121 (2009)
120. R.A. Lothe, P. Peltomaki, G.I. Meling, L.A. Aaltonen, M. Nystrom-Lahti, L. Pylkkanen, K. Heimdal, T.I. Andersen, P. Moller, T.O. Rognum et al., Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. *Cancer Res.* **53**, 5849–5852 (1993)
121. M.B. Loughrey, P.M. Waring, A. Tan, M. Trivett, S. Kovalenko, V. Beshay, M.A. Young, G. McArthur, A. Boussioutas, A. Dobrovic, Incorporation of somatic BRAF mutation testing into an algorithm for the investigation of hereditary non-polyposis colorectal cancer. *Fam Cancer* **6**, 301–310 (2007)
122. H.T. Lynch, M.W. Shaw, C.W. Magnuson, A.L. Larsen, A.J. Krush, Hereditary factors in cancer. Study of two large midwestern kindreds. *Arch. Intern. Med.* **117**, 206–212 (1966)
123. R. Magrini, M.R. Bionde, M.L. Hanski, M. Notter, H. Scherubl, C.R. Boland, M. Zeit, C. Hanski, Cellular effects of CPT-11 on colon carcinoma cells: dependence on p53 and hMLH1 status. *Int. J. Cancer* **101**, 23–31 (2002)
124. A. Malesci, L. Laghi, P. Bianchi, G. Delconte, A. Randolph, V. Torri, C. Carnaghi, R. Doci, R. Rosati, M. Montorsi, M. Roncalli, L. Gennari, A. Santoro, Reduced likelihood of metastases in patients with microsatellite-unstable colorectal cancer. *Clin. Cancer Res.* **13**, 3831–3839 (2007)
125. S. Markowitz, J. Wang, L. Myeroff, R. Parsons, L. Sun, J. Lutterbaugh, R.S. Fan, E. Zborowska, K.W. Kinzler, B. Vogelstein et al., Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science* **268**, 1336–1338 (1995)
126. G. Marra, I. Iaccharino, T. Lettieri, G. Roscilli, P. Delmastro, J. Jiricny, G.T. Marsischky, S. Lee, J. Griffith, R.D. Kolodner, Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc. Natl Acad. Sci. USA* **95**, 8568–8573 (1998)
127. G.T. Marsischky, N. Filosi, M.F. Kane, R. Kolodner, Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes Dev.* **10**, 407–420 (1996)
128. G.T. Marsischky, R.D. Kolodner, Biochemical characterization of the interaction between the *Saccharomyces cerevisiae* MSH2-MSH6 complex and mispaired bases in DNA. *J. Biol. Chem.* **274**, 26668–26682 (1999)
129. A. McGovern, C.V. Wynter, V.L. Whitehall, T. Kambara, K.J. Spring, M.D. Walsh, M.A. Barker, S. Arnold, L.A. Simms, B.A. Leggett, J. Young, J.R. Jass, Promoter hypermethylation frequency and BRAF mutations distinguish hereditary non-polyposis colon cancer from sporadic MSI-H colon cancer. *Fam Cancer* **3**, 101–107 (2004)
130. M. Meyers, M.W. Wagner, H.S. Hwang, T.J. Kinsella, D.A. Boothman, Role of the hMLH1 DNA mismatch repair protein in fluoropyrimidine-mediated cell death and cell cycle responses. *Cancer Res.* **61**, 5193–5201 (2001)
131. J.M. Michael-Robinson, A. Biemer-Huttmann, D.M. Purdie, M. D. Walsh, L.A. Simms, K.G. Bide, J.P. Young, B.A. Leggett, J. R. Jass, G.L. Radford-Smith, Tumour infiltrating lymphocytes and apoptosis are independent features in colorectal cancer stratified according to microsatellite instability status. *Gut* **48**, 360–366 (2001)
132. C. Miquel, F. Borroni, S. Grandjouan, A. Auperin, J. Viguier, V. Velasco, P. Duvillard, F. Praz, J.C. Sabourin, Role of bax mutations in apoptosis in colorectal cancers with microsatellite instability. *Am. J. Clin. Pathol.* **123**, 562–570 (2005)
133. C. Miquel, S. Jacob, S. Grandjouan, A. Aime, J. Viguier, J.C. Sabourin, A. Sarasin, A. Duval, F. Praz, Frequent alteration of DNA damage signalling and repair pathways in human colorectal cancers with microsatellite instability. *Oncogene* **26**, 5919–5926 (2007)
134. P. Modrich, R. Lahue, Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**, 101–133 (1996)
135. P. Mongiat-Artus, C. Miquel, M. Van der Aa, O. Buhard, R. Hamelin, H. Soliman, C. Bangma, A. Janin, P. Teillac, T. van Kwast, F. Praz, Microsatellite instability and mutation analysis of candidate genes in urothelial cell carcinomas of upper urinary tract. *Oncogene* **25**, 2113–2118 (2006)
136. C.I. Muller, K. Schulmann, A. Reinacher-Schick, N. Andre, D. Arnold, A. Tannapfel, H. Arkenau, S.A. Hahn, S.H. Schmoll, R. Porschen, W. Schmieg, U. Graeven, Predictive and prognostic value of microsatellite instability in patients with advanced colorectal cancer treated with a fluoropyrimidine and oxaliplatin containing first-line chemotherapy. A report of the AIO Colorectal Study Group. *Int. J. Colorectal Dis.* **23**, 1033–1039 (2008)
137. T. Nakagawa, A. Datta, R.D. Kolodner, I. Iaccharino, G. Marra, F. Palombo, J. Jiricny, Multiple functions of MutS- and MutL-related heterodimers. *Proc. Natl Acad. Sci. USA* **96**, 14186–14188 (1999)
138. N.C. Nicolaides, N. Papadopoulos, B. Liu, Y.F. Wei, K.C. Carter, S.M. Ruben, C.A. Rosen, W.A. Haseltine, R.D. Fleischmann, C. M. Fraser et al., Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* **371**, 75–80 (1994)
139. M. Nystrom-Lahti, R. Parsons, P. Sistonen, L. Pylkkanen, L.A. Aaltonen, F.S. Leach, S.R. Hamilton, P. Watson, E. Bronson, R. Fusaro et al., Mismatch repair genes on chromosomes 2p and 3p account for a major share of hereditary nonpolyposis colorectal cancer families evaluable by linkage. *Am. J. Hum. Genet.* **55**, 659–665 (1994)
140. M. Nystrom-Lahti, P. Sistonen, J.P. Mecklin, L. Pylkkanen, L.A. Aaltonen, H. Jarvinen, J. Weissenbach, A. de la Chapelle, P. Peltomaki, Close linkage to chromosome 3p and conservation of ancestral founding haplotype in hereditary nonpolyposis colorectal cancer families. *Proc. Natl Acad. Sci. USA* **91**, 6054–6058 (1994)
141. V. O'Brien, R. Brown, Signalling cell cycle arrest and cell death through the MMR system. *Carcinogenesis* **27**, 682–692 (2006)
142. S. Ogino, K. Nosho, G.J. Kirkner, K. Shima, N. Irahara, S. Kure, A.T. Chan, J.A. Engelman, P. Kraft, L.C. Cantley, E.L. Giovannucci, C.S. Fuchs, PIK3CA mutation is associated with poor prognosis among patients with curatively resected colon cancer. *J. Clin. Oncol.* **27**, 1477–1484 (2009)
143. L.I. Overbeek, M.J. Ligtenberg, R.W. Willems, R.P. Hermens, W.A. Blokx, S.V. Dubois, H. van der Linden, J.W. Meijer, M.L. Mlynck-Kersjes, N. Hoogerbrugge, K.M. Hebeda, J.H. van Krieken, Interpretation of immunohistochemistry for mismatch

- repair proteins is only reliable in a specialized setting. *Am. J. Surg. Pathol.* **32**, 1246–1251 (2008)
144. F. Palombo, I. Iaccarino, E. Nakajima, M. Ikejima, T. Shimada, J. Jiricny, hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.* **6**, 1181–1184 (1996)
 145. R. Parsons, L.L. Myeroff, B. Liu, J.K. Willson, S.D. Markowitz, K.W. Kinzler, B. Vogelstein, Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. *Cancer Res.* **55**, 5548–5550 (1995)
 146. T.M. Pawlik, C.P. Raut, M.A. Rodriguez-Bigas, Colorectal carcinogenesis: MSI-H versus MSI-L. *Dis. Markers* **20**, 199–206 (2004)
 147. P. Peltomaki, Lynch syndrome genes. *Fam Cancer* **4**, 227–232 (2005)
 148. P. Peltomaki, L.A. Aaltonen, P. Sistonen, L. Pylkanen, J.P. Mecklin, H. Jarvinen, J.S. Green, J.R. Jass, J.L. Weber, F.S. Leach et al., Genetic mapping of a locus predisposing to human colorectal cancer. *Science* **260**, 810–812 (1993)
 149. P. Peltomaki, H. Vasen, Mutations associated with HNPCC predisposition – Update of ICG-HNPCC/INSiGHT mutation database. *Dis. Markers* **20**, 269–276 (2004)
 150. S.M. Phillips, A. Banerjee, R. Feakins, S.R. Li, S.A. Bustin, S. Dorudi, Tumour-infiltrating lymphocytes in colorectal cancer with microsatellite instability are activated and cytotoxic. *Br. J. Surg.* **91**, 469–475 (2004)
 151. M. Pocard, R. Bras-Goncalves, R. Hamelin, J. Northover, M.F. Poupon, Response to 5-fluorouracil of orthotopically xenografted human colon cancers with a microsatellite instability: influence of P53 status. *Anticancer Res.* **20**, 85–90 (2000)
 152. S. Popat, R. Hubner, R.S. Houlston, Systematic review of microsatellite instability and colorectal cancer prognosis. *J. Clin. Oncol.* **23**, 609–618 (2005)
 153. T.A. Prolla, DNA mismatch repair and cancer. *Curr. Opin. Cell Biol.* **10**, 311–316 (1998)
 154. M. Radman, R. Wagner, B.W. Glickman, Mismatch repair in *Escherichia coli*. *Annu. Rev. Genet.* **20**, 523–538 (1986)
 155. M. Raschle, G. Marra, M. Nystrom-Lahti, P. Schar, J. Jiricny, Identification of hMutLbeta, a heterodimer of hMLH1 and hPMS1. *J. Biol. Chem.* **274**, 32368–32375 (1999)
 156. C. Rayssiguier, D.S. Thaler, M. Radman, The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**, 396–401 (1989)
 157. C.M. Ribic, D.J. Sargent, M.J. Moore, S.N. Thibodeau, A.J. French, R.M. Goldberg, S.R. Hamilton, P. Laurent-Puig, R. Gryfe, L.E. Shepherd, D. Tu, M. Redston, S. Gallinger, Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N. Engl. J. Med.* **349**, 247–257 (2003)
 158. R. Rodriguez, L.T. Hansen, G. Phear, J. Scorah, M. Spang-Thomsen, A. Cox, T. Helleday, M. Meuth, Thymidine selectively enhances growth suppressive effects of camptothecin/irinotecan in MSI+ cells and tumors containing a mutation of MRE11. *Clin. Cancer Res.* **14**, 5476–5483 (2008)
 159. R. Roman, M. Verdu, M. Calvo, A. Vidal, X. Sanjuan, M. Jimeno, A. Salas, J. Autonell, I. Trias, M. Gonzalez, B. Garcia, N. Rodon, X. Puig, Microsatellite instability of the colorectal carcinoma can be predicted in the conventional pathologic examination. A prospective multicentric study and the statistical analysis of 615 cases consolidate our previously proposed logistic regression model. *Virchows Arch.* **456**, 533–541 (2010)
 160. P. Ross-Macdonald, G.S. Roeder, Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* **79**, 1069–1080 (1994)
 161. B. Rousseau, B. Chibaudel, J.B. Bachet, A.K. Larsen, C. Tournigand, C. Louvet, T. Andre, A. de Gramont, Stage II and stage III colon cancer: treatment advances and future directions. *Cancer J.* **16**, 202–209 (2010)
 162. S. Salahshor, U. Kressner, H. Fischer, G. Lindmark, B. Glimelius, L. Pahlman, A. Lindblom, Microsatellite instability in sporadic colorectal cancer is not an independent prognostic factor. *Br. J. Cancer* **81**, 190–193 (1999)
 163. P. Salama, M. Phillips, F. Grien, M. Morris, N. Zeps, D. Joseph, C. Platell, B. Iacopetta, Tumor-infiltrating FOXP3+ T regulatory cells show strong prognostic significance in colorectal cancer. *J. Clin. Oncol.* **27**, 186–192 (2009)
 164. L.B. Saltz, D. Niedzwiecki, D. Hollis, R.M. Goldberg, A. Hantel, J.P. Thomas, A.L. Fields, R.J. Mayer, Irinotecan fluorouracil plus leucovorin is not superior to fluorouracil plus leucovorin alone as adjuvant treatment for stage III colon cancer: results of CALGB 89803. *J. Clin. Oncol.* **25**, 3456–3461 (2007)
 165. W.S. Samowitz, K. Curtin, K.N. Ma, D. Schaffer, L.W. Coleman, M. Leppert, M.L. Slattery, Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. *Cancer Epidemiol. Biomark. Prev.* **10**, 917–923 (2001)
 166. D.J. Sargent, S. Marsoni, G. Monges, S.N. Thibodeau, R. Labianca, S.R. Hamilton, A.J. French, B. Kabat, N.R. Foster, V. Torri, C. Ribic, A. Grothey, M. Moore, A. Zaniboni, J.F. Seitz, F. Sinicrope and S. Gallinger, Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J. Clin. Oncol.* (2010)
 167. P. Schweizer, A.L. Moisio, S.A. Kuismanen, K. Truninger, R. Vierumaki, R. Salovaara, J. Arola, R. Butzow, J. Jiricny, P. Peltomaki, M. Nystrom-Lahti, Lack of MSH2 and MSH6 characterizes endometrial but not colon carcinomas in hereditary nonpolyposis colorectal cancer. *Cancer Res.* **61**, 2813–2815 (2001)
 168. Y. Schwitalle, M. Kloor, S. Eiermann, M. Linnebacher, P. Kienle, H.P. Knaebel, M. Tariverdian, A. Benner, M. von Knebel Doeberitz, Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. *Gastroenterology* **134**, 988–997 (2008)
 169. S.N. Shah, S.E. Hile, K.A. Eckert, Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes. *Cancer Res.* **70**, 431–435 (2010)
 170. J. Shia, L.H. Tang, E. Vakiani, J.G. Guillem, Z.K. Stadler, R.A. Soslow, N. Katabi, M.R. Weiser, P.B. Paty, L.K. Temple, G.M. Nash, W.D. Wong, K. Offit, D.S. Klimstra, Immunohistochemistry as first-line screening for detecting colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome: a 2-antibody panel may be as predictive as a 4-antibody panel. *Am. J. Surg. Pathol.* **33**, 1639–1645 (2009)
 171. F.A. Sinicrope, DNA mismatch repair and adjuvant chemotherapy in sporadic colon cancer. *Nat. Rev. Clin. Oncol.* **7**, 174–177 (2010)
 172. F.A. Sinicrope, R.L. Rego, M.M. Garrity-Park, N.R. Foster, D.J. Sargent, R.M. Goldberg, M. Wiesenfeld, T.E. Witzig, S.N. Thibodeau, L.J. Burgart, Alterations in cell proliferation and apoptosis in colon cancers with microsatellite instability. *Int. J. Cancer* **120**, 1232–1238 (2007)
 173. F.A. Sinicrope, R.L. Rego, K.C. Halling, N. Foster, D.J. Sargent, B. La Plant, A.J. French, J.A. Laurie, R.M. Goldberg, S.N. Thibodeau, T.E. Witzig, Prognostic impact of microsatellite instability and DNA ploidy in human colon carcinoma patients. *Gastroenterology* **131**, 729–737 (2006)
 174. E. Stoffel, B. Mukherjee, V.M. Raymond, N. Tayob, F. Kastrinos, J. Sparr, F. Wang, P. Bandipalliam, S. Syngal, S.B. Gruber, Calculation of risk of colorectal and endometrial cancer

- among patients with Lynch syndrome. *Gastroenterology* **137**, 1621–1627 (2009)
175. L. Stojic, R. Brun, J. Jiricny, Mismatch repair and DNA damage signalling. *DNA Repair (Amst.)* **3**, 1091–1101 (2004)
176. N. Sugawara, F. Paques, M. Colaiacovo, J.E. Haber, Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc. Natl Acad. Sci. USA* **94**, 9214–9219 (1997)
177. H. Takemura, V.A. Rao, O. Sordet, T. Furuta, Z.H. Miao, L. Meng, H. Zhang, Y. Pommier, Defective Mre11-dependent activation of Chk2 by ataxia telangiectasia mutated in colorectal carcinoma cells in response to replication-dependent DNA double strand breaks. *J. Biol. Chem.* **281**, 30814–30823 (2006)
178. J.P. Terdiman, J.R. Gum Jr., P.G. Conrad, G.A. Miller, V. Weinberg, S.C. Crawley, T.R. Levin, C. Reeves, A. Schmitt, M. Hepburn, M.H. Sleisenger, Y.S. Kim, Efficient detection of hereditary nonpolyposis colorectal cancer gene carriers by screening for tumor microsatellite instability before germline genetic testing. *Gastroenterology* **120**, 21–30 (2001)
179. S.N. Thibodeau, G. Bren, D. Schaid, Microsatellite instability in cancer of the proximal colon. *Science* **260**, 816–819 (1993)
180. A. Umar, C.R. Boland, J.P. Terdiman, S. Syngal, A. de la Chapelle, J. Ruschoff, R. Fishel, N.M. Lindor, L.J. Burgart, R. Hamelin, S.R. Hamilton, R.A. Hiatt, J. Jass, A. Lindblom, H.T. Lynch, P. Peltomaki, S.D. Ramsey, M.A. Rodriguez-Bigas, H.F. Vasen, E.T. Hawk, J.C. Barrett, A.N. Freedman, S. Srivastava, Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J. Natl Cancer Inst.* **96**, 261–268 (2004)
181. A. Vaisman, M. Varchenko, A. Umar, T.A. Kunkel, J.I. Risinger, J.C. Barrett, T.C. Hamilton, S.G. Chaney, The role of hMLH1, hMSH3, and hMSH6 defects in cisplatin and oxaliplatin resistance: correlation with replicative bypass of platinum-DNA adducts. *Cancer Res.* **58**, 3579–3585 (1998)
182. E. Van Cutsem, R. Labianca, G. Bodoky, C. Barone, E. Aranda, B. Nordlinger, C. Topham, J. Tabernero, T. Andre, A.F. Sobrero, E. Mini, R. Greil, F. Di Costanzo, L. Collette, L. Cisar, X. Zhang, D. Khayat, C. Bokemeyer, A.D. Roth, D. Cunningham, Randomized phase III trial comparing biweekly infusional fluorouracil/leucovorin alone or with irinotecan in the adjuvant treatment of stage III colon cancer: PETACC-3. *J. Clin. Oncol.* **27**, 3117–3125 (2009)
183. H.F. Vasen, J.P. Mecklin, P.M. Khan, H.T. Lynch, The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis. Colon Rectum* **34**, 424–425 (1991)
184. H.F. Vasen, G. Moslein, A. Alonso, S. Aretz, I. Bernstein, L. Bertario, I. Blanco, S. Bulow, J. Burn, G. Capella, C. Colas, C. Engel, I. Frayling, N. Rahner, F.J. Hes, S. Hodgson, J.P. Mecklin, P. Moller, T. Myrhorj, F.M. Nagengast, Y. Parc, M. Ponz de Leon, L. Renkonen-Sinisalo, J.R. Sampson, A. Stormorken, S. Tejpar, H.J. Thomas, J. Wijnen, J. Lubinski, H. Jarvinen, E. Claes, K. Heinimann, J.A. Karagiannis, A. Lindblom, I. Dove-Edwin, H. Muller, Recommendations to improve identification of hereditary and familial colorectal cancer in Europe. *Fam Cancer* **9**, 109–115 (2010)
185. H.F. Vasen, G. Moslein, A. Alonso, I. Bernstein, L. Bertario, I. Blanco, J. Burn, G. Capella, C. Engel, I. Frayling, W. Friedl, F.J. Hes, S. Hodgson, J.P. Mecklin, P. Moller, F. Nagengast, Y. Parc, L. Renkonen-Sinisalo, J.R. Sampson, A. Stormorken, J. Wijnen, Guidelines for the clinical management of Lynch syndrome (hereditary non-polyposis cancer). *J. Med. Genet.* **44**, 353–362 (2007)
186. H.F. Vasen, P. Watson, J.P. Mecklin, H.T. Lynch, New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* **116**, 1453–1456 (1999)
187. V. Vassileva, A. Millar, L. Briollais, W. Chapman, B. Bapat, Genes involved in DNA repair are mutational targets in endometrial cancers with microsatellite instability. *Cancer Res.* **62**, 4095–4099 (2002)
188. S. Velho, C. Oliveira, A. Ferreira, A.C. Ferreira, G. Suriano, S. Schwartz Jr., A. Duval, F. Carneiro, J.C. Machado, R. Hamelin, R. Seruca, The prevalence of PIK3CA mutations in gastric and colon cancer. *Eur. J. Cancer* **41**, 1649–1654 (2005)
189. E. Vilar, S.B. Gruber, Microsatellite instability in colorectal cancer—the stable evidence. *Nat. Rev. Clin. Oncol.* **7**, 153–162 (2010)
190. E. Vilar, M. Scaltriti, J. Balmana, C. Saura, M. Guzman, J. Arribas, J. Baselga, J. Tabernero, Microsatellite instability due to hMLH1 deficiency is associated with increased cytotoxicity to irinotecan in human colorectal cancer cell lines. *Br. J. Cancer* **99**, 1607–1612 (2008)
191. A. Walther, R. Houlston, I. Tomlinson, Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut* **57**, 941–950 (2008)
192. L. Wang, J.M. Cunningham, J.L. Winters, J.C. Guenther, A.J. French, L.A. Boardman, L.J. Burgart, S.K. McDonnell, D.J. Schaid, S.N. Thibodeau, BRAF mutations in colon cancer are not likely attributable to defective DNA mismatch repair. *Cancer Res.* **63**, 5209–5212 (2003)
193. R. Ward, A. Meagher, I. Tomlinson, T. O'Connor, M. Norrie, R. Wu, N. Hawkins, Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. *Gut* **48**, 821–829 (2001)
194. A.S. Warthin, Heredity with reference to carcinoma. *Arch. Intern. Med.* **12**, 546–555 (1913)
195. P. Watson, H.F. Vasen, J.P. Mecklin, I. Bernstein, M. Aarnio, H.J. Jarvinen, T. Myrhorj, L. Sunde, J.T. Wijnen, H.T. Lynch, The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. *Int. J. Cancer* **123**, 444–449 (2008)
196. D.J. Weisenberger, K.D. Siegmund, M. Campan, J. Young, T.I. Long, M.A. Faasse, G.H. Kang, M. Widschwendter, D. Weener, D. Buchanan, H. Koh, L. Simms, M. Barker, B. Leggett, J. Levine, M. Kim, A.J. French, S.N. Thibodeau, J. Jass, R. Haile, P.W. Laird, CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.* **38**, 787–793 (2006)
197. K.M. Welsh, A.L. Lu, S. Clark, P. Modrich, Isolation and characterization of the *Escherichia coli* mutH gene product. *J. Biol. Chem.* **262**, 15624–15629 (1987)
198. J.L. Westra, M. Schaapveld, H. Hollema, J.P. de Boer, M.M. Kraak, D. de Jong, A. ter Elst, N.H. Mulder, C.H. Buys, R.M. Hofstra, J.T. Plukker, Determination of TP53 mutation is more relevant than microsatellite instability status for the prediction of disease-free survival in adjuvant-treated stage III colon cancer patients. *J. Clin. Oncol.* **23**, 5635–5643 (2005)
199. S.M. Woerner, A. Benner, C. Sutter, M. Schiller, Y.P. Yuan, G. Keller, P. Bork, M.K. Doeberitz, J.F. Gebert, Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes. *Oncogene* **22**, 2226–2235 (2003)
200. S.M. Woerner, Y.P. Yuan, A. Benner, S. Korff, M. von Knebel Doeberitz, P. Bork, SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Res.* **38**, D682–689 (2010)
201. C.M. Wright, O.F. Dent, M. Barker, R.C. Newland, P.H. Chapuis, E.L. Bokey, J.P. Young, B.A. Leggett, J.R. Jass, G.A. Macdonald, Prognostic significance of extensive microsatellite

- instability in sporadic clinicopathological stage C colorectal cancer. *Br. J. Surg.* **87**, 1197–1202 (2000)
202. M. Yamada, E. O'Regan, R. Brown, P. Karran, Selective recognition of a cisplatin-DNA adduct by human mismatch repair proteins. *Nucleic Acids Res.* **25**, 491–496 (1997)
203. A. Zaanan, P. Cuilliere-Dartigues, A. Guilloux, Y. Parc, C. Louvet, A. de Gramont, E. Tiret, S. Dumont, B. Gayet, P. Validire, J.F. Flejou, A. Duval, F. Praz, Impact of p53 expression and microsatellite instability on stage III colon cancer disease-free survival in patients treated by 5-fluorouracil and leucovorin with or without oxaliplatin. *Ann. Oncol.* **21**, 772–780 (2010)
204. Z.Z. Zdraveski, J.A. Mello, C.K. Farinelli, J.M. Essigmann, M. G. Marinus, MutS preferentially recognizes cisplatin- over oxaliplatin-modified DNA. *J. Biol. Chem.* **277**, 1255–1260 (2002)

RESEARCH ARTICLE

Intestinal epithelial stem cells do not protect their genome by assymmetric chromosome segregation

Marion Escobar, Pierre Nicolas, Fatiha Sangar, Sabine Laurent-Chabalier, Philippe
Clair, Dominique Joubert, Philippe Jay and Catherine Legraverend

Nature Communications

2011

ARTICLE

Received 19 Oct 2010 | Accepted 2 Mar 2011 | Published 29 Mar 2011

DOI: 10.1038/ncomms1260

Intestinal epithelial stem cells do not protect their genome by asymmetric chromosome segregation

Marion Escobar^{1,2,3,*}, Pierre Nicolas^{4,*}, Fatiha Sangar^{1,2,3,†}, Sabine Laurent-Chabalier⁵, Philippe Clair^{1,2,3,†}, Dominique Joubert^{1,2,3}, Philippe Jay^{1,2,3} & Catherine Legraverend^{1,2,3}

The idea that stem cells of adult tissues with high turnover are protected from DNA replication-induced mutations by maintaining the same 'immortal' template DNA strands together through successive divisions has been tested in several tissues. In the epithelium of the small intestine, the provided evidence was based on the assumption that stem cells are located above Paneth cells. The results of genetic lineage-tracing experiments point instead to crypt base columnar cells intercalated between Paneth cells as *bona fide* stem cells. Here we show that these cells segregate most, if not all, of their chromosomes randomly, both in the intact and in the regenerating epithelium. Therefore, the 'immortal' template DNA strand hypothesis does not apply to intestinal epithelial stem cells, which must rely on other strategies to avoid accumulating mutations.

¹ CNRS, UMR-5203, Institut de Génétique Fonctionnelle, F-34000 Montpellier, France. ² INSERM, U661, F-34000 Montpellier, France. ³ Universités de Montpellier 1&2, UMR-5203, F-34000 Montpellier, France. ⁴ INRA, UR-1077, Mathématique Informatique et Génome, F-78000 Jouy-en-Josas, France. ⁵ CNRS, UPR-1142, Institut de Génétique Humaine, F-34000 Montpellier, France. *These authors contributed equally to this work. †Present addresses: INSERM, UMR-S938-UPMC, Centre de Recherche Saint-Antoine, F-75000 Paris, France (F.S.); Université de Montpellier 2, Plateforme qPCR Haut débit Montpellier GenomiX, cc91, F-34000 Montpellier, France (P.C.). Correspondence and requests for materials should be addressed to C.L. (email: catherine.legraverend@igf.cnrs.fr).

After semiconservative replication of eukaryotic chromosomes, each sister chromatid inherits one of the two original template DNA strands, and segregation of the new chromosomes between daughter cells is thought to occur randomly. However, past the period of rapid growth characteristic of development, the stem cells of some tissues are suspected to switch to a nonrandom mode of chromosome segregation (CS; reviewed in ref. 1). From that point of time onwards, one of the two template DNA strands of each chromosome, the ‘immortal strand’, according to the Cairns hypothesis², is thought to be selectively and repeatedly transmitted to the ‘stem’ daughter cell, division after division (Fig. 1a, asterisk), whereas the other template is inherited by the other, shorter-lived, daughter cell that will eventually migrate out of the stem compartment. This CS mode is assumed to prevent the accumulation of replication errors in the genome of stem cells². Alternatively, if only a subset of chromosomes are involved, nonrandom CS could link the transmission of different epigenetic marks to differences in daughter-cell fate (reviewed in ref. 3). In several mammalian tissues, including the small intestine⁴, breast⁵, brain⁶ and muscle^{7,8}, nonrandom CS has been linked to self-renewing divisions of a fraction of stem cells. The main evidence provided in support of asymmetric CS in brain, breast and muscle cells was the observation of asymmetric label segregation in cultured cells. Two other experimental paradigms were examined in the intestine and the muscle: juvenile growth^{4,7} and regeneration after injury in adult animals^{4,7,8}. In the latter case, stem or progenitor cells of the injured tissue are assumed to transiently switch back to a symmetric mode of cell division involving random CS. This presumably enables incorporation of nucleic-acid base analogues into both DNA templates. Following the regenerative response, stem or progenitor cells are thought to resume nonrandom CS, and the repeated and selective inheritance of the ‘immortal’ strands is inferred from the long-term retention of labelled DNA in their nucleus, during successive divisions in the absence of label.

The crypt of the intestinal epithelium is indeed seen as a potential candidate site for nonrandom CS because of its high cell turnover and capacity for regeneration sustained by a high number of stem cells^{9,10}, as well as the fact that no quiescent intestinal stem cell has been identified so far, which could have a role of guardian of the genome¹¹. However, the precise location of stem cells has long remained contentious (reviewed in ref. 12). Until very recently, the so-called cell position +4 (cp4) model placed stem cells at the base of the transit-amplifying progenitor compartment, just above the last Paneth cells in the small intestine¹³. The other model, supported by lineage studies^{14,15}, assigned the status of stem cells to the immature proliferative crypt base columnar (CBC) cells intercalated between Paneth cells in the small intestine¹⁶. These cells were recently shown to express the leucine-rich repeat-containing G-protein-coupled receptor 5 (Lgr5) along the entire gastrointestinal tract¹⁵. Supporting both the cp4 model and the Cairns hypothesis was the finding that tritiated thymidine incorporated during post-irradiation crypt regeneration could be retained for at least 18 days in cycling cells, preferentially localized above Paneth cells⁴. Furthermore, some recently discovered multipotent stem cells expressing the B-cell-specific Maloney Murine Leukaemia Virus integration site 1 (Bmi-1) ring finger nuclear protein were reported to be preferentially located around cp4 (ref. 17). However, Bmi-1 is not expressed in the distal half of the small intestine and in the colon¹⁷, and its expression was later shown to overlap with that of several CBC stem-cell markers, including Lgr5 (refs 18, 19). Hence, Bmi-1 and Lgr5 may mark overlapping cell populations. As CBC cells were shown to generate all epithelial lineages along the entire gastrointestinal tract, both *in vivo*^{14,15,20} and in culture²¹, we decided to investigate their mode of CS in both the intact and regenerating intestinal epithelium of adult mice, and compare it with that of the neighbouring cp4 cells. Here we demonstrate that, unlike what is commonly thought, cp4 cells are

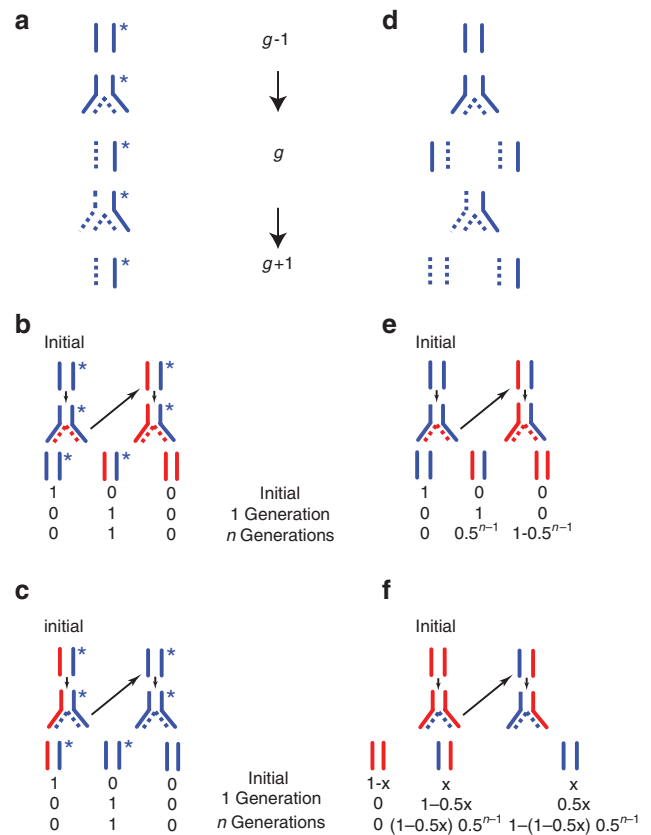


Figure 1 | Chromosome fates and labelling in the stem-cell lineages

depending on the mode of chromosome segregation (CS). Solid lines represent parental template DNA strands and dotted lines represent newly synthesized strands. **(a)** Nonrandom chromosome segregation (CS) implies an asymmetric mode of cell division generating stem and non-stem daughter cells. The ‘immortal’ DNA template (*) inherited by a stem cell at generation $g-1$, is again inherited by the stem progeny at generations g and $g+1$. **(b)** In the case of nonrandom CS, and no matter how long the labelling period (1 to n generations), only one strand of each chromosome can be labelled (red) in stem cells as they systematically inherit the chromatid containing the initially unlabelled ‘immortal’ strand (blue*). **(c)** In the case of nonrandom CS, stem cells replicating their DNA in the absence of label are unlabelled from the first generation. **(d)** On the contrary, because random CS implies that both templates of generation $g-1$ can be transmitted to generation $g+1$ with equal probability, the proportion of labelled chromosomes steadily increases with each generation of stem cells produced in the presence of label **(e)**, and steadily decreases with each generation of stem cells produced during the chase period **(f)**.

not long-term label-retaining cells (LRCs) and our results suggest that both CBC and cp4 cells segregate their chromosomes randomly under homeostatic conditions and when recovering from a lesion.

Results

Defining the experimental approach. The best available test capable of discriminating a nonrandom from a random CS mode involving all chromosomes is the DNA label dilution assay. The nonrandom CS mode has two implications, which are key to understanding the simulations described below. First, only one DNA strand per chromosome can be labelled in stem cells, regardless of the length of the labelling period (Fig. 1a,b). Second, all chromosomes of all stem cells should become unlabelled as soon as the ‘immortal’ strands serve as templates in the absence of label (Fig. 1c, 1 generation). In contrast, both chromosome strands can be labelled in the case

of random CS and the probability that all chromosomes become unlabelled after the first division occurring during the chase is negligible (Fig. 1d–f). Therefore, the fraction of labelled CBC cell nuclei detected after two or more cell cycles under both scenarios should mainly depend on two parameters, namely, the length of the labelling period relative to that of the cell cycle, and the threshold level of detection, hence explaining the importance of determining the cytokinetic parameters of CBC cells in order to optimize the labelling protocol and simulate the respective rates of decay in the labelling ratio under both scenarios.

Cytokinetic parameters of CBC cells. To determine the average duration of the CBC cell cycle, we subjected adult mice to a protocol of cumulative CldU labelling²². In an asynchronously distributed cell population, the proportion of CldU⁺ cells scored after the first CldU pulse should reflect the length of the S phase (Fig. 2a, solid black line) relative to the total length of the cell cycle (dotted circle)²². Each new pulse of CldU labels cells that have entered the S phase since the last injection (red line). As a result, the fraction of labelled cells increases with time and a number (n) of pulses are needed to label all cycling cells. We then performed immunohistochemistry (IHC) on thin paraffin sections of rolled intestine, using Ki67 as a marker of cycling cells and phospho (ser10)-histone H3 to evaluate the percentage of cells in the G2/M transition. For cell scoring, we selected cells flanked by Paneth cells in longitudinal sections featuring part or all of the crypt lumen²³. Analysis of the relationship between the ratio of CldU⁺ Ki67⁺ CBC cells to the total number of Ki67⁺ CBC cells and time revealed a linear increase in labelled CBC cells until CldU labelling became saturated (Fig. 2b). It follows that CBC cells represent a single population in terms of S phase (9.4 h), G2/M phase (3.4 h) and cell cycle duration (28.5 h).

CBC and cp4 cells segregate chromosomes randomly. In addition to the cell cycle length and the CS mode (random versus nonrandom), the rate of decrease in the proportion of labelled cell nuclei should also depend on the rate of migration of the cell progeny out of the stem compartment, and the threshold of label detection δ . We built a mathematical model that integrates all the parameters cited above, the predictions of which are shown in Fig. 3a–d (see Methods). According to this model, nonrandom CS should easily be distinguished from random CS by examining the time-dependent decrease in the proportion of CldU⁺ CBC cells, starting from a fully labelled cell population, and provided that δ is equal or inferior to 0.25 labelled strand per chromosome (see Methods). This threshold allows the detection of cells in which at least one-fourth of the chromosomes contain one labelled strand (green curves in Fig. 3c,d). Indeed, if these conditions are met, very few labelled nuclei should be detected in crypt bottoms after three days of chase (> 2 generations) in case of nonrandom CS (Fig. 3c), whereas over 90% labelled CBC cell nuclei should still be detected in the case of random CS (Fig. 3d).

Using the timeframe predicted by our model, we designed a DNA labelling–dilution experiment. To avoid the stress inflicted by repeated injections, adult mice were administered one initial injection of CldU, followed by CldU in drinking water for 2.5 days (~2 generations), a duration that we knew was more than sufficient to label 100% cells. Using CldU instead of BrdU allowed us to pulse mice with IdU at the end of the labelling period, and we found that the proportion of CBC cells transiting through the DNA S phase (36%) was unchanged compared with untreated control mice. For cell scoring, we distinguished Ki67⁺ CBC cells flanked by Paneth cells from the first Ki67⁺ located above the last Paneth cells, at the average +4 position (cp4) from the base²⁴. The results show that after 2 days of chase, which exceeds the time required to complete the first round of cell divisions in the absence of label, a CldU signal is detected in nearly 100% of CBC and cp4 cells (Fig. 3e and

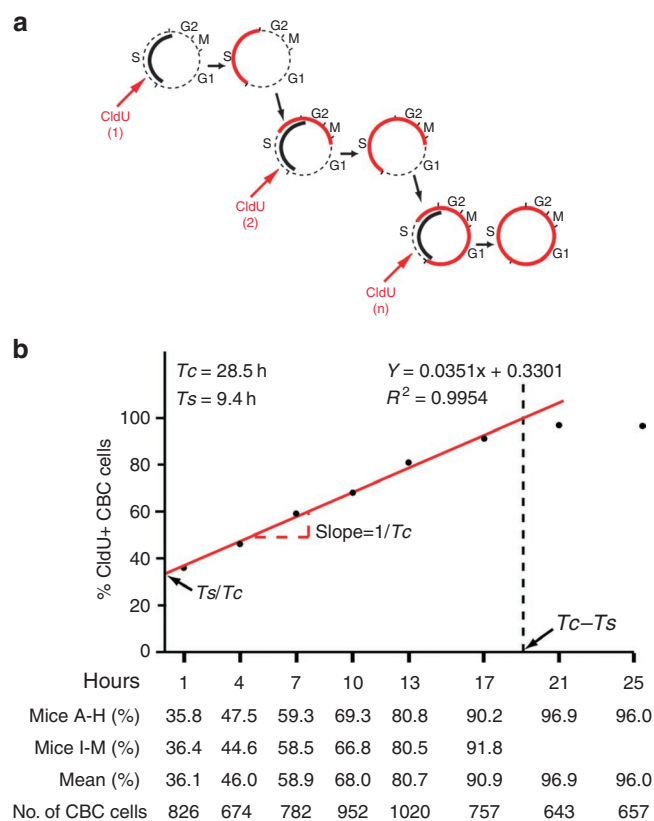


Figure 2 | Cytokinetic parameters of CBC cells. (a) Schematic representation of the protocol of cumulative CldU labelling. The solid black line represents the length of the S phase relative to the other phases of the cycle. The solid red line represents the distribution of cells within the cycle, which have been labelled with CldU. (b) Kinetics of cumulative CldU incorporation in CBC stem cells. We performed CldU immunohistochemistry on thin paraffin sections of rolled intestine, using Ki67 as a marker of cycling cells and phospho (ser10)-histone H3 to evaluate the percentage of cells in the G2/M transition. Analysis of the relationship between the ratio of CldU⁺ Ki67⁺ CBC cells to the total number of Ki67⁺ CBC cells and time revealed a linear increase in labelled CBC cells until CldU labelling became saturated. Data points represent mean percentages of Ki67⁺ CldU⁺ CBC cell nuclei from two mice (A–F and I–M), except for the 21-h and 25-h time points (mice G,H). The average length of the cell cycle ($T_c = 28.5$ h) is given by the slope of the regression line ($1/0.0351$), and that of the DNA synthetic phase ($T_s = 9.4$ h) is derived from the y-intercept $T_s/T_c = 0.3301$. $T_c - T_s$ is the time needed to reach the maximum labelling value. The length of the G2/M phase was calculated using the following equation: $T_{G2/M} = (\% \text{ Histone H3-P}^+ \text{ CBC cells} \times T_c)$.

Supplementary Table S1). By day 3, all cells have been through two successive divisions in the absence of label and yet a clear signal can be detected in 94–97% of CBC cells and 93–96% of cp4 cells (Fig. 3e). These experimental data are incompatible with the non-random CS segregation scenario and are instead in broad agreement with the predictions of the random CS model, with a detection threshold of 0.09 labelled strand per chromosome (Fig. 3f, orange line, $r^2 = 0.74$). The observed kinetics of label dilution nevertheless suggests that modelling the detection of the label using a single parameter (δ)²⁵ is oversimplified (orange line in Fig. 3g). Realistically, other factors than the amount of label in the cell should influence the results, such as the orientation of the cell relative to the crypt section. Therefore, the data are better explained with a two-parameter model accounting for some degree of randomness in the decision of calling a cell CldU⁺ (Fig. 3f,g, blue lines, $r^2 = 0.96$).

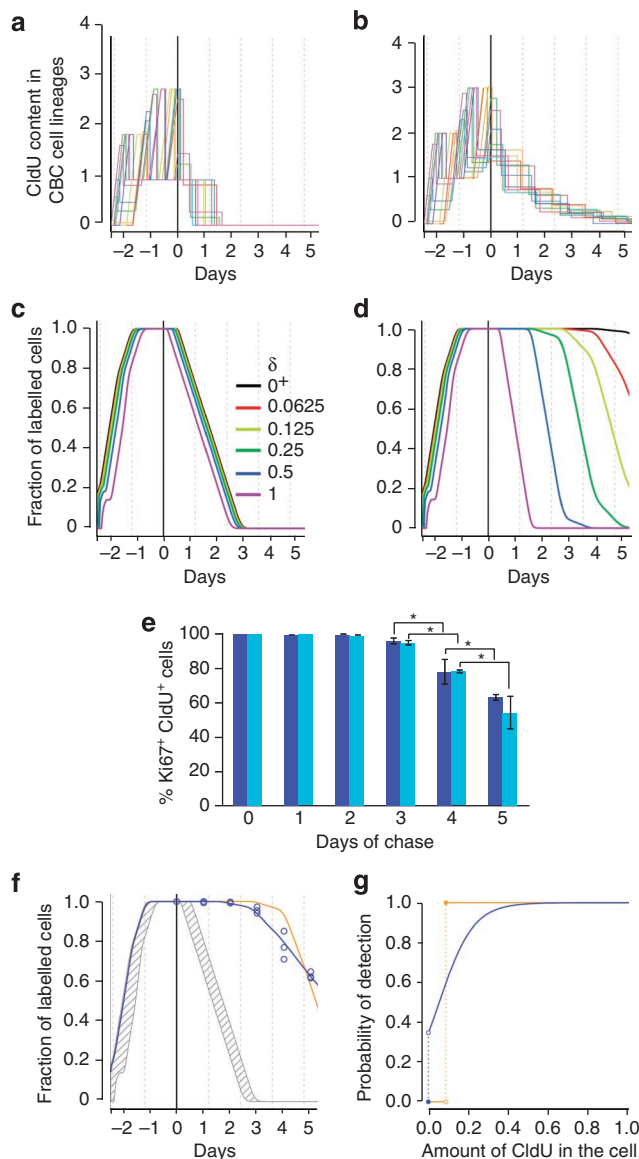


Figure 3 | Analysis of the CBC and cp4 cell-modes of chromosome segregation during a labelling-dilution experiment. (a,b) Simulation of the CldU content fluctuations in ten individual CBC cell lineages (one colour per lineage) during a pulse-chase experiment. A CldU content beyond 2 cannot be reached before the beginning of the S phase, when ploidy increases. (a) In the case of nonrandom CS, the number of labelled DNA strands per chromosome (CldU content) should substantially increase after 2 days of labelling (-2 to 0) but should decrease abruptly within 2–3 days of chase. (b) In the case of random CS, the CldU content should decay at a slower rate and labelled DNA should still be present after 5 days of chase. (c,d) Mathematical modelling of the DNA label dilution kinetics. The proportion of CldU⁺ cells depends on the minimal amount of CldU-labelled DNA that can be detected by immunofluorescence. The parameter δ corresponds to this threshold in our simple detection model²⁵. Predictions are represented for the nonrandom and random CS scenarios and six thresholds of detection ranging from $\delta=1$ (CldU content $>$ one labelled strand per chromosome) to the limit when δ approaches 0 (black line perfect detection). (c) In the case of nonrandom CS, CldU should be lost from the stem cell compartment after 3 days of chase, whereas in the case of random CS (d), CldU should still be present and detected provided that $\delta < 0.5$. (e) Kinetics of CldU dilution in the fully labelled CBC (dark blue) and cp4 (turquoise blue) cell populations. Mean $\pm$ s.d. ($n=3$). One initial injection of CldU was followed by CldU in drinking water for 60 h ($> 2T_c$). Mice were killed at the end of the labelling period and every day for 5 consecutive days. The differences between day 4 and day 3 and between day 5 and day 4 are significant ($P < 0.05$, Fisher's exact test). (f) Model fit for CBC cells. The observed kinetics (data represented by open circles) is compared with the nonrandom CS scenario (grey hatched area) and the random CS scenario after estimating the label-detection threshold (orange line) and a more sophisticated two-parameter model of label detection (blue line). The fraction of variance (r^2) explained by this second model reaches 0.96 compared with 0.74 with the simple model. (g) Modelling the probability of label detection as a function of the CldU content. The orange line represents the probability of detection, using the simple model estimated on the experimental data ($\delta=0.088$). The second model represented by the blue line accounts for randomness in the detection process. It involves two parameters, δ and ϵ , estimated from the experimental data ($\delta=0.055$ and $\epsilon=0.085$).

As illustrated in Fig. 4, the CldU-labelling pattern changed from bright and homogeneous at the end of the labelling period (T_0) to spotty after 24 h in the absence of CldU. The largest CldU-labelled spots (Fig. 4i–k) likely correspond to clusters of major and minor satellite repeats of centromeres (also called chromocenters) that remain associated through interphase and S phase, and dissociate only in prophase^{26,27}. Between 1 and 3 days of chase (Fig. 4h–j), the number but not the intensity of the spots steadily decreased, attesting the progressive loss of labelled chromosomes through successive divisions. Taken together, the results indicated that in the unlesioned epithelium of adult mice, CBC stem cells as well as cp4 cells segregate most if not all their chromosomes randomly.

CBC and cp4 cells are not LRCs. There are several reports on the capacity of stem cells of the adult small intestinal epithelium to retain DNA label for long periods of time during the period of regeneration that follows a lesion^{4,28–30}. Although a clear correlation between cell position and differentiation or proliferation of LRCs was not evident in these studies, long-term label retention has been considered until now as the exclusive property of stem cells in the epithelium of the small intestine^{4,28–30}. We therefore decided to perform a label retention assay that includes an IHC technique (see Methods)

capable of discriminating putative cycling or quiescent undifferentiated LRCs from post-mitotic differentiated LRCs present in the crypt base (Paneth, entero-endocrine and tuft cells¹²).

To be as discriminatory as possible in the detection of stem cells, we used Lgr5-EGFP mice¹⁵ in which GFP is expressed in 20–30% of the crypts and cycling CBC cells sandwiched between Paneth cells are GFP^{high}¹⁵. Lgr5-EGFP mice were irradiated with a dose of 6 Gy of Cs¹³⁷ to the whole body, after which mice were administered BrdU in drinking water for 5 days. Control and irradiated mice were killed 2, 5, 14 and 23 days post irradiation. Both GFP⁺ and GFP[−] CBC cells were counted and the data pertaining to this experiment can be found in Supplementary Tables S2–S5. The efficiency of irradiation was attested at day 2 by a significant decrease in the number of GFP⁺ crypt sections (Fig. 5a) and in the number of Ki67⁺ CBC cells (Fig. 5b), whereas the Ki67⁺ cp4 cell score was unchanged (Fig. 5b). The complete recovery of GFP expression and CBC cells was observed 5 days post irradiation (Fig. 5a,b). In addition, the CBC and cp4 cell populations were fully labelled as in control mice (Fig. 5c). Fluorescent photomicrographs acquired at low as well as high magnification illustrate the radiation-induced changes in the CBC cell population observed during the BrdU-labelling period (Fig. 6). As predicted by the model of random CS (Supplementary Fig. S1), BrdU⁺ CBC and BrdU⁺ cp4 cells could still be detected after 9 days of chase but not at day 23 after 18 days of chase (Fig. 7). After 9 days of chase (day 14 post-irradiation), two types of BrdU⁺

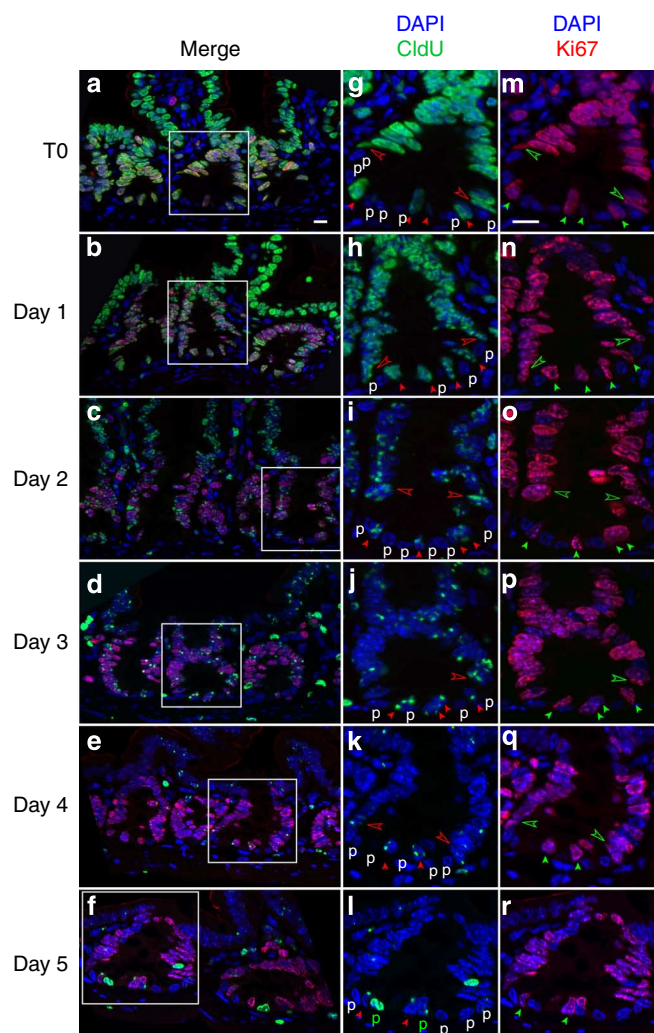


Figure 4 | CldU and Ki67 immunostaining of the CBC and cp4 cell populations. Mice were killed at the beginning of the chase period (T0) and every day for 5 consecutive days. **(a–f)** Merged lower power views of ileal crypt sections stained for CldU (green) and Ki67 (red), and counterstained with 4′-6-diamidino-2-phenylindole (DAPI) on **(a)** day 0, **(b)** day 1, **(c)** day 2, **(d)** day 3, **(e)** day 4 and **(f)** day 5. **(g–r)** Enlarged single-channel views of the crypt sections present in the insets. **(g–l)** CldU staining on **(g)** day 0, **(h)** day 1, **(i)** day 2, **(j)** day 3, **(k)** day 4 and **(l)** day 5. Note that the number but not the intensity of the spots steadily decreased from T0 **(g)** to day 5 **(l)**, attesting the progressive loss of labelled chromosomes through successive divisions. **(m–r)** Ki67 staining on **(m)** day 0, **(n)** day 1, **(o)** day 2, **(p)** day 3, **(q)** day 4 and **(r)** day 5. Closed and open arrowheads point to CBC cell and cp4 cell nuclei, respectively. P, unlabelled (white) and CldU-labelled (green) Paneth cell nuclei. Scale bars, 10 μ m.

nuclei were present in the stem compartment and could easily be distinguished on the basis of the intensity and staining pattern. Both in GFP⁺ and GFP[−] crypts, all BrdU^{high} nuclei without exception belonged to post-mitotic lineage⁺ cells, whereas BrdU^{low} nuclei were predominantly found in CBC or cp4 cells at a frequency expected in case of random CS after eight cell generations (Fig. 7). In line with their relatively longer turnover time (8 weeks)³¹, Paneth cell nuclei were over-represented in the BrdU^{high} group, compared with entero-endocrine and tuft cells, with turnover times of 4–10 days^{32,33} (Fig. 7 and Supplementary Table S5).

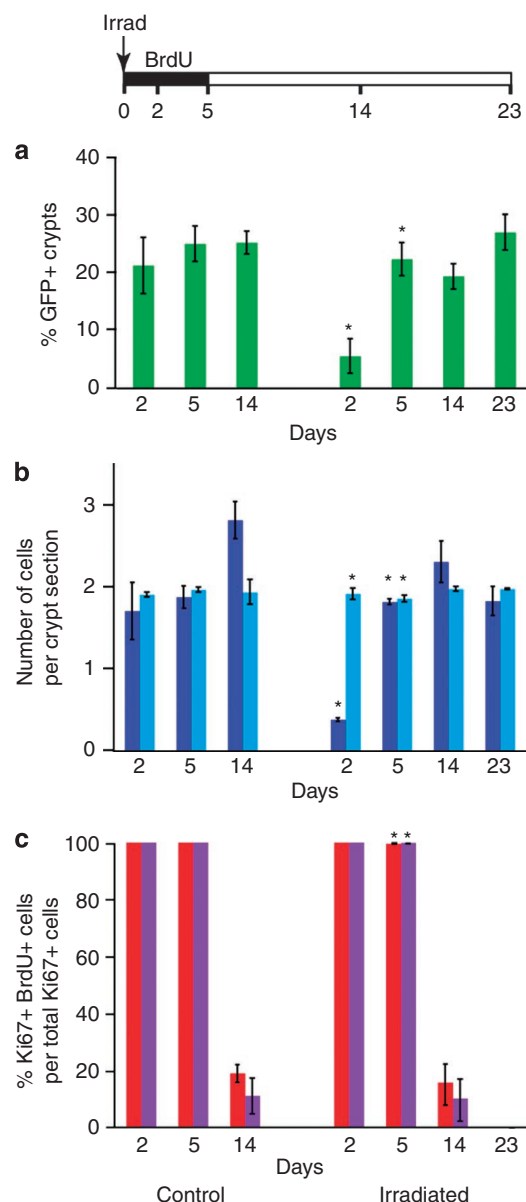


Figure 5 | Quantification of CBC cell depletion, followed by recovery in the ileum of irradiated mice. Lgr5-EGFP mice were irradiated (or not) with a dose of 6 Gy of Cs¹³⁷ to the whole body. They were then administered BrdU in drinking water for 5 days and killed 2, 5, 14 and 23 days post irradiation. **(a)** Temporary decrease followed by recovery in GFP expression in the crypt epithelium of irradiated mice. **(b)** Transient irradiation-induced depletion in CBC cells (dark blue) but not cp4 cells (turquoise blue) 2 days post irradiation. **(c)** Similar rates of BrdU incorporation (day 5) and BrdU dilution (day 14) in CBC cells (red) and cp4 cells (purple) in irradiated versus control mice. Note that 2 days post irradiation, 100% of surviving CBC cells were Ki67⁺ BrdU⁺. Data are presented as mean with error bars. Error bars with * ($n=3$) correspond to standard deviations and bars without * ($n=2$) indicate the range. The values obtained from distinct mice for the same time point (see Supplementary Tables S2–S5) were compared using Fisher's exact test.

Discussion

Before Lgr5 was coined as a specific marker of intestinal epithelial stem cells¹⁵, CBC cells were known to be more radiosensitive, to cycle more slowly and not to display the circadian rhythm evident for cells at higher positions, suggesting that they represent a distinct

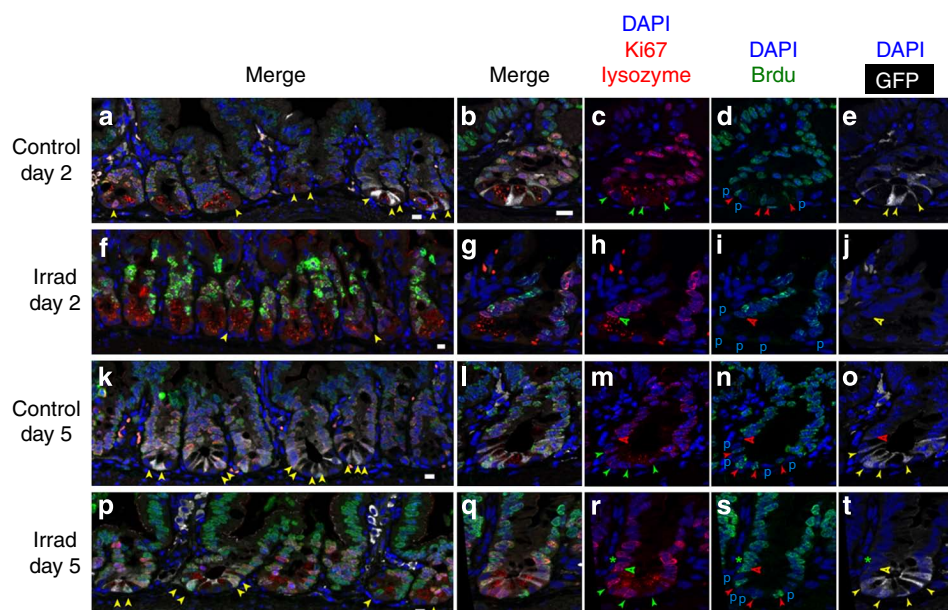


Figure 6 | Immunofluorescence CBC cell BrdU-labelling at 2 and 5 days post irradiation. Control (a–e,k,o) and irradiated (f–j,p–t) Lgr5-EGFP mice were continuously exposed to BrdU administered in drinking water and killed after 2 and 5 days. Ileum sections were simultaneously stained for nuclear BrdU (green), nuclear Ki67 (red), cytoplasmic lysozyme (red) and cytoplasmic green fluorescent protein (GFP, white). Nuclei were counterstained with DAPI (blue). Closed and open arrowheads point at CBC and to cp4 cells, respectively. (a) Merge picture of neighbouring GFP⁺ and GFP[−] crypt sections. (b) Merge picture of a crypt section with four CBC cells staining positive for Ki67 (c, green arrowheads) BrdU (d, red arrowheads) and GFP (e, yellow arrowheads). (f) Merge picture of crypt sections devoid of CBC cells. (g) Merge picture of a single crypt section. The cp4 cell of this crypt (open arrowhead) is Ki67⁺ (h), BrdU⁺ (i) and GFP[−] (j). (k) Merge picture of GFP⁺ crypt sections. (l) Merge picture of a single crypt section with four CBC cells and one cp4 cell, all of which are Ki67⁺ (m), BrdU⁺ (n) and GFP⁺ (o). (p) Merge picture of a succession of GFP⁺ and GFP[−] crypt sections. (q) Merge picture of a single GFP⁺ crypt section with four CBC cells, all of which are Ki67⁺ (r), BrdU⁺ (s) and GFP⁺ (t). Note the presence of a Ki67⁺ BrdU⁺ GFP[−] cp4 cell and the BrdU⁺ Ki67[−] nucleus (*) of a post-mitotic cell sandwiched between Ki67⁺ BrdU⁺ progenitor nuclei. P, Paneth cell nuclei. Scale bar, 10 μm. DAPI, 4',6-diamidino-2-phenylindole.

cell population (reviewed in refs 12, 34). Here we confirmed that, at least in the small intestine, CBC cells do represent a single population in terms of proliferation kinetics and in terms of radiosensitivity. The return to a normal number of CBC cells within 5 days post irradiation highlights the outstanding capacity of the crypt to respond to acute damage by recruiting surviving CBC cells or progenitors of the transit-amplifying zone so as to replenish the stem-cell pool. The results of our label dilution and label retention assays suggest that intestinal epithelial stem cells segregate most if not all their chromosomes randomly, both under homeostatic conditions and when recovering from a lesion. In addition, we demonstrate that unlike what is commonly thought, all label-retaining cells found in the lower half of the crypts 18 days after irradiation are post-mitotic, terminally differentiated cells, and that cycling cp4 cells are not LRCs. The observed longer retention of BrdU in Paneth cell nuclei compared with CBC cells and other types of differentiated cells was expected considering the slow rate of Paneth cell replacement compared with the other cell types.

With the type of pulse-chase experiments used to test the immortal strand hypothesis in our study and in previous studies, it is difficult to answer the question of whether one or a small number of chromosomes segregate asymmetrically, a potential mechanism proposed by others to regulate lineage cell commitment^{3,8}. This question was recently addressed in the unlesioned mouse colon epithelium, using the technique of chromosome orientation fluorescent *in situ* hybridization to follow the fate of parental template strand inheritance in mitotic cell pairs³⁵. Because the orientation of major satellite DNA relative to telomeric DNA is fixed, probes hybridized to major satellite repeats were used to arbitrarily define a Watson (major A-rich satellite) and a Crick (major T-rich satellite) DNA template. The results revealed a higher frequency of daughter-cell pairs with asymmetry, than predicted by simulated random segregation. This

was interpreted as a possible indication that nonrandom CS contributes to cell-fate decisions³⁵. However, CS could, in theory, be non-random with respect to chromosome orientation (that is, all Crick or all Watson strands inherited by the same daughter cell) and yet, cell fate be independent of this phenomenon. Conversely, if retention of the same parental template strand of one or several chromosomes is linked to cell fate, and if this mechanism is independent of DNA orientation, the current chromosome orientation fluorescent *in situ* hybridization method could not possibly detect it.

We have examined some possible scenarios that deviate from the purely random CS model and consider the possibility of nonrandom CS concerning a fraction of chromosomes, or occurring in a fraction of the cell population (Supplementary Fig. S2 and Supplementary Discussion). Whereas rare occurrences of either scenario cannot be completely ruled out, we reached the conclusion that to be compatible with our data and the sensitivity of our technique (detection in the transit amplifying zone after 3 days of chase, that is, after five rounds of division and a 32-fold dilution of the label), the fraction of cells or chromosomes subjected to nonrandom CS would have to be small in order to escape detection in the post-irradiation label retention assay.

Our post-irradiation label retention assay is the first of the kind in which a clear correlation between cell position, differentiation and proliferation for each label-retaining and each non-LRC could be established among the hundreds of cells examined. A significant point of concern regarding previous studies was the lack of reference to Ki67 or PCNA staining, which could have been used to unambiguously establish the proliferative status of LRCs^{4,28–30}. In one instance, the close proximity of labelled and unlabelled nuclei was used to distinguish mitotic pairs from unrelated nuclei in whole-mount preparations³⁰. However, in our opinion, deciding whether two nuclei form a mitotic pair or belong to neighbouring cells are difficult distinctions to make *in situ*, when based only on the

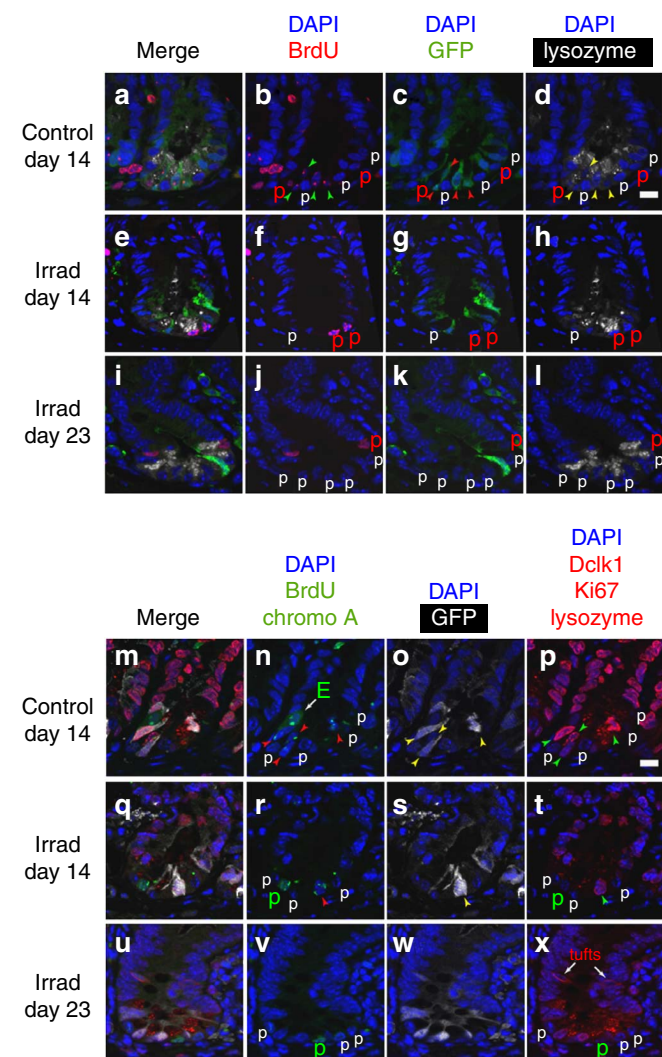


Figure 7 | CBC and cp4 cells of irradiated Lgr5-EGFP mice are not label-retaining cells. Control and irradiated Lgr5-EGFP mice were continuously exposed to BrdU for 5 days and killed 14 days (9 days of chase) and 23 days (18 days of chase) after the beginning of the treatment. Ileum sections were either simultaneously stained for BrdU, GFP and lysozyme (**a–l**) or BrdU, GFP, lysozyme, chromogranin A, Dcl1 and Ki67 (**m–x**). (**a–d**) Single crypt section from a control mouse. (**a**) Merge. (**b**) BrdU⁺ CBC cell nuclei (green arrowheads) sandwiched between BrdU⁺ and BrdU[−] Paneth cell nuclei (p). The CBC cells of this crypt section are all GFP⁺ (arrowheads in **c**) and the Paneth cells are lysozyme⁺ (**d**). (**e–h**) Single crypt section from an irradiated mouse killed 14 days post irradiation. (**e**) Merge. The two nuclei that stain positive for BrdU (**f**) and negative for GFP (**g**) belong to Paneth cells with cytoplasmic lysozyme granules (**h**). (**i–l**) Single crypt section from an irradiated mouse killed 23 days post irradiation. (**i**) Merge. The BrdU⁺ nucleus (**j**) belongs to a GFP[−] (**k**) lysozyme⁺ (**l**) Paneth cell. (**m–p**) Single crypt from a control mouse. (**m**) Merge. (**n**) The small arrow points at an entero-endocrine cell (E) with nuclear BrdU and cytoplasmic chromogranin A staining. Arrowheads point at BrdU⁺ nuclei that belong to GFP⁺ (**o**), Ki67⁺ (**p**) CBC cells. (**q–t**) Single crypt section from an irradiated mouse killed 14 days post irradiation. (**q**) Merge. (**r**) Two BrdU⁺ nuclei, one of which belongs to a GFP⁺ cell (arrowhead in **s**) and the other to a lysozyme⁺ Paneth cell (**t**). (**u–x**) Single crypt section from an irradiated mouse killed 23 days post irradiation. (**u**) Merge. (**v**) The BrdU⁺ nucleus belongs to a GFP[−] (**w**) lysozyme⁺ (**x**) Paneth cell. White arrows in (**x**) point at BrdU[−] Dcl1⁺ tuft cells. Scale bar, 10 μ m. DAPI, 4'-6-diamidino-2-phenylindole.

distance between nuclei. In contrast, the distance between nuclei is a valid parameter when looking at dividing cells *in vitro*^{6–8}. We assume that the differences listed above between our experimental setup and that of previous studies could explain why we reached substantially different conclusions. Importantly, our results agree with the conclusion of two recent studies in which monoclonal conversion of intestinal epithelial crypts was analysed at single stem cell resolution^{36,37}. In both studies, clone-size expansion or contraction was observed before monoclonal conversion was complete, and this early dynamics was found to have the ‘scaling’ behaviour expected for a scenario of neutral drift within a population of stem cells. According to their data the authors state that if the population size is 14–18, a figure they believe to be likely, then symmetric divisions in which both stem cell daughters have the same fate should be the rule rather than the exception^{36,37}, and this argues against long-term template-strand retention in the stem-cell lineage.

In conclusion, similar to what was previously demonstrated for the haematopoietic system and the skin^{25,38,39}, both tissues undergoing continuous turnover during adult life and engaged in wound repair in response to injury, our study does not support the chromosome-sorting model as a major mechanism of genome integrity preservation in the adult intestinal epithelium.

Methods

Mice and chemicals. Adult (13- to 16-week old) male C57BL/6 mice were purchased from Charles River Laboratories (Arbresle, France) and housed under standard laboratory conditions on a light–dark cycle of 0800 to 2000 hours of light, with free access to food and water. Lgr5-EGFP-IRES-creERT2 (Lgr5-EGFP) mice were provided by Dr Eduard Batlle at the Institute for Research in Biomedicine of Barcelona, Spain. Experiments were performed according to the principles of laboratory animal care in compliance with the European and French law. Chemicals and antibodies are listed in Supplementary Methods and Supplementary Table S6.

Measurement of CBC cell cytokinetic parameters. Adult C57BL/6 mice were administered five successive intraperitoneal injections of CldU (35 mg per kg of body weight) in Hanks buffered saline solution at 3- or 4-h intervals for a total period of 24 h. The first and last injections were administered at 0700 h. Two mice were killed 1 h after each CldU injection. The small intestine was flushed with phosphate-buffered saline (PBS) and 10-cm long segments were everted on 3-mm OD wooden skewers before fixation in 3% paraformaldehyde (PFA) for 3 h, followed by overnight incubation in ice-cold PBS. Intestines were released from the skewers by a longitudinal incision and rolled up using wooden toothpicks. Rolls were then processed for IHC on paraffin sections, according to standard protocols⁴⁰. Thin (1–2 μ m) sections were stained for CldU and Ki67, and the proportions of Ki67⁺ CldU⁺ and CldU[−] CBC cell nuclei were determined from a collection of fluorescence images acquired with a Zeiss Axio Imager microscope equipped with a $\times 40$ objective and an ApoTome slider.

CldU-labelling-dilution and CldU-IdU double labelling assays. Adult C57BL/6 mice were administered one initial intraperitoneal injection of 32.56 mM CldU in Hanks buffered saline solution (85.5 mg per kg of body weight), followed by CldU in drinking water (1 mg ml^{−1}) for 60 h, a duration that is more than sufficient to label 100% CBC cells. The lack of cytotoxicity induced by continuous CldU exposure was attested, after one IP injection of 32.56 mM IdU (115.3 mg kg^{−1} of body weight), by the unchanged level of nuclear cleaved caspase 3, compared with untreated control mice (not shown) and the unchanged proportion (36%) of CBC cells transiting through the DNA S phase. This procedure was adapted from a previously described protocol⁴¹. Antigen retrieval in 10 mM sodium citrate buffer, pH 6.0 for 20 min, was not followed by incubation with HCl (2N). Incubation with the primary antibodies overnight at 4 °C was followed by incubation with 0.5 M NaCl for 6 min to eliminate nonspecific binding of the anti-IdU antibody. The anti-Ki67 and anti-lysozyme antibodies were then reapplied for 1 h at room temperature before rinsing in PBS-Tween and incubation for 1 h at room temperature with secondary antibodies⁴¹.

BrdU-labelling during post-irradiation crypt regeneration. Lgr5-EGFP mice were administered one initial injection of BrdU (60 mg kg^{−1}) followed (or not) by irradiation with a dose of 6 Gy of Cs¹³⁷ to the whole body (3.5 Gy min^{−1}), with a therapeutic irradiator. They were then administered BrdU in drinking water (1 mg ml^{−1}) for 5 days and groups of control and irradiated mice were killed 2, 5, 14 and 23 days post irradiation.

CBC and cp4 cell scoring procedure. For cell scoring, we distinguished Ki67⁺ CBC cells flanked by Paneth cells from the first Ki67⁺ located above the last Paneth cells, at the average +4 position (cp4) from the base²⁴. To facilitate CBC cell spot-

ting, we restricted our analysis to the ileum, at which the most abundant Paneth cell and CBC cell populations are found. As all CBC cell nuclei stain positive for Ki67, to determine the fraction of CBC cells with labelled nuclei (BrdU, CldU or IdU), the following criteria were applied during counting: CBC cell nuclei located between Paneth cells (not above) were scored; Ki67⁺/BrdU⁺ and Ki67⁺/BrdU⁻ nuclei were scored in well-oriented crypt sections, defined as crypts that contain a single layer of epithelial cells surrounding their lumen, and those that are adjacent to the muscularis mucosa. Two independent persons performed scoring.

Scoring procedure for long-term BrdU-retaining cell nuclei. We performed IHC on thin paraffin sections of rolled intestine, using antibodies directed against GFP and Ki67 (cycling cells), BrdU, and a cocktail of primary antibodies directed against lineage-specific antigens: chromogranin A for entero-endocrine cells; doublecortin-like kinase 1 (Dclk1) for tuft cells¹¹; and lysozyme for Paneth cells. A donkey anti-sheep-A488 conjugate was used to detect nuclear BrdU and cytoplasmic chromogranin A in entero-endocrine cells. A donkey anti-rabbit-Cy3 conjugate was used to detect nuclear Ki67, lysozyme⁺ granules in the cytoplasm of Paneth cells and cytoplasmic dclk1 in tuft cells. Revealing several antigens with a single conjugate was possible because of the different subcellular localizations of the antigens (nuclear/cytoplasmic, diffuse cytoplasmic/cytoplasmic granules), and because each labelled cell could be unambiguously identified by its shape and localization along the crypt–villus axis. Importantly, this procedure was absolutely required to minimize the risk of overlooking the presence of rare cycling or quiescent lineage-negative BrdU⁺ cells.

Statistical analysis. Proportions of Ki67⁺ CBC and cp4 cells labelled with CldU or BrdU obtained from distinct mice for the same timepoint were compared using Fisher's exact test. When not statistically significantly different (P -values > 0.05), counts from distinct mice were aggregated before timepoint comparison, also based on Fisher's exact test.

Stochastic modelling of the proportion of labelled CBC cells. In the CldU label dilution assay, the expected proportions of CldU⁺ CBC daughter cells under the random and nonrandom CS scenarios were estimated using Monte-Carlo simulations of stem cell lineages, according to the model we describe below. Additional information is provided in the Supplementary Methods and Supplementary Software.

The length of the cell cycle has mean $\mu = 28.5$ h and standard deviation $\sigma = 2$ h. The cell cycle is subdivided into three parts with respect to ploidy. The S phase (ploidy increases from 2 to 4) accounts for a fraction $\beta_3 = 9.4$ h/28.5 h of the cycle. The time after the S phase (ploidy = 4) accounts for $\beta_2 = 3.4$ h/28.5 h of the cycle. The time before the S phase (ploidy = 2) accounts for $\beta_1 = 1 - \beta_3 - \beta_2$ of the cycle.

As the number of CBC cells per crypt does not vary significantly during the day, production of new cells is continuously compensated by upward migration of older cells during this time. The daughter cells cannot be distinguished from one another until one of them migrates, and both have thus been included in the final count. Given that the maximum amount of time spent by each new generation should not exceed the duration of the CBC cell cycle, we modelled the two daughter cells as remaining for the same length of time in the crypt base, which is the most unfavourable situation to distinguish random from nonrandom CS in our pulse-chase experiment.

Just before the g th generation, the stem cell contains $K = 2n + 2$ chromosomes ($n = 19$) made of two DNA strands: one strand was the template, whereas the other was synthesized during generation $g - 1$. After completion of the S phase, the cell contains two copies of the K chromosomes but only one copy of each chromosome contains the template DNA strand of generation $g - 1$. At the beginning of generation $g + 1$, each daughter cell inherits one copy of the K chromosomes. For each chromosome, the copy that contains the template DNA strand of generation $g - 1$ remains in the simulated stem cell lineage with a probability P . The random and nonrandom CS scenarios compared in this study correspond to $P = 0.5$ and $P = 1$, respectively (Fig. 1a,d).

The amount of CldU incorporated into the cell DNA increases at a constant rate during the S phase and as our unit of measure, we used the average number of labelled strands per chromosome. Therefore, one unit corresponds to one labelled DNA strand per chromosome in a diploid cell (40 labelled strands in the mouse) and two units to either one labelled DNA strand per chromosome in a tetraploid cell (at the end of the S phase) or two labelled strands per chromosome in a diploid cell (80 labelled strands in both cases). Figure 3a,b illustrate the variations of CldU content in stem cell lineages simulated with this model.

The parameter δ is the amount of labelled DNA that a cell must contain to be detected and scored as CldU⁺²⁵. It is expressed in the unit of measure described above. The output of our simulations is the proportion of CldU⁺ cells detected for δ varying between 0.0625 and 1. We used the simulations to estimate δ and to investigate the fit of a more realistic model that incorporates a certain level of randomness in the detection of the label by means of a second parameter ϵ (Fig. 3f,g).

Modelling the post-irradiation BrdU dilution assay. Computations were carried out by the same procedure as that for the CldU dilution assay. As stem cells are supposed to adopt a random mode of CS during tissue regeneration before resuming

nonrandom CS, the dynamics of label incorporation were not simulated. For simplicity, we assumed that all chromosomes were fully labelled after 5 days of pulse.

References

1. Rando, T. A. The immortal strand hypothesis: segregation and reconstruction. *Cell* **129**, 1239–1243 (2007).
2. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197–200 (1975).
3. Lansdorp, P. M. Immortal strands? Give me a break. *Cell* **129**, 1244–1247 (2007).
4. Potten, C. S., Owen, G. & Booth, D. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J. Cell Sci.* **115**, 2381–2388 (2002).
5. Smith, G. H. Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands. *Development* **132**, 681–687 (2005).
6. Karpowicz, P. *et al.* Support for the immortal strand hypothesis: neural stem cells partition DNA asymmetrically *in vitro*. *J. Cell Biol.* **170**, 721–732 (2005).
7. Shinin, V., Gayraud-Morel, B., Gomes, D. & Tajbakhsh, S. Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. *Nat. Cell Biol.* **8**, 677–687 (2006).
8. Conboy, M. J., Karasov, A. O. & Rando, T. A. High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. *PLoS Biol.* **5**, e102 (2007).
9. Shibata, D. & Tavare, S. Counting divisions in a human somatic cell tree: how, what and why? *Cell Cycle* **5**, 610–614 (2006).
10. Nicolas, P., Kim, K. M., Shibata, D. & Tavare, S. The stem cell population of the human colon crypt: analysis via methylation patterns. *PLoS Comput. Biol.* **3**, e28 (2007).
11. Gerbe, F., Brulin, B., Makrini, L., Legerverend, C. & Jay, P. DCAMKL-1 expression identifies Tuft cells rather than stem cells in the adult mouse intestinal epithelium. *Gastroenterology* **137**, 2179–2180; author reply 2180–2181 (2009).
12. Bjerknes, M. & Cheng, H. Intestinal epithelial stem cells and progenitors. *Methods Enzymol.* **419**, 337–383 (2006).
13. Marshman, E., Booth, C. & Potten, C. S. The intestinal epithelial stem cell. *BioEssays* **24**, 91–98 (2002).
14. Bjerknes, M. & Cheng, H. Clonal analysis of mouse intestinal epithelial progenitors. *Gastroenterology* **116**, 7–14 (1999).
15. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007).
16. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian theory of the origin of the four epithelial cell types. *Am. J. Anat.* **141**, 537–561 (1974).
17. Sangiorgi, E. & Capecchi, M. R. Bmi1 is expressed *in vivo* in intestinal stem cells. *Nat. Genet.* **40**, 915–920 (2008).
18. Snippert, H. J. *et al.* Prominin-1/CD133 marks stem cells and early progenitors in mouse small intestine. *Gastroenterology* **136**, 2187–2194 e2181 (2009).
19. van der Flier, L. G. *et al.* Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* **136**, 903–912 (2009).
20. Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. III. Evidence from columnar, enteroendocrine, and mucous cells in the adult mouse. *Am. J. Anat.* **160**, 77–91 (1981).
21. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
22. Nowakowski, R. S., Lewin, S. B. & Miller, M. W. Bromodeoxyuridine immunohistochemical determination of the lengths of the cell cycle and the DNA-synthetic phase for an anatomically defined population. *J. Neurocytol.* **18**, 311–318 (1989).
23. Stappenbeck, T. S., Mills, J. C. & Gordon, J. I. Molecular features of adult mouse small intestinal epithelial progenitors. *Proc. Natl Acad. Sci. USA* **100**, 1004–1009 (2003).
24. Potten, C. S., Gandara, R., Mahida, Y. R., Loeffler, M. & Wright, N. A. The stem cells of small intestinal crypts: where are they? *Cell Prolif.* **42**, 731–750 (2009).
25. Kiel, M. J. *et al.* Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU. *Nature* **449**, 238–242 (2007).
26. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* **166**, 493–505 (2004).
27. Wu, R., Singh, P. B. & Gilbert, D. M. Uncoupling global and fine-tuning replication timing determinants for mouse pericentric heterochromatin. *J. Cell Biol.* **174**, 185–194 (2006).
28. He, X. C. *et al.* BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-beta-catenin signaling. *Nat. Genet.* **36**, 1117–1121 (2004).
29. Holmberg, J. *et al.* EphB receptors coordinate migration and proliferation in the intestinal stem cell niche. *Cell* **125**, 1151–1163 (2006).
30. Quyn, A. J. *et al.* Spindle orientation bias in gut epithelial stem cell compartments is lost in precancerous tissue. *Cell Stem. Cell.* **6**, 175–181 (2010).

31. Ireland, H., Houghton, C., Howard, L. & Winton, D. J. Cellular inheritance of a Cre-activated reporter gene to determine Paneth cell longevity in the murine small intestine. *Dev. Dyn.* **233**, 1332–1336 (2005).
32. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. III. Entero-endocrine cells. *Am. J. Anat.* **141**, 503–519 (1974).
33. Thompson, E. M., Price, Y. E. & Wright, N. A. Kinetics of enteroendocrine cells with implications for their origin: a study of the cholecystokinin and gastrin subpopulations combining tritiated thymidine labelling with immunocytochemistry in the mouse. *Gut* **31**, 406–411 (1990).
34. Bjerknes, M. & Cheng, H. Gastrointestinal stem cells. II. Intestinal stem cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **289**, G381–387 (2005).
35. Falconer, E. *et al.* Identification of sister chromatids by DNA template strand sequences. *Nature* **463**, 93–97 (2010).
36. Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
37. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 8222–825 (2010).
38. Waghmare, S. K. *et al.* Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells. *EMBO J* **27**, 1309–1320 (2008).
39. Sotiropoulou, P. A., Candi, A. & Blanpain, C. The majority of multipotent epidermal stem cells do not protect their genome by asymmetrical chromosome segregation. *Stem Cells* **26**, 2964–2973 (2008).
40. Bastide, P. *et al.* Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *J. Cell Biol.* **178**, 635–648 (2007).
41. Svetlova, M. *et al.* Differential incorporation of halogenated deoxyuridines during UV-induced DNA repair synthesis in human cells. *DNA Repair (Amst)* **4**, 359–366 (2005).

Acknowledgments

This work was supported by the Agence Nationale pour la Recherche (ANR-09-BLAN-0368-01), Institut National du Cancer (INCa PLBIO09-070) and Association pour la Recherche Contre le Cancer (ARC 5050). M.E. was granted a MENRT fellowship. We thank Daniel Fisher, Frederic Hollande and Ueli Schibler for helpful comments on the manuscript.

Author contributions

M.E., P.J. and C.L. conceived and designed the experiments. M.E., F.S. and P.C. performed the experiments. P.N. performed the simulations. M.E., S.L.-C. and C.L. analysed the data. S.L. and P.N. performed the statistical analysis. D.J. gave conceptual advice. C.L. and P.N. wrote the paper.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Escobar, M. *et al.* Intestinal epithelial stem cells do not protect their genome by asymmetric chromosome segregation. *Nat. Commun.* 2:258 doi: 10.1038/ncomms1260 (2011).

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

BIBLIOGRAPHIE

1. McMahon, H.T., and Boucrot, E. (2011). Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nature reviews. Molecular cell biology* 12, 517-533.
2. Woerner, S.M., Yuan, Y.P., Benner, A., Korff, S., von Knebel Doeberitz, M., and Bork, P. (2010). SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Res* 38, D682-689.
3. Montagnac, G., Sibarita, J.B., Loubery, S., Daviet, L., Romao, M., Raposo, G., and Chavrier, P. (2009). ARF6 Interacts with JIP4 to control a motor switch mechanism regulating endosome traffic in cytokinesis. *Curr Biol* 19, 184-195.
4. Duval, A., and Hamelin, R. (2003). [Replication error repair, microsatellites, and cancer]. *Med Sci (Paris)* 19, 55-62.
5. Conacci-Sorrell, M., Zhurinsky, J., and Ben-Ze'ev, A. (2002). The cadherin-catenin adhesion system in signaling and cancer. *J Clin Invest* 109, 987-991.
6. Kahn, R.A., Bruford, E., Inoue, H., Logsdon, J.M., Jr., Nie, Z., Premont, R.T., Randazzo, P.A., Satake, M., Theibert, A.B., Zapp, M.L., et al. (2008). Consensus nomenclature for the human ArfGAP domain-containing proteins. *J Cell Biol* 182, 1039-1044.
7. Arjonen, A., Alanko, J., Veltel, S., and Ivaska, J. (2012). Distinct recycling of active and inactive beta1 integrins. *Traffic*.
8. Edeling, M.A., Smith, C., and Owen, D. (2006). Life of a clathrin coat: insights from clathrin and AP structures. *Nature reviews. Molecular cell biology* 7, 32-44.
9. Montagnac, G., Echard, A., and Chavrier, P. (2008). Endocytic traffic in animal cell cytokinesis. *Curr Opin Cell Biol* 20, 454-461.
10. Grant, B.D., and Donaldson, J.G. (2009). Pathways and mechanisms of endocytic recycling. *Nature reviews. Molecular cell biology* 10, 597-608.
11. Jacob, S., and Praz, F. (2002). DNA mismatch repair defects: role in colorectal carcinogenesis. *Biochimie* 84, 27-47.
12. Margadant, C., Monsuur, H.N., Norman, J.C., and Sonnenberg, A. (2011). Mechanisms of integrin activation and trafficking. *Curr Opin Cell Biol* 23, 607-614.
13. Gillingham, A.K., and Munro, S. (2007). The small G proteins of the Arf family and their regulators. *Annu Rev Cell Dev Biol* 23, 579-611.
14. Bryant, D.M., and Stow, J.L. (2004). The ins and outs of E-cadherin trafficking. *Trends Cell Biol* 14, 427-434.
15. Izumi, G., Sakisaka, T., Baba, T., Tanaka, S., Morimoto, K., and Takai, Y. (2004). Endocytosis of E-cadherin regulated by Rac and Cdc42 small G proteins through IQGAP1 and actin filaments. *J Cell Biol* 166, 237-248.
16. Casanova, J.E. (2007). Regulation of Arf activation: the Sec7 family of guanine nucleotide exchange factors. *Traffic* 8, 1476-1485.
17. Zaanen, A., Meunier, K., Sangar, F., Flejou, J.F., and Praz, F. (2011). Microsatellite instability in colorectal cancer: from molecular oncogenic mechanisms to clinical implications. *Cell Oncol (Dordr)* 34, 155-176.
18. Wang, X., Yue, T.L., White, R.F., Barone, F.C., and Feuerstein, G.Z. (1995). Transforming growth factor-beta 1 exhibits delayed gene expression following focal cerebral ischemia. *Brain Res Bull* 36, 607-609.

19. Hempen, P.M., Zhang, L., Bansal, R.K., Iacobuzio-Donahue, C.A., Murphy, K.M., Maitra, A., Vogelstein, B., Whitehead, R.H., Markowitz, S.D., Willson, J.K., et al. (2003). Evidence of selection for clones having genetic inactivation of the activin A type II receptor (ACVR2) gene in gastrointestinal cancers. *Cancer Res* 63, 994-999.
20. Deacu, E., Mori, Y., Sato, F., Yin, J., Olaru, A., Sterian, A., Xu, Y., Wang, S., Schulmann, K., Berki, A., et al. (2004). Activin type II receptor restoration in ACVR2-deficient colon cancer cells induces transforming growth factor-beta response pathway genes. *Cancer Res* 64, 7690-7696.
21. El-Bchiri, J., Guilloux, A., Dartigues, P., Loire, E., Mercier, D., Buhard, O., Sobhani, I., de la Grange, P., Auboeuf, D., Praz, F., et al. (2008). Nonsense-mediated mRNA decay impacts MSI-driven carcinogenesis and anti-tumor immunity in colorectal cancers. *PLoS One* 3, e2583.
22. Ionov, Y., Peinado, M.A., Malkhosyan, S., Shibata, D., and Perucho, M. (1993). Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561.
23. Aaltonen, L.A., Peltomaki, P., Leach, F.S., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Powell, S.M., Jen, J., Hamilton, S.R., et al. (1993). Clues to the pathogenesis of familial colorectal cancer. *Science* 260, 812-816.
24. Peltomaki, P., Aaltonen, L.A., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Green, J.S., Jass, J.R., Weber, J.L., Leach, F.S., et al. (1993). Genetic mapping of a locus predisposing to human colorectal cancer. *Science* 260, 810-812.
25. Thibodeau, S.N., Bren, G., and Schaid, D. (1993). Microsatellite instability in cancer of the proximal colon. *Science* 260, 816-819.
26. Modrich, P. (1989). Methyl-directed DNA mismatch correction. *J Biol Chem* 264, 6597-6600.
27. Galio, L., Bouquet, C., and Brooks, P. (1999). ATP hydrolysis-dependent formation of a dynamic ternary nucleoprotein complex with MutS and MutL. *Nucleic Acids Res* 27, 2325-2331.
28. Ban, C., Junop, M., and Yang, W. (1999). Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell* 97, 85-97.
29. Welsh, K.M., Lu, A.L., Clark, S., and Modrich, P. (1987). Isolation and characterization of the *Escherichia coli* mutH gene product. *J Biol Chem* 262, 15624-15629.
30. Hall, M.C., and Matson, S.W. (1999). The *Escherichia coli* MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity. *J Biol Chem* 274, 1306-1312.
31. Dao, V., and Modrich, P. (1998). Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. *J Biol Chem* 273, 9202-9207.
32. Yamaguchi, M., Dao, V., and Modrich, P. (1998). MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J Biol Chem* 273, 9197-9201.
33. Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S.T., and Modrich, P. (2001). In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc Natl Acad Sci U S A* 98, 6765-6770.
34. Williamson, M.S., Game, J.C., and Fogel, S. (1985). Meiotic gene conversion mutants in *Saccharomyces cerevisiae*. I. Isolation and characterization of pms1-1 and pms1-2. *Genetics* 110, 609-646.

35. Reenan, R.A., and Kolodner, R.D. (1992). Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics* 132, 963-973.
36. Reenan, R.A., and Kolodner, R.D. (1992). Characterization of insertion mutations in the *Saccharomyces cerevisiae* MSH1 and MSH2 genes: evidence for separate mitochondrial and nuclear functions. *Genetics* 132, 975-985.
37. New, L., Liu, K., and Crouse, G.F. (1993). The yeast gene MSH3 defines a new class of eukaryotic MutS homologues. *Mol Gen Genet* 239, 97-108.
38. Ross-Macdonald, P., and Roeder, G.S. (1994). Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* 79, 1069-1080.
39. Marsischky, G.T., Filosi, N., Kane, M.F., and Kolodner, R. (1996). Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes & development* 10, 407-420.
40. Acharya, S., Wilson, T., Gradia, S., Kane, M.F., Guerrette, S., Marsischky, G.T., Kolodner, R., and Fishel, R. (1996). hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc Natl Acad Sci U S A* 93, 13629-13634.
41. Genschel, J., Littman, S.J., Drummond, J.T., and Modrich, P. (1998). Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J Biol Chem* 273, 19895-19901.
42. Marsischky, G.T., and Kolodner, R.D. (1999). Biochemical characterization of the interaction between the *Saccharomyces cerevisiae* MSH2-MSH6 complex and mispaired bases in DNA. *J Biol Chem* 274, 26668-26682.
43. Alani, E. (1996). The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol Cell Biol* 16, 5604-5615.
44. Strand, M., Prolla, T.A., Liskay, R.M., and Petes, T.D. (1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* 365, 274-276.
45. Kramer, W., Kramer, B., Williamson, M.S., and Fogel, S. (1989). Cloning and nucleotide sequence of DNA mismatch repair gene PMS1 from *Saccharomyces cerevisiae*: homology of PMS1 to procaryotic MutL and HexB. *J Bacteriol* 171, 5339-5346.
46. Umar, A., Buermeier, A.B., Simon, J.A., Thomas, D.C., Clark, A.B., Liskay, R.M., and Kunkel, T.A. (1996). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell* 87, 65-73.
47. Schmutte, C., Sadoff, M.M., Shim, K.S., Acharya, S., and Fishel, R. (2001). The interaction of DNA mismatch repair proteins with human exonuclease I. *J Biol Chem* 276, 33011-33018.
48. Iyer, R.R., Pluciennik, A., Burdett, V., and Modrich, P.L. (2006). DNA mismatch repair: functions and mechanisms. *Chem Rev* 106, 302-323.
49. Montecucco, A., Rossi, R., Levin, D.S., Gary, R., Park, M.S., Motycka, T.A., Ciarrocchi, G., Villa, A., Biamonti, G., and Tomkinson, A.E. (1998). DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories. *Embo J* 17, 3786-3795.
50. Lynch, H.T., Shaw, M.W., Magnuson, C.W., Larsen, A.L., and Krush, A.J. (1966). Hereditary factors in cancer. Study of two large midwestern kindreds. *Arch Intern Med* 117, 206-212.

51. Lynch, H.T., Krush, A.J., and Larsen, A.L. (1967). Heredity and multiple primary malignant neoplasms: six cancer families. *Am J Med Sci* 254, 322-329.
52. Lindblom, A., Tannergard, P., Werelius, B., and Nordenskjold, M. (1993). Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nat Genet* 5, 279-282.
53. Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., and Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* 75, 1027-1038.
54. Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., et al. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75, 1215-1225.
55. Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., et al. (1994). Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* 368, 258-261.
56. Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., et al. (1994). Mutation of a mutL homolog in hereditary colon cancer. *Science* 263, 1625-1629.
57. Bhattacharyya, N.P., Skandalis, A., Ganesh, A., Groden, J., and Meuth, M. (1994). Mutator phenotypes in human colorectal carcinoma cell lines. *Proc Natl Acad Sci U S A* 91, 6319-6323.
58. Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., et al. (1994). Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* 371, 75-80.
59. Liu, B., Nicolaides, N.C., Markowitz, S., Willson, J.K., Parsons, R.E., Jen, J., Papadopoulos, N., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., et al. (1995). Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat Genet* 9, 48-55.
60. Ishiguro, K., Kawakami, K., Maeda, K., Ishida, Y., Omura, K., and Watanabe, G. (2003). Microsatellite instability in gastric cancer is closely associated with hMLH1 hypermethylation at the proximal region of the promoter. *Int J Mol Med* 12, 603-608.
61. Grady, W.M., Rajput, A., Lutterbaugh, J.D., and Markowitz, S.D. (2001). Detection of aberrantly methylated hMLH1 promoter DNA in the serum of patients with microsatellite unstable colon cancer. *Cancer research* 61, 900-902.
62. Yamamoto, H., Itoh, F., Nakamura, H., Fukushima, H., Sasaki, S., Perucho, M., and Imai, K. (2001). Genetic and clinical features of human pancreatic ductal adenocarcinomas with widespread microsatellite instability. *Cancer research* 61, 3139-3144.
63. Park, S.J., Rashid, A., Lee, J.H., Kim, S.G., Hamilton, S.R., and Wu, T.T. (2003). Frequent CpG island methylation in serrated adenomas of the colorectum. *Am J Pathol* 162, 815-822.
64. Jass, J.R., Do, K.A., Simms, L.A., Iino, H., Wynter, C., Pillay, S.P., Searle, J., Radford-Smith, G., Young, J., and Leggett, B. (1998). Morphology of sporadic colorectal cancer with DNA replication errors. *Gut* 42, 673-679.

65. Benatti, P., Gafa, R., Barana, D., Marino, M., Scarselli, A., Pedroni, M., Maestri, I., Guerzoni, L., Roncucci, L., Menigatti, M., et al. (2005). Microsatellite instability and colorectal cancer prognosis. *Clin Cancer Res* 11, 8332-8340.
66. Lanza, G., Gafa, R., Santini, A., Maestri, I., Guerzoni, L., and Cavazzini, L. (2006). Immunohistochemical test for MLH1 and MSH2 expression predicts clinical outcome in stage II and III colorectal cancer patients. *J Clin Oncol* 24, 2359-2367.
67. Kuebler, J.P., Wieand, H.S., O'Connell, M.J., Smith, R.E., Colangelo, L.H., Yothers, G., Petrelli, N.J., Findlay, M.P., Seay, T.E., Atkins, J.N., et al. (2007). Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: results from NSABP C-07. *J Clin Oncol* 25, 2198-2204.
68. Andre, T., Boni, C., Navarro, M., Tabernero, J., Hickish, T., Topham, C., Bonetti, A., Clingan, P., Bridgewater, J., Rivera, F., et al. (2009). Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J Clin Oncol* 27, 3109-3116.
69. Zaan, A., Flejou, J.F., Emile, J.F., Des, G.G., Cuilliere-Dartigues, P., Malka, D., Lecaille, C., Validire, P., Louvet, C., Rougier, P., et al. (2011). Defective mismatch repair status as a prognostic biomarker of disease-free survival in stage III colon cancer patients treated with adjuvant FOLFOX chemotherapy. *Clin Cancer Res* 17, 7470-7478.
70. Sargent, D.J., Marsoni, S., Monges, G., Thibodeau, S.N., Labianca, R., Hamilton, S.R., French, A.J., Kabat, B., Foster, N.R., Torri, V., et al. (2010). Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol* 28, 3219-3226.
71. Zaan, A., Cuilliere-Dartigues, P., Guilloux, A., Parc, Y., Louvet, C., de Gramont, A., Tiret, E., Dumont, S., Gayet, B., Validire, P., et al. (2010). Impact of p53 expression and microsatellite instability on stage III colon cancer disease-free survival in patients treated by 5-fluorouracil and leucovorin with or without oxaliplatin. *Ann Oncol* 21, 772-780.
72. Giannini, G., Rinaldi, C., Ristori, E., Ambrosini, M.I., Cerignoli, F., Viel, A., Bidoli, E., Berni, S., D'Amati, G., Scambia, G., et al. (2004). Mutations of an intronic repeat induce impaired MRE11 expression in primary human cancer with microsatellite instability. *Oncogene* 23, 2640-2647.
73. Ham, M.F., Takakuwa, T., Luo, W.J., Liu, A., Horii, A., and Aozasa, K. (2006). Impairment of double-strand breaks repair and aberrant splicing of ATM and MRE11 in leukemia-lymphoma cell lines with microsatellite instability. *Cancer Sci* 97, 226-234.
74. Dorard, C., de Thonel, A., Collura, A., Marisa, L., Svrcek, M., Lagrange, A., Jegou, G., Wanherdrick, K., Joly, A.L., Buhard, O., et al. (2011). Expression of a mutant HSP110 sensitizes colorectal cancer cells to chemotherapy and improves disease prognosis. *Nat Med* 17, 1283-1289.
75. Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Dev* 21, 1833-1856.
76. Marcos, I., Borrego, S., Rodriguez de Cordoba, S., Galan, J.J., and Antinolo, G. (2002). Cloning, characterization and chromosome mapping of the human SMAP1 gene. *Gene* 292, 167-171.

77. Tanabe, K., Torii, T., Natsume, W., Braesch-Andersen, S., Watanabe, T., and Satake, M. (2005). A novel GTPase-activating protein for ARF6 directly interacts with clathrin and regulates clathrin-dependent endocytosis. *Mol Biol Cell* **16**, 1617-1628.
78. Natsume, W., Tanabe, K., Kon, S., Yoshida, N., Watanabe, T., Torii, T., and Satake, M. (2006). SMAP2, a novel ARF GTPase-activating protein, interacts with clathrin and clathrin assembly protein and functions on the AP-1-positive early endosome/trans-Golgi network. *Molecular biology of the cell* **17**, 2592-2603.
79. Kon, S., Tanabe, K., Watanabe, T., Sabe, H., and Satake, M. (2008). Clathrin dependent endocytosis of E-cadherin is regulated by the Arf6GAP isoform SMAP1. *Exp Cell Res* **314**, 1415-1428.
80. Tanabe, K., Kon, S., Ichijo, N., Funaki, T., Natsume, W., Watanabe, T., and Satake, M. (2008). A SMAP gene family encoding ARF GTPase-activating proteins and its implication in membrane trafficking. *Methods Enzymol* **438**, 155-170.
81. Sakakura, I., Tanabe, K., Nouki, N., Suzuki, M., Satake, M., and Watanabe, T. (2011). The carboxy-terminal region of SMAP2 directs subcellular localization as well as Arf protein specificity. *Biochem Biophys Res Commun* **404**, 661-666.
82. Yanai, N., Sato, Y., and Obinata, M. (1997). A new type-II membrane protein in erythropoietic organs enhances erythropoiesis. *Leukemia* **11 Suppl 3**, 484-485.
83. Sato, Y., Hong, H.N., Yanai, N., and Obinata, M. (1998). Involvement of stromal membrane-associated protein (SMAP-1) in erythropoietic microenvironment. *J Biochem* **124**, 209-216.
84. Meyer, C., Schneider, B., Reichel, M., Angermueller, S., Strehl, S., Schnittger, S., Schoch, C., Jansen, M.W., van Dongen, J.J., Pieters, R., et al. (2005). Diagnostic tool for the identification of MLL rearrangements including unknown partner genes. *Proc Natl Acad Sci U S A* **102**, 449-454.
85. Montagnac, G., de Forges, H., Smythe, E., Gueudry, C., Romao, M., Salamero, J., and Chavrier, P. (2011). Decoupling of activation and effector binding underlies ARF6 priming of fast endocytic recycling. *Curr Biol* **21**, 574-579.
86. Palacios, F., Price, L., Schweitzer, J., Collard, J.G., and D'Souza-Schorey, C. (2001). An essential role for ARF6-regulated membrane traffic in adherens junction turnover and epithelial cell migration. *Embo J* **20**, 4973-4986.
87. Schweitzer, J.K., Sedgwick, A.E., and D'Souza-Schorey, C. (2011). ARF6-mediated endocytic recycling impacts cell movement, cell division and lipid homeostasis. *Semin Cell Dev Biol* **22**, 39-47.
88. Hashimoto, S., Onodera, Y., Hashimoto, A., Tanaka, M., Hamaguchi, M., Yamada, A., and Sabe, H. (2004). Requirement for Arf6 in breast cancer invasive activities. *Proc Natl Acad Sci U S A* **101**, 6647-6652.
89. D'Souza-Schorey, C., and Chavrier, P. (2006). ARF proteins: roles in membrane traffic and beyond. *Nat Rev Mol Cell Biol* **7**, 347-358.
90. Chun, J., Shapovalova, Z., Dejgaard, S.Y., Presley, J.F., and Melancon, P. (2008). Characterization of class I and II ADP-ribosylation factors (Arfs) in live cells: GDP-bound class II Arfs associate with the ER-Golgi intermediate compartment independently of GBF1. *Molecular biology of the cell* **19**, 3488-3500.

91. Volpicelli-Daley, L.A., Li, Y., Zhang, C.J., and Kahn, R.A. (2005). Isoform-selective effects of the depletion of ADP-ribosylation factors 1-5 on membrane traffic. *Molecular biology of the cell* **16**, 4495-4508.
92. Kahn, R.A., and Gilman, A.G. (1986). The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein. *J Biol Chem* **261**, 7906-7911.
93. Cohen, L.A., Honda, A., Varnai, P., Brown, F.D., Balla, T., and Donaldson, J.G. (2007). Active Arf6 recruits ARNO/cytohesin GEFs to the PM by binding their PH domains. *Molecular biology of the cell* **18**, 2244-2253.
94. Weiss, O., Holden, J., Rulka, C., and Kahn, R.A. (1989). Nucleotide binding and cofactor activities of purified bovine brain and bacterially expressed ADP-ribosylation factor. *J Biol Chem* **264**, 21066-21072.
95. Pasqualato, S., Menetrey, J., Franco, M., and Cherfils, J. (2001). The structural GDP/GTP cycle of human Arf6. *EMBO Rep* **2**, 234-238.
96. Mossessova, E., Gulbis, J.M., and Goldberg, J. (1998). Structure of the guanine nucleotide exchange factor Sec7 domain of human arno and analysis of the interaction with ARF GTPase. *Cell* **92**, 415-423.
97. Peyroche, A., Antonny, B., Robineau, S., Acker, J., Cherfils, J., and Jackson, C.L. (1999). Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: involvement of specific residues of the Sec7 domain. *Mol Cell* **3**, 275-285.
98. Peters, P.J., Hsu, V.W., Ooi, C.E., Finazzi, D., Teal, S.B., Oorschot, V., Donaldson, J.G., and Klausner, R.D. (1995). Overexpression of wild-type and mutant ARF1 and ARF6: distinct perturbations of nonoverlapping membrane compartments. *J Cell Biol* **128**, 1003-1017.
99. Kawasaki, M., Nakayama, K., and Wakatsuki, S. (2005). Membrane recruitment of effector proteins by Arf and Rab GTPases. *Curr Opin Struct Biol* **15**, 681-689.
100. Shiba, T., Kawasaki, M., Takatsu, H., Nogi, T., Matsugaki, N., Igarashi, N., Suzuki, M., Kato, R., Nakayama, K., and Wakatsuki, S. (2003). Molecular mechanism of membrane recruitment of GGA by ARF in lysosomal protein transport. *Nat Struct Biol* **10**, 386-393.
101. Bonifacino, J.S., and Glick, B.S. (2004). The mechanisms of vesicle budding and fusion. *Cell* **116**, 153-166.
102. Kirchhausen, T. (2000). Three ways to make a vesicle. *Nature reviews. Molecular cell biology* **1**, 187-198.
103. Donaldson, J.G. (2003). Multiple roles for Arf6: sorting, structuring, and signaling at the plasma membrane. *J Biol Chem* **278**, 41573-41576.
104. Klein, S., Franco, M., Chardin, P., and Luton, F. (2006). Role of the Arf6 GDP/GTP cycle and Arf6 GTPase-activating proteins in actin remodeling and intracellular transport. *J Biol Chem* **281**, 12352-12361.
105. Pearse, B.M. (1975). Coated vesicles from pig brain: purification and biochemical characterization. *J Mol Biol* **97**, 93-98.
106. Pearse, B.M. (1976). Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles. *Proc Natl Acad Sci U S A* **73**, 1255-1259.
107. Rothberg, K.G., Heuser, J.E., Donzell, W.C., Ying, Y.S., Glenney, J.R., and Anderson, R.G. (1992). Caveolin, a protein component of caveolae membrane coats. *Cell* **68**, 673-682.
108. Yamada, E. (1955). The fine structure of the gall bladder epithelium of the mouse. *J Biophys Biochem Cytol* **1**, 445-458.

109. Ehrlich, M., Boll, W., Van Oijen, A., Hariharan, R., Chandran, K., Nibert, M.L., and Kirchhausen, T. (2004). Endocytosis by random initiation and stabilization of clathrin-coated pits. *Cell* **118**, 591-605.
110. Ungewickell, E.J., and Hinrichsen, L. (2007). Endocytosis: clathrin-mediated membrane budding. *Curr Opin Cell Biol* **19**, 417-425.
111. Henne, W.M., Boucrot, E., Meinecke, M., Evergren, E., Vallis, Y., Mittal, R., and McMahon, H.T. (2010). FCHO proteins are nucleators of clathrin-mediated endocytosis. *Science* **328**, 1281-1284.
112. Hawryluk, M.J., Keyel, P.A., Mishra, S.K., Watkins, S.C., Heuser, J.E., and Traub, L.M. (2006). Epsin 1 is a polyubiquitin-selective clathrin-associated sorting protein. *Traffic* **7**, 262-281.
113. Chen, H., and De Camilli, P. (2005). The association of epsin with ubiquitinated cargo along the endocytic pathway is negatively regulated by its interaction with clathrin. *Proc Natl Acad Sci U S A* **102**, 2766-2771.
114. Ramanan, V., Agrawal, N.J., Liu, J., Engles, S., Toy, R., and Radhakrishnan, R. (2011). Systems biology and physical biology of clathrin-mediated endocytosis. *Integr Biol (Camb)* **3**, 803-815.
115. Owen, D.J., and Evans, P.R. (1998). A structural explanation for the recognition of tyrosine-based endocytotic signals. *Science* **282**, 1327-1332.
116. Olusanya, O., Andrews, P.D., Swedlow, J.R., and Smythe, E. (2001). Phosphorylation of threonine 156 of the mu2 subunit of the AP2 complex is essential for endocytosis in vitro and in vivo. *Curr Biol* **11**, 896-900.
117. Conner, S.D., and Schmid, S.L. (2002). Identification of an adaptor-associated kinase, AAK1, as a regulator of clathrin-mediated endocytosis. *J Cell Biol* **156**, 921-929.
118. Krauss, M., Kukhtina, V., Pechstein, A., and Haucke, V. (2006). Stimulation of phosphatidylinositol kinase type I-mediated phosphatidylinositol (4,5)-bisphosphate synthesis by AP-2mu-cargo complexes. *Proc Natl Acad Sci U S A* **103**, 11934-11939.
119. Crowther, R.A., and Pearse, B.M. (1981). Assembly and packing of clathrin into coats. *J Cell Biol* **91**, 790-797.
120. Kirchhausen, T., and Harrison, S.C. (1981). Protein organization in clathrin trimers. *Cell* **23**, 755-761.
121. Ungewickell, E., and Branton, D. (1981). Assembly units of clathrin coats. *Nature* **289**, 420-422.
122. Fotin, A., Cheng, Y., Sliz, P., Grigorieff, N., Harrison, S.C., Kirchhausen, T., and Walz, T. (2004). Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **432**, 573-579.
123. McMahon, H.T., Wigge, P., and Smith, C. (1997). Clathrin interacts specifically with amphiphysin and is displaced by dynamin. *FEBS Lett* **413**, 319-322.
124. Danino, D., Moon, K.H., and Hinshaw, J.E. (2004). Rapid constriction of lipid bilayers by the mechanochemical enzyme dynamin. *J Struct Biol* **147**, 259-267.
125. Roux, A., Uyhazi, K., Frost, A., and De Camilli, P. (2006). GTP-dependent twisting of dynamin implicates constriction and tension in membrane fission. *Nature* **441**, 528-531.
126. Spudich, G., Chibalina, M.V., Au, J.S., Arden, S.D., Buss, F., and Kendrick-Jones, J. (2007). Myosin VI targeting to clathrin-coated structures and dimerization is mediated by binding to Disabled-2 and PtdIns(4,5)P2. *Nat Cell Biol* **9**, 176-183.

127. Krendel, M., Osterweil, E.K., and Mooseker, M.S. (2007). Myosin 1E interacts with synaptojanin-1 and dynamin and is involved in endocytosis. *FEBS Lett* **581**, 644-650.
128. Lee, D.W., Wu, X., Eisenberg, E., and Greene, L.E. (2006). Recruitment dynamics of GAK and auxilin to clathrin-coated pits during endocytosis. *J Cell Sci* **119**, 3502-3512.
129. Schlossman, D.M., Schmid, S.L., Braell, W.A., and Rothman, J.E. (1984). An enzyme that removes clathrin coats: purification of an uncoating ATPase. *J Cell Biol* **99**, 723-733.
130. Ungewickell, E., Ungewickell, H., Holstein, S.E., Lindner, R., Prasad, K., Barouch, W., Martin, B., Greene, L.E., and Eisenberg, E. (1995). Role of auxilin in uncoating clathrin-coated vesicles. *Nature* **378**, 632-635.
131. Hopkins, C.R., and Trowbridge, I.S. (1983). Internalization and processing of transferrin and the transferrin receptor in human carcinoma A431 cells. *J Cell Biol* **97**, 508-521.
132. Maxfield, F.R., and McGraw, T.E. (2004). Endocytic recycling. *Nature reviews. Molecular cell biology* **5**, 121-132.
133. Jovic, M., Sharma, M., Rahajeng, J., and Caplan, S. (2010). The early endosome: a busy sorting station for proteins at the crossroads. *Histol Histopathol* **25**, 99-112.
134. Dunn, K.W., McGraw, T.E., and Maxfield, F.R. (1989). Iterative fractionation of recycling receptors from lysosomally destined ligands in an early sorting endosome. *J Cell Biol* **109**, 3303-3314.
135. Bucci, C., Parton, R.G., Mather, I.H., Stunnenberg, H., Simons, K., Hoflack, B., and Zerial, M. (1992). The small GTPase rab5 functions as a regulatory factor in the early endocytic pathway. *Cell* **70**, 715-728.
136. Christoforidis, S., Miaczynska, M., Ashman, K., Wilm, M., Zhao, L., Yip, S.C., Waterfield, M.D., Backer, J.M., and Zerial, M. (1999). Phosphatidylinositol-3-OH kinases are Rab5 effectors. *Nat Cell Biol* **1**, 249-252.
137. Lawe, D.C., Patki, V., Heller-Harrison, R., Lambright, D., and Corvera, S. (2000). The FYVE domain of early endosome antigen 1 is required for both phosphatidylinositol 3-phosphate and Rab5 binding. Critical role of this dual interaction for endosomal localization. *J Biol Chem* **275**, 3699-3705.
138. McBride, H.M., Rybin, V., Murphy, C., Giner, A., Teasdale, R., and Zerial, M. (1999). Oligomeric complexes link Rab5 effectors with NSF and drive membrane fusion via interactions between EEA1 and syntaxin 13. *Cell* **98**, 377-386.
139. Simonsen, A., Gaullier, J.M., D'Arrigo, A., and Stenmark, H. (1999). The Rab5 effector EEA1 interacts directly with syntaxin-6. *J Biol Chem* **274**, 28857-28860.
140. Van Der Sluijs, P., Hull, M., Zahraoui, A., Tavitian, A., Goud, B., and Mellman, I. (1991). The small GTP-binding protein rab4 is associated with early endosomes. *Proc Natl Acad Sci U S A* **88**, 6313-6317.
141. Kouranti, I., Sachse, M., Arouche, N., Goud, B., and Echard, A. (2006). Rab35 regulates an endocytic recycling pathway essential for the terminal steps of cytokinesis. *Curr Biol* **16**, 1719-1725.
142. Ito, M., Yoshioka, K., Akechi, M., Yamashita, S., Takamatsu, N., Sugiyama, K., Hibi, M., Nakabeppu, Y., Shiba, T., and Yamamoto, K.I. (1999). JSAP1, a novel jun N-terminal protein kinase (JNK)-binding protein that functions as a Scaffold factor in the JNK signaling pathway. *Mol Cell Biol* **19**, 7539-7548.

143. Nguyen, Q., Lee, C.M., Le, A., and Reddy, E.P. (2005). JLP associates with kinesin light chain 1 through a novel leucine zipper-like domain. *J Biol Chem* **280**, 30185-30191.
144. Sonnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J., and Zerial, M. (2000). Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *J Cell Biol* **149**, 901-914.
145. Mohrmann, K., and van der Sluijs, P. (1999). Regulation of membrane transport through the endocytic pathway by rabGTPases. *Mol Membr Biol* **16**, 81-87.
146. Dai, J., Li, J., Bos, E., Porcionatto, M., Premont, R.T., Bourgoin, S., Peters, P.J., and Hsu, V.W. (2004). ACAP1 promotes endocytic recycling by recognizing recycling sorting signals. *Dev Cell* **7**, 771-776.
147. Boucrot, E., and Kirchhausen, T. (2007). Endosomal recycling controls plasma membrane area during mitosis. *Proc Natl Acad Sci U S A* **104**, 7939-7944.
148. Schweitzer, J.K., Burke, E.E., Goodson, H.V., and D'Souza-Schorey, C. (2005). Endocytosis resumes during late mitosis and is required for cytokinesis. *J Biol Chem* **280**, 41628-41635.
149. Fielding, A.B., Schonteich, E., Matheson, J., Wilson, G., Yu, X., Hickson, G.R., Srivastava, S., Baldwin, S.A., Prekeris, R., and Gould, G.W. (2005). Rab11-FIP3 and FIP4 interact with Arf6 and the exocyst to control membrane traffic in cytokinesis. *Embo J* **24**, 3389-3399.
150. Schweitzer, J.K., and D'Souza-Schorey, C. (2002). Localization and activation of the ARF6 GTPase during cleavage furrow ingression and cytokinesis. *J Biol Chem* **277**, 27210-27216.
151. Prigent, M., Dubois, T., Raposo, G., Derrien, V., Tenza, D., Rosse, C., Camonis, J., and Chavrier, P. (2003). ARF6 controls post-endocytic recycling through its downstream exocyst complex effector. *J Cell Biol* **163**, 1111-1121.
152. Zhang, X.M., Ellis, S., Sriratana, A., Mitchell, C.A., and Rowe, T. (2004). Sec15 is an effector for the Rab11 GTPase in mammalian cells. *J Biol Chem* **279**, 43027-43034.
153. Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871-890.
154. Hu, B., Shi, B., Jarzynka, M.J., Yiin, J.J., D'Souza-Schorey, C., and Cheng, S.Y. (2009). ADP-ribosylation factor 6 regulates glioma cell invasion through the IQ-domain GTPase-activating protein 1-Rac1-mediated pathway. *Cancer research* **69**, 794-801.
155. Muralidharan-Chari, V., Hoover, H., Clancy, J., Schweitzer, J., Suckow, M.A., Schroeder, V., Castellino, F.J., Schorey, J.S., and D'Souza-Schorey, C. (2009). ADP-ribosylation factor 6 regulates tumorigenic and invasive properties in vivo. *Cancer research* **69**, 2201-2209.
156. Hiroi, T., Someya, A., Thompson, W., Moss, J., and Vaughan, M. (2006). GEP100/BRAG2: activator of ADP-ribosylation factor 6 for regulation of cell adhesion and actin cytoskeleton via E-cadherin and alpha-catenin. *Proc Natl Acad Sci U S A* **103**, 10672-10677.
157. Jones, M.C., Caswell, P.T., and Norman, J.C. (2006). Endocytic recycling pathways: emerging regulators of cell migration. *Curr Opin Cell Biol* **18**, 549-557.
158. Greco, C., Bralet, M.P., Ailane, N., Dubart-Kupperschmitt, A., Rubinstein, E., Le Naour, F., and Boucheix, C. (2010). E-cadherin/p120-catenin and

- tetraspanin Co-029 cooperate for cell motility control in human colon carcinoma. *Cancer research* 70, 7674-7683.
159. Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* 2, 76-83.
 160. Hajra, K.M., Chen, D.Y., and Fearon, E.R. (2002). The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer research* 62, 1613-1618.
 161. ten Berge, D., Koole, W., Fuerer, C., Fish, M., Eroglu, E., and Nusse, R. (2008). Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell* 3, 508-518.
 162. Figueiredo, J., Simoes-Correia, J., Soderberg, O., Suriano, G., and Seruca, R. (2011). ADP-ribosylation factor 6 mediates E-cadherin recovery by chemical chaperones. *PLoS One* 6, e23188.
 163. Shin, O.H., and Exton, J.H. (2001). Differential binding of arfaptin 2/POR1 to ADP-ribosylation factors and Rac1. *Biochem Biophys Res Commun* 285, 1267-1273.
 164. D'Souza-Schorey, C. (2005). Disassembling adherens junctions: breaking up is hard to do. *Trends Cell Biol* 15, 19-26.
 165. Daniel, J.M., and Reynolds, A.B. (1997). Tyrosine phosphorylation and cadherin/catenin function. *Bioessays* 19, 883-891.
 166. Fujita, Y., Krause, G., Scheffner, M., Zechner, D., Leddy, H.E., Behrens, J., Sommer, T., and Birchmeier, W. (2002). Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nat Cell Biol* 4, 222-231.
 167. Pece, S., and Gutkind, J.S. (2002). E-cadherin and Hakai: signalling, remodeling or destruction? *Nat Cell Biol* 4, E72-74.
 168. Palacios, F., Tushir, J.S., Fujita, Y., and D'Souza-Schorey, C. (2005). Lysosomal targeting of E-cadherin: a unique mechanism for the down-regulation of cell-cell adhesion during epithelial to mesenchymal transitions. *Mol Cell Biol* 25, 389-402.
 169. Steeg, P.S., Bevilacqua, G., Kopper, L., Thorgeirsson, U.P., Talmadge, J.E., Liotta, L.A., and Sobel, M.E. (1988). Evidence for a novel gene associated with low tumor metastatic potential. *J Natl Cancer Inst* 80, 200-204.
 170. Boissan, M., Wendum, D., Arnaud-Dabernat, S., Munier, A., Debray, M., Lascu, I., Daniel, J.Y., and Lacombe, M.L. (2005). Increased lung metastasis in transgenic NM23-Null/SV40 mice with hepatocellular carcinoma. *J Natl Cancer Inst* 97, 836-845.
 171. Palacios, F., Schweitzer, J.K., Boshans, R.L., and D'Souza-Schorey, C. (2002). ARF6-GTP recruits Nm23-H1 to facilitate dynamin-mediated endocytosis during adherens junctions disassembly. *Nat Cell Biol* 4, 929-936.
 172. Boissan, M., De Wever, O., Lizarraga, F., Wendum, D., Poincloux, R., Chignard, N., Desbois-Mouthon, C., Dufour, S., Nawrocki-Raby, B., Birembaut, P., et al. (2010). Implication of metastasis suppressor NM23-H1 in maintaining adherens junctions and limiting the invasive potential of human cancer cells. *Cancer research* 70, 7710-7722.
 173. Palacios, F., and D'Souza-Schorey, C. (2003). Modulation of Rac1 and ARF6 activation during epithelial cell scattering. *J Biol Chem* 278, 17395-17400.

174. Radhakrishna, H., Al-Awar, O., Khachikian, Z., and Donaldson, J.G. (1999). ARF6 requirement for Rac ruffling suggests a role for membrane trafficking in cortical actin rearrangements. *J Cell Sci* 112 (Pt 6), 855-866.
175. Powelka, A.M., Sun, J., Li, J., Gao, M., Shaw, L.M., Sonnenberg, A., and Hsu, V.W. (2004). Stimulation-dependent recycling of integrin beta1 regulated by ARF6 and Rab11. *Traffic* 5, 20-36.
176. Al-Awar, O., Radhakrishna, H., Powell, N.N., and Donaldson, J.G. (2000). Separation of membrane trafficking and actin remodeling functions of ARF6 with an effector domain mutant. *Mol Cell Biol* 20, 5998-6007.
177. Cuilliere-Dartigues, P., El-Bchiri, J., Krimi, A., Buhard, O., Fontanges, P., Flejou, J.F., Hamelin, R., and Duval, A. (2006). TCF-4 isoforms absent in TCF-4 mutated MSI-H colorectal cancer cells colocalize with nuclear CtBP and repress TCF-4-mediated transcription. *Oncogene* 25, 4441-4448.
178. Conrad, M.E., and Umbreit, J.N. (1993). A concise review: iron absorption--the mucin-mobilferrin-integrin pathway. A competitive pathway for metal absorption. *Am J Hematol* 42, 67-73.
179. Wilson, G.M., Fielding, A.B., Simon, G.C., Yu, X., Andrews, P.D., Hames, R.S., Frey, A.M., Peden, A.A., Gould, G.W., and Prekeris, R. (2005). The FIP3-Rab11 protein complex regulates recycling endosome targeting to the cleavage furrow during late cytokinesis. *Molecular biology of the cell* 16, 849-860.
180. Morin, P.J., Sparks, A.B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B., and Kinzler, K.W. (1997). Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* 275, 1787-1790.
181. Roepstorff, K., Grandal, M.V., Henriksen, L., Knudsen, S.L., Lerdrup, M., Grovdal, L., Willumsen, B.M., and van Deurs, B. (2009). Differential effects of EGFR ligands on endocytic sorting of the receptor. *Traffic* 10, 1115-1127.