



UNIVERSITE DE STRASBOURG

Ecole Doctorale des Sciences de la Vie et de la Santé

THESE

**En vue de l'obtention du diplôme de
Docteur en Sciences de l'Université de Strasbourg**

Discipline : Biologie Cellulaire et Moléculaire

Présentée le 1^{er} Octobre 2010 par

Agnès NEUVILLE

**Carcinogenèse colique : sous-typage moléculaire
des étapes précoces du cancer du côlon
de phénotype non-MIN.**

Jury

Monsieur le Pr Pierre OUDET, examinateur
Madame le Dr Elisabeth LABOUESSE-GEORGES, rapporteur interne
Madame le Pr Marie-Danièle DIEBOLD, rapporteur externe
Monsieur le Dr Alex DUVAL, rapporteur externe
Monsieur le Dr Bernard DENIS, examinateur
Madame le Dr Marie-Pierre GAUB, Directeur de thèse

Thèse préparée dans l'Unité INSERM U682 de Madame le Dr Michèle KEDINGER puis dans l'Equipe d'Accueil E4438 de Monsieur le Pr LUDES, Responsable : Mme le Dr Dominique Guénot.

TABLE DES MATIERES

REMERCIEMENTS.....	4
ABREVIATIONS.....	5
RESUME.....	6
INTRODUCTION	8
1. Polypes et cancer du côlon	8
1.1 Epidémiologie	8
1.1.1 Dans le monde	8
1.1.2 En France.....	8
1.2 Histoire naturelle du cancer	10
1.2.1 Foyer de cryptes aberrantes	10
1.2.2 Adénomes	10
1.2.3 Adénocarcinome in situ	11
1.2.4 Adénocarcinome infiltrant	11
1.2.5 Evolution	14
1.3 Dépistage du cancer du côlon	14
1.4 Anatomie pathologique.....	15
1.4.1 Les polypes non épithéliaux.....	16
1.4.2 Les polypes épithéliaux	16
2. La carcinogenèse colique.....	26
2.1 Généralités	26
2.2 Voies moléculaires impliquées	29
2.2.1 Le gène suppresseur de tumeur <i>TP53</i>	29
2.2.2 Voie Wnt/APC/ β caténine.....	30
2.2.3 Voie TGF- β /SMAD	32
2.2.4 Voie RAS/MAPK.....	33
2.2.5 Voie PI3K/Akt	34
2.3 Génétique et cancer du côlon	35
2.3.1 Instabilité chromosomique	35
2.3.2 Instabilité des microsatellites	39
2.3.3 Formes héréditaires	42
2.3.4 Hyperméthylation	44
2.4 Modèles de carcinogenèse.....	48
2.4.1 Modèle de Vogelstein	48
2.4.2 Voie festonnée	49
2.4.3 MGMT	50
2.4.4 Sous-typage moléculaire du phénotype CIN.....	50
2. Conséquences cliniques du diagnostic d'adénome colique	52
3.1 Facteurs pronostiques de récurrence des adénomes coliques.....	52
3.2 Prise en charge des patients porteurs d'adénome.....	53
3.3.1 Indications chirurgicales.....	53
3.3.2 Périodicité de la surveillance	54
3.3.3 Consultation d'oncogénétique.....	54
3.2 Variabilité inter-observateurs du diagnostic des adénomes coliques.....	55
OBJECTIF DU TRAVAIL.....	56

RESULTATS	59
1. Mise en place de la cohorte de polypes coliques	59
2. Sous-typage moléculaire des polypes coliques non-MIN	61
3. Identification de sous-types moléculaires par association d'altérations	62
3.1. Polypes avec AI	62
3.2. Polypes sans AI	63
Publication 1.....	64
4. Proposition de séquences d'initiation adénomateuse	65
5. Altérations moléculaires selon la localisation.....	66
5.1. Polypes avec AI	66
5.2 Polypes sans AI	67
6. Altérations moléculaires sur l'ensemble des chromosomes	68
6.1 Polypes de type non-CIN non-MIN	68
6.2 Identification d'une micro-délétion	69
Publication 2.....	71
7. Proposition de distribution des altérations en fonction de la localisation.....	72
DISCUSSION	73
PERSPECTIVES	80
CONCLUSION	83
REFERENCES BIBLIOGRAPHIQUES	84
ANNEXES	95
Publication 1.....	95
Publication 2.....	96
Publication 3.....	97

REMERCIEMENTS

Je tiens à remercier le Dr Marie-Pierre Gaub, mon directeur de thèse, à qui je voue une immense admiration. Elle m'a accueilli dans son équipe, m'a initié à la biologie moléculaire et a toujours cru en moi. Depuis 1995, nous avons tissé des liens entre la pathologie et la biologie moléculaire que rien ne pourra effacer. Merci Marie-Pierre.

Au Professeur Pierre Oudet, merci pour toutes nos discussions enrichissantes et stimulantes.

Au Professeur Marie-Danièle Diebold.

Au Docteur Alex Duval.

Merci d'accepter de juger ce travail.

Au Docteur Georges-Labouesse et au Docteur Denis, merci d'avoir accepté de participer à ce jury de thèse.

Elisabeth, notre collaboration fut de courte durée mais un challenge technique !

Bernard, ta persévérance dans l'aboutissement d'un projet est un exemple pour moi.

Merci au Dr Dominique Guénot pour sa grande disponibilité et son soutien.

Merci à Céline Nicolet pour son entrain et son aide.

Merci au Dr Nicolas Meyer, statisticien à l'écoute.

Merci à Madame Michèle Kédinger pour son soutien et à l'unité inserm U682 qui m'a intégré dès mon DEA en 2000.

Un grand merci à toute l'équipe EA 4438 pour son aspect hétéroclite.

Merci au Professeur Bellocq pour m'avoir accueillie dans le Département de Pathologie de Strasbourg et m'avoir donné la possibilité d'initier cette thèse.

Merci au Professeur Coindre et à l'équipe du Département de Pathologie de l'Institut Bergonié pour m'octroyer du temps pour la soutenance de cette thèse.

Je remercie mes amis et collègues du Département de Pathologie et du Centre de Ressources Biologiques qui n'ont cessé de me soutenir dans cette entreprise périlleuse.

Enfin, ce travail n'aurait pas pu avoir lieu sans l'aide, l'amour et la compréhension d'un mari incroyable.

Je dédie cette thèse (la dernière !) à mes enfants Bastien et Julia que j'aime, et à mes parents que j'adore.

ABREVIATIONS

APC	Adenomatous Polyposis Coli
CC	Cancer colique
CDK	Cyclin-Dependent Kinase
CGH	Comparative Genomic Hybridization
CIMP	CpG Island Methylator Phenotype
CIN	Chromosomal Instability
DNMT	DNA Methyltransférase
EGF	Epidermal Growth Factor
EMT	Epithelial-Mesenchymal Transition
FAP	Familial Adenomatous Polyposis
FCA	Foyer de cryptes aberrantes
FISH	Fluorescent In Situ Hybridization
GSK3- β	Glycogen Synthase Kinase 3 Beta
GST	Gène suppresseur de tumeur
HGF	Hepatocyte Growth Factor
HIF-1	Hypoxia-Induced Factor 1
HNPCC	Hereditary Non Polyposis Colorectal Cancer
LOH	Loss Of Heterozygosity
MAP	MUTYH Associated Polyposis
MAPK	Mitogen-Activated Protein (MAP) Kinase
MET	Mesenchymal-Epithelial Transition
MGMT	Méthyle guanine méthyle transférase
MIN	Microsatellite Instability
MMP	Matrix Metalloprotease
MMR	Mismatch Repair
MSI	Microsatellite Instability
PDGF	Platelet-Derived Growth Factor
RER	Replication Error
SNP	Single Nucleotide Polymorphism
TGF- β	Tumor Growth Factor beta
VEGF	Vascular Endothelial Growth Factor

RESUME

La carcinogenèse colique implique des mutations et/ou des pertes d'hétérozygoties de gènes suppresseurs de tumeurs et des mutations d'oncogènes. Deux voies de carcinogenèse sont classiquement décrites i) une voie dans laquelle des mutations de gènes de réparation de l'ADN entraînent une instabilité de microsatellites et un fort indice de mutation dans le génome (phénotype microsatellite instable ou MIN), ii) une voie dans laquelle des mutations de gènes génèrent des mécanismes entraînant une instabilité de chromosomes avec de nombreuses pertes d'hétérozygotie (LOH) qui définit le phénotype CIN. Iii) Une troisième voie a été plus récemment décrite impliquant des altérations épigénétiques (phénotype méthylateur des îlots CpG ou CIMP).

Un modèle de progression tumorale a été établi par Vogelstein et al. à partir de données d'altérations génomiques observées dans le cancer colorectal. Cependant, ce modèle n'a pas pris en compte l'hétérogénéité des cohortes, quant à la localisation des lésions et quant à l'existence de sous-types classés sur la base de la fréquence des altérations et leur combinatoire.

A ce jour, ces modèles n'ont pas permis d'établir des marqueurs pronostiques validés dans la pratique clinique. L'existence de sous-types moléculaires au sein du phénotype CIN permettrait peut-être d'expliquer ce manque de marqueurs moléculaires pronostiques.

Afin d'étudier l'existence de ces sous-types, nous avons évalué, par allélotypage ciblant 33 microsatellites localisés sur 18 des chromosomes les plus fréquemment altérés dans les cancers coliques, l'incidence d'altérations génomiques proposées par Vogelstein, dans une cohorte de 159 polypes coliques de phénotypes non-MIN et localisés dans le côlon proximal et distal, en excluant le rectum. De plus, le statut de *KRAS* et la méthylation du promoteur du gène *MGMT* ont permis de compléter la caractérisation des anomalies génétiques et épigénétiques. Afin de préciser la nature des altérations observées par allélotypage (gain ou perte) et d'étendre l'analyse à l'ensemble des chromosomes, une analyse par hybridation génomique comparative (CGH array) a été réalisée sur 110 adénomes de la cohorte.

Comme pour les carcinomes, la clusterisation des données d'allélotypage des polypes a identifié 4 sous-types selon la fréquence globale d'altération : non-CIN non-MIN sans aucune altération des sites analysés (38%), CIN-Very Low avec une fréquence inférieure à 25% (49,5%), CIN-Low moyennement altérés avec une fréquence moyenne de 30% (5,5%) et CIN-High fortement altéré avec une fréquence moyenne de 66% (7%) qui regroupe en majorité, des adénomes tubulo-villeux et de haut grade.

32% des adénomes présentent une mutation de *KRAS* et 39% présentent une méthylation du promoteur du gène *MGMT*. Il est intéressant de noter que 14% des adénomes ne présentent aucune des modifications étudiées.

Parmi les microsatellites altérés, ceux ciblant les chromosomes 1q, 6q et 9q sont plus fréquemment associés à des adénomes de haut grade et de type tubulo-villeux, suggérant que les loci associés sont des facteurs pronostiques de progression.

L'analyse statistique préliminaire des puces CGH a permis d'identifier un gain du chromosome 20q et une délétion de la région génomique localisée en 1p36.11-12. Parmi les gènes localisés en 1p36.11-12, le facteur de transcription *E2F2* est un gène intéressant de par son rôle de régulateur du cycle cellulaire. Ainsi ces deux altérations semblent des évènements précoces dans l'initiation tissu sain-adénome et adénome-carcinome.

Notre travail a montré une grande hétérogénéité des combinaisons d'altérations moléculaires étudiées, en faveur de différents mécanismes d'initiation tumorale des lésions coliques pré néoplasiques de phénotype non MIN. Des altérations spécifiques sont corrélées au type histologique tubulo-villeux et au caractère haut grade, ce qui pourraient représenter des cibles pour identifier des facteurs d'évolutivité.

INTRODUCTION

1. Polypes et cancer du côlon

1.1 Epidémiologie

1.1.1 Dans le monde

L'incidence du cancer du côlon (CC) dans le monde est extrêmement variable selon les pays avec un rapport de 1 à 25 entre l'Inde, pays le moins touché, et l'Amérique du Nord, pays le plus touché. Les pays riches et industrialisés sont les plus concernés par ce cancer avec une forte fréquence dans le Nord et l'Ouest de l'Europe (Cancer Mondial, <http://www-dep.iarc.fr/>). Dans la dernière étude Eurocare-4 du registre européen des cancers portant sur la période 2000-2002, le taux de survie relative à 5 ans est passé à 59.9% en France, taux parmi les plus élevés d'Europe, aux côtés de la Belgique (60.7%), l'Allemagne (61.2%) et la Suisse (63.8%) (Verdecchia, Francisci et al. 2007).

1.1.2 En France

En France, le CC représente environ 15% de l'ensemble des cancers. Sa fréquence le place en 3^{ème} position pour les deux sexes confondus, après les cancers du sein et de la prostate. Chez la femme, c'est la seconde localisation après le cancer du sein, avec une incidence de 24,6 nouveaux cas par an pour 100 000 habitants. Chez l'homme, c'est la troisième localisation après le cancer de la prostate et du poumon, avec une incidence annuelle de 39,1 nouveaux cas pour 100 000 habitants. Depuis 20 ans, son incidence est en constante augmentation, due en partie au vieillissement de la population. Les dernières données épidémiologiques publiées et concernant des patients diagnostiqués de 1980 à 2005, estimaient à 37 413 le nombre de nouveaux cas en 2005 (Belot, Grosclaude et al. 2008). L'augmentation moyenne de l'incidence calculée sur la période 1978-2000 est d'environ 1% par an ce qui permet d'établir une projection de 45.000 nouveaux cas en 2020. En France, l'Alsace est l'une des régions les plus touchées par ce cancer avec plus de 1000 cancers diagnostiqués et 480 morts par an (www.adeca-alsace.fr).

L'incidence du CC augmente significativement après l'âge de 50 ans ; elle est comparable pour les deux sexes jusqu'à l'âge de 60 ans puis augmente plus fortement chez l'homme passé cet âge. L'âge moyen au diagnostic est de 70 ans.

Néanmoins, si la survie globale s'est allongée de manière significative au cours de la dernière décennie, le CC représente toujours un problème majeur de santé publique par sa fréquence et sa gravité, ainsi que son coût. En effet, près de 16 000 personnes décèdent chaque année de ce cancer en France, la survie étant directement corrélée au stade de la maladie au moment du diagnostic.

Le CC n'occasionne des symptômes que tardivement, lorsqu'il est déjà évolué. Sur 100 patients ayant un CC, 15 à 20 ont des métastases hépatiques synchrones et 20 développeront des métastases métachrones dans les 5 ans.

En recoupant les données épidémiologiques des différents réseaux français, on peut estimer que 25 à 30% des CC diagnostiqués chaque année le sont au stade IV (métastatique), soit environ 12.000 CC, soit 50% de l'ensemble des cas de CC diagnostiqués dans notre pays. Un chiffre cohérent avec le nombre de décès estimé puisque les patients atteints de CC décèdent de leurs métastases.

Une étude réalisée en 2000 par le Registre Européen des Cancers montre que la survie à 5 ans est de 77% lorsque le cancer est diagnostiqué à un stade précoce mais seulement de 35 % lorsque le cancer est diagnostiqué à un stade plus avancé (Gatta, Capocaccia et al. 2000).

Enfin, l'augmentation de l'incidence et la baisse relative de la mortalité sont des indicateurs de l'amélioration des traitements au cours des dernières années, en particulier pour les stades les plus avancés de la maladie, avec un dépistage plus précoce, et une meilleure prévention nutritionnelle. En effet, il est maintenant bien démontré que l'incidence d'un CC est en partie favorisée par une alimentation trop abondante et trop riche en graisses animales associée à la sédentarité.

1.2 Histoire naturelle du cancer

1.2.1 Foyer de cryptes aberrantes

La première étape du développement du CC est l'apparition d'une hyperprolifération de l'épithélium colique. Au plan morphologique, la première anomalie décelable est le foyer de cryptes aberrantes (FCA) (Pretlow, Barrow et al. 1991). Il s'agit de cryptes au calibre élargi entourées d'un épithélium épaissi avec une mucosécrétion diminuée. Ces FCA existent sous la forme d'une hyperplasie, c'est à dire une augmentation de la prolifération anormale des cellules normales dans un tissu, ou d'une dysplasie, caractérisée par des anomalies cellulaires, nucléaires et cytoplasmiques associées à des anomalies architecturales des glandes. Dans le côlon, la dysplasie est caractérisée par une stratification des noyaux, des mitoses dans les deux tiers supérieurs des cryptes et une raréfaction des cellules caliciformes et représentent à peu près 5% des FCA et est l'équivalent de micro-adénome. Les FCA ne sont pas visibles en routine en coloscopie. Leur détection nécessite l'utilisation de vidéo-endoscopes à optique grossissante (zoom) et de colorants (chromoendoscopie au bleu de méthylène). La prolifération cellulaire anormale tend à l'accumulation des cellules dans la muqueuse et aboutit à la formation d'un polype.

1.2.2 Adénomes

Les adénomes sont des lésions néoplasiques qui peuvent évoluer vers un carcinome. Ils peuvent être classés selon deux critères : l'histologie et le degré de dysplasie (voir § 1.4).

La prévalence des adénomes augmente avec l'âge et atteint 30% chez des sujets de 65 ans.

La plupart des CC sont des adénocarcinomes et se développent à partir d'un adénome et on estime que 5% des adénomes progressent vers un cancer (Stryker, Wolff et al. 1987). La durée de la séquence « adénome - carcinome » est en moyenne de 10 à 15 ans (**Figure 1**).

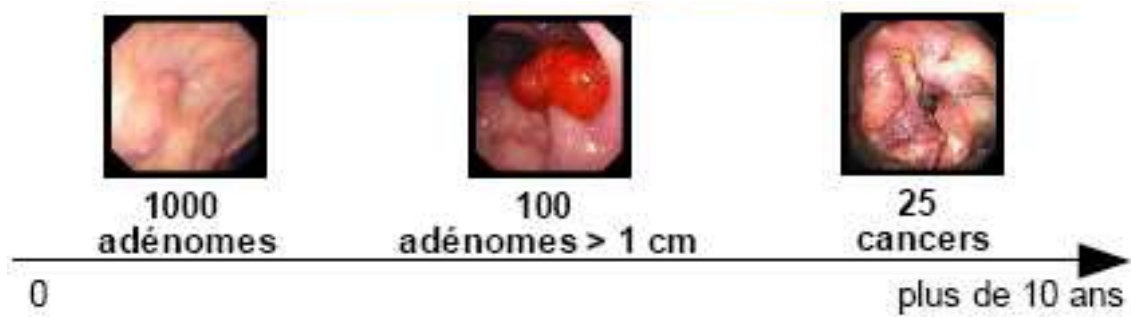


Figure 1 : Evolution du cancer colorectal dans le temps.

1.2.3 Adénocarcinome in situ

Le développement d'un carcinome est d'abord superficiel, strictement intra-épithélial puis intramuqueux, sans franchissement de la lame basale et est désigné sous le terme « *in situ* », stade précoce de cancer.

1.2.4 Adénocarcinome infiltrant

Le terme de carcinome invasif s'applique aux lésions entraînant une rupture de la lame basale et dépassant la musculaire muqueuse. Le carcinome s'étend ensuite progressivement en profondeur dans la paroi colique, dans la sous-muqueuse, dans la musculature, puis dans la sous-séreuse et la séreuse (**Figure 2**).

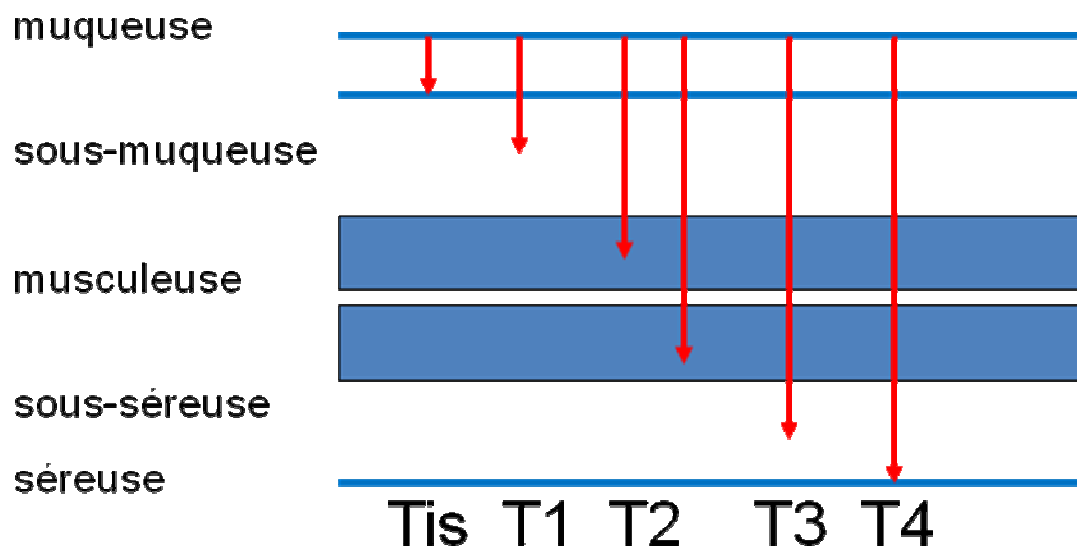


Figure 2 : Stades T de la classification TNM 2009. Extension de l'adénocarcinome à travers les tuniques pariétales du côlon.

Conformément aux recommandations de la conférence de consensus de 1998 sur le CC, un des facteurs pronostiques les plus importants concerne l'extension du carcinome en profondeur qui doit être évaluée selon la classification TNM. Cette classification évalue l'extension d'un cancer à partir de trois éléments : T pour la profondeur d'envahissement de la tumeur primitive (**Figure 2**), N pour la présence/absence d'atteintes ganglionnaires et M pour la présence/absence de métastases (**Tableau 1**).

Stade pTNM	Signification	Stade clinique
Tis	Carcinome in situ : limité à la muqueuse et à la musculaire muqueuse	0
T1	Envahissement de la sous-muqueuse	I
T2	Envahissement de la musculature	I
T3	Envahissement de la sous-séreuse	IIa
T4a	Envahissement de la séreuse	IIb
T4b	Envahissement par contiguïté d'un organe de voisinage	IIb
N1a	1 ganglion métastasé	III
N1b	2 à 3 ganglions métastasés	III
N1c	1 nodule métastatique sans structure ganglionnaire résiduelle	III
N2a	4 à 6 ganglions métastasés	III
N2b	≥ 7 ganglions métastasés	III
M1a	Métastase viscérale d'1 organe	IV
M1b	Métastase viscérale de plus d'1 organe ou métastase péritonéale	IV

Tableau 1 : Classification TNM 2009 du cancer colo-rectal et correspondances cliniques.

Les adénocarcinomes conventionnels, de type lieberkühien ou intestinal, se caractérisent également par leur degré de différenciation : le grade 1 ou bien différencié, pour des carcinomes constitués à plus de 95% de glandes coliques, le grade 2 ou modérément différencié pour des carcinomes constitués de 50 à 95% de glandes, le grade 3 ou peu différencié pour des carcinomes constitués de 5 à 50% de glandes et enfin le grade 4 ou indifférencié pour des carcinomes constitués de moins de 5% de glandes coliques. Ces grades sont indicateurs de pronostic. Ueno et al. rapportent un caractère plus agressif pour les lésions malignes peu différenciées (Ueno, Mochizuki et al. 2004). En pratique courante, la distinction des grades 1 et 2 étant très subjective, une classification simplifiée a été proposée et comprend

2 grades : les adénocarcinomes différenciés, constitués de plus de 50% de glandes et les adénocarcinomes peu différenciés, constitués de moins de 50% de glandes.

1.2.5 Evolution

A un stade tardif de l'évolution du CC, certains clones acquièrent la capacité de disséminer et de former des métastases. Dans 75% des cas, les tumeurs primaires coliques métastasent au niveau du foie. Les autres localisations sont les métastases pulmonaires, osseuses ou cérébrales.

1.3 Dépistage du cancer du côlon

Le dépistage du cancer du côlon est possible et utile et s'adresse à deux catégories de personnes : les personnes ayant un risque moyen de développer un cancer colique et les personnes ayant un risque élevé ou très élevé de développer un cancer colique (**Tableau 2**).

Risque de développer un cancer colique	Type de population
moyen	Personnes de plus de 50 ans
élevé	<ul style="list-style-type: none"> - Sujets ayant un parent au 1er degré de moins de 60 ans ou plusieurs parents au 1er degré, atteints d'un cancer colorectal ou d'un adénome avancé - Patients atteints de maladie inflammatoire chronique de l'intestin, rectocolite ulcérohémorragique et maladie de Crohn, en cas de pancolite d'évolution prolongée - Patients atteints d'acromégalie
très élevé	Sujets appartenant à une famille atteinte de cancers à transmission héréditaire

Tableau 2 : Définitions des risques de développer un cancer colique.

L'objectif du dépistage du cancer du côlon chez la population à risque moyen est de diminuer de 15 à 20 % la mortalité liée à cette maladie. Ce dépistage est possible par la recherche de saignement occulte dans les selles. Le dépistage organisé a démontré son efficacité pour réduire la mortalité par cancer (Gross, Andersen et al. 2006; Levin, Lieberman et al. 2008).

Un dépistage gratuit du cancer colorectal est proposé aux Français âgés de 50 à 74 ans.

Mis en place depuis 2002 dans des départements pilotes, dont celui du Haut-Rhin en 2003, le programme de dépistage organisé du cancer colorectal a été généralisé à l'ensemble du territoire en 2009.

Le dépistage du cancer colorectal se fait par la recherche de sang occulte dans les selles avec le teste HémocultII®.

Un test négatif signifie qu'aucun saignement n'a été détecté. Il faut ensuite renouveler le test deux ans plus tard.

Un test positif signifie que du sang a été détecté dans les selles. Une coloscopie est alors prescrite.

Le taux de participation moyen est de 42% et la coloscopie est réalisée dans 87% des cas de tests positifs. Le taux de détection pour 1000 personnes détectées, évalué sur les 22 départements pilotes de la première campagne de dépistage, est de 7,5% d'adénomes, 3,7% d'adénomes avancés et 2,3% de carcinome dont plus de 60% était à un stade clinique précoce I/II (INVS, évaluation épidémiologique du dépistage organisé du cancer du côlon et du rectum en France, avril 2009).

1.4 Anatomie pathologique

Les polypes du tube digestif et plus particulièrement du côlon représentent un ensemble d'entités diverses.

Le terme polype a une signification macroscopique. Il correspond à une excroissance tissulaire faisant protrusion dans la lumière intestinale. Il faut donc lui adjoindre un terme histologique pour déterminer la nature du polype : épithélial ou non épithélial.

1.4.1 Les polypes non épithéliaux

Ils peuvent être de type inflammatoire par prolapsus de la muqueuse, souvent de localisation sigmoïdienne ou rectale, ou correspondre à une tumeur mésenchymateuse qui soulève la muqueuse, comme par exemple un lipome, un léiomyome, une tumeur stromale intestinale (**Figure 3**).

1.4.2 Les polypes épithéliaux

Ils sont classiquement divisés en polypes adénomateux et non adénomateux.

1.4.2.1 Les polypes non adénomateux (**Figure 3**)

- Le plus fréquent des polypes non adénomateux est le polype hyperplasique : prolifération épithéliale bénigne de petite taille (moins de 1 cm), il siège le plus souvent au niveau du côlon sigmoïde et du rectum. Les glandes sont allongées, bordées de cellules mucipares et non mucipares, avec une architecture dentelée dans la moitié supérieure des glandes.

Certains polypes hyperplasiques sont de plus grande taille (plus de 1 cm) siégeant au niveau du côlon droit.

- Le polype juvénile siège principalement au niveau du rectum, chez le sujet jeune. Ce polype hamartomateux est constitué de glandes tortueuses au revêtement hyperplasique associées à un chorion inflammatoire. Le plus souvent sporadique, il n'a alors pas de potentiel de transformation cancéreuse. Ce potentiel est par contre élevé lorsqu'il est associé à une polypose juvénile.

- Le polype juvénile-like caractérise le syndrome de Cronkhite Canada. Ce polype hamartomateux de type juvénile siège au niveau colique, intestinale grêle et gastrique, sans potentiel de dégénérescence maligne et sans prédisposition familiale.

- Le polype de Peutz-Jeghers est un polype hamartomateux et siège le plus souvent au niveau de l'intestin grêle et de l'estomac. Ce polype sessile ou pédiculé a une architecture caractéristique en arbre (aussi dénommé « Christmas tree »), avec un revêtement mucipare de type hyperplasique et un chorion parfois occupé par des fibres musculaires lisses

perpendiculaires à la lumière intestinale. Le revêtement est parfois mixte hyperplasique et adénomateux avec un risque de transformation cancéreuse.

- Le polype hamartomateux du syndrome de Cowden est constitué de glandes désorganisées au sein d'un chorion fibreux parfois parcouru de fibres musculaires lisses.

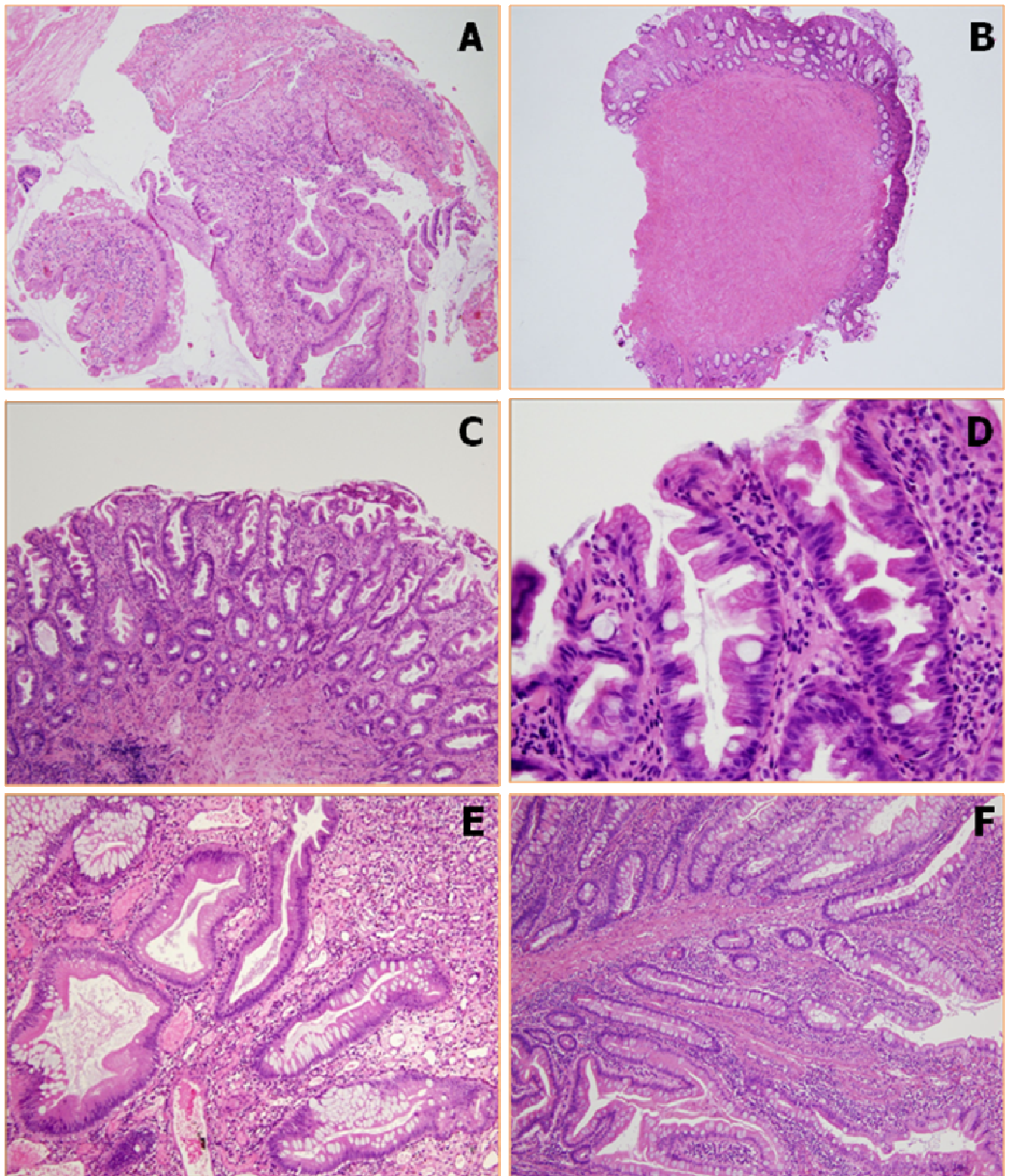


Figure 3 : Histologie de polypes non-épithéliaux et non-adénomateux (H.E. x10 pour A-C-E-F, x20 pour D, x4 pour B)

A : Polype inflammatoire : muqueuse ulcérée, chorion inflammatoire

B : Léiomyome sous-muqueux

C : Polype hyperplasique

D : Polype hyperplasique : revêtement festonné sans atypies

E : Polype juvénile : glandes ectasiées à revêtement mucipare, chorion inflammatoire

F : Polype de Peutz-Jeghers

1.4.2.2 Les polypes adénomateux

Les adénomes coliques sont des proliférations intra-épithéliales de cellules néoplasiques, ce qui signifie que tout adénome comporte de la dysplasie. La taille des adénomes varie d'un millimètre à plusieurs centimètres.

Ces adénomes se présentent dans plus de 90% des cas sous forme de polypes (Park, Kim et al. 2008). On parle de polype pédiculé lorsque la taille de la base d'implantation du polype est inférieure à celle du polype et de polype sessile lorsque la taille de la base d'implantation du polype est supérieure à celle du polype.

Lorsqu'ils ne sont pas polypeux, on parle d'adénomes plans. Cette distinction macroscopique est importante pour l'évaluation des critères pronostiques des adénomes de haut grade (voir ci-dessous).

Lors de l'établissement d'un diagnostic d'adénome, il est nécessaire de préciser le type histologique et le degré de dysplasie.

Au niveau histologique, l'Organisation Mondiale de la Santé individualise trois types de lésions adénomateuses avec un risque croissant de transformation tumorale : tubuleux, tubulo-villeux et vilieux (**Figure 4**). L'adénome tubuleux est le plus fréquent, observé dans 65 à 87% des cas, l'adénome vilieux dans 8 à 25% des cas et l'adénome tubulo-villeux, associant les deux formes précédentes, dans 5 à 10% des cas.

A ces trois aspects classiques d'adénomes, s'ajoute un quatrième type histologique : l'adénome festonné (serrated adenoma). Décrit en 1990 par Longacre et Fenoglio (Longacre et Fenoglio-Preiser, 1990), ce type d'adénome n'a vraiment été reconnu qu'à partir de 2002 (Sawyer, Cerar et al. 2002). Cette entité, découverte à l'occasion d'études génomiques, est caractérisée par des voies de carcinogenèse différentes et comprend trois sous-types histologiques : le polype hyperplasique, l'adénome festonné sessile et l'adénome festonné traditionnel (**Figure 4**). Ces trois lésions ont en commun l'architecture festonnée du revêtement des glandes mais diffère par les contours des glandes et par leur revêtement plus ou moins adénomateux et prolifératif (**Tableau 3**).

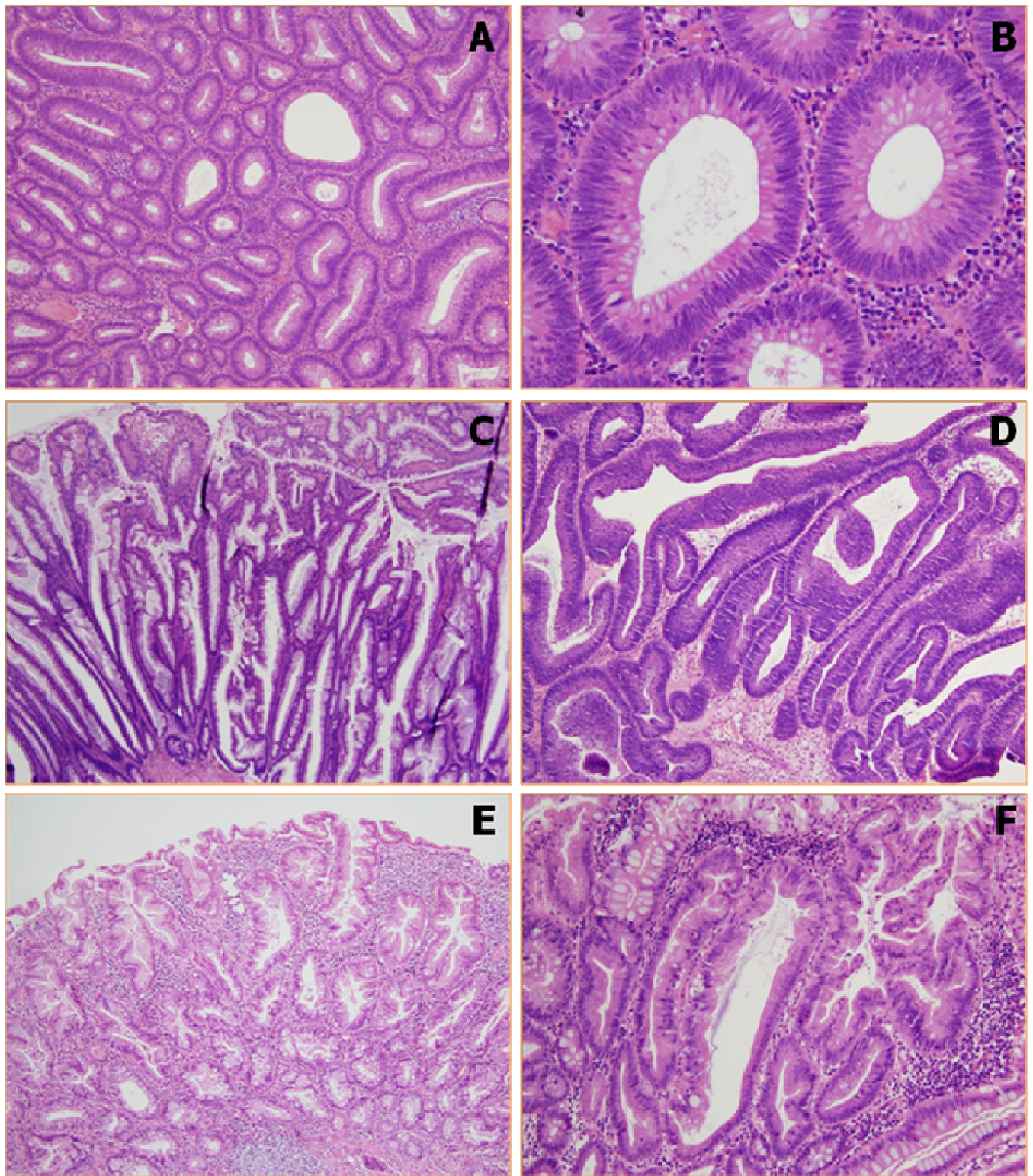


Figure 4 : Histologie des adénomes (H.E. x10 pour A-C-D-E-F, x20 pour B)
 A : Adénome tubuleux : plus de 80% des glandes ont un aspect tubuleux
 B : Adénome avec dysplasie de bas grade
 C : Adénome vilieux : plus de 80% des glandes ont un aspect vilieux
 D : Adénome tubulo-vilieux : plus de 20% des glandes sont de l'un et de l'autre type
 E : Adénome festonné sessile
 F : Adénome festonné traditionnel

Histologie	Caractéristiques
Polype hyperplasique Type à cellules Goblet Type microvésiculaire	Glandes allongées rectilignes au niveau du fond des cryptes Aspect festonné au niveau du haut des cryptes Prédominance de cellules Goblet Prédominance de cellules cylindriques microvésiculaires
Adénome festonné sessile	Glandes d'aspect festonné, dilatées et branchées jusqu'au fond des cryptes. Prolifération \pm anormale du revêtement.
Adénome festonné traditionnel	Glandes d'aspect festonné, dilatées et branchées jusqu'au fond des cryptes. Revêtement stratifié atypique.
Polype mixte	Aspect de polype hyperplasique et d'adénome.

Tableau 3 : Différentes formes de polypes festonnés (selon Snover, Jass et al. 2005).

Le deuxième critère définissant un adénome est la dysplasie. Elle comprend des anomalies cellulaires (augmentation du rapport nucléo-cytoplasmique, hyperchromatisme, apparition d'un nucléole) et architecturales (pluri-stratification, perte de polarité, fusion des glandes). Il s'agit d'une lésion limitée à la muqueuse et à la musculaire muqueuse, sans infiltration de la sous-muqueuse. Les anomalies sont classées en degrés de dysplasie : bas ou haut.

Le risque d'évoluer vers un cancer est d'autant plus grand que la dysplasie est de haut grade. Ainsi selon le degré de dysplasie on distinguera des adénomes pré-cancéreux dits de bas grade et des adénomes cancéreux dits de haut grade équivalents à de l'adénocarcinome in situ. Ce dernier est morphologiquement divisé en deux sous-types : l'adénocarcinome intra-épithélial et l'adénocarcinome intra-muqueux (**Figure 5**).

La distinction histologique de ces deux lésions est difficile et malgré la différence fondamentale de dépassement ou non de la membrane basale, elle n'entraîne aucune implication clinique puisque le risque métastatique de l'adénocarcinome intra-épithélial et de l'adénocarcinome intra-muqueux sont identiques. Ces deux états sont donc tous les deux classés en pTis dans la classification pTNM (**Tableau 1**).

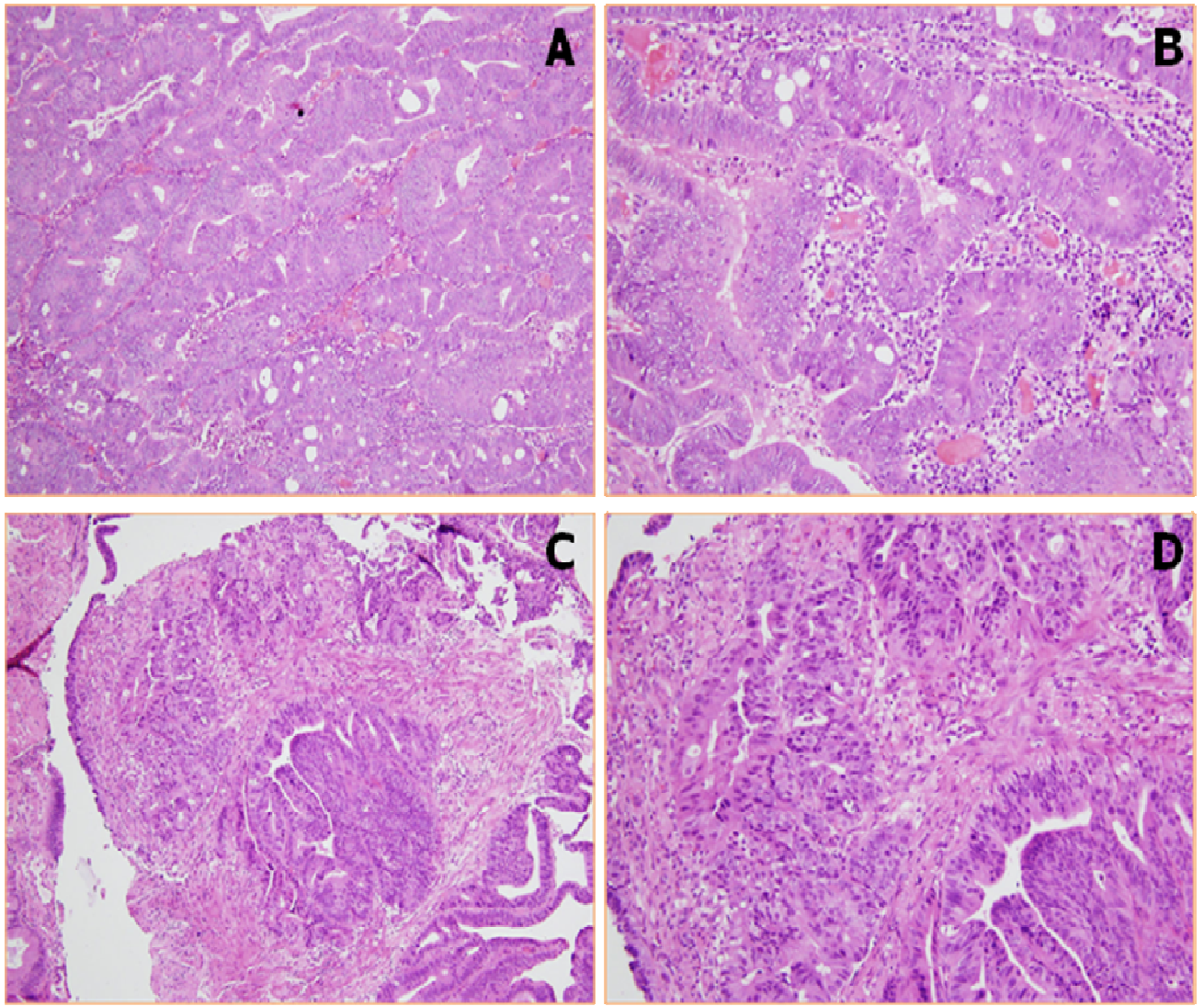


Figure 5 : Histologie des adénomes de haut grade (H.E. x10 pour A-C, x20 pour B-D)

A : Adénome de haut grade : adénocarcinome intra-glandulaire

B : Dysplasie de haut grade : modifications architecturales et cytologiques

C : Adénome de haut grade : adénocarcinome intra-muqueux

D : Dysplasie de haut grade : modifications architecturales et cytologiques, avec stroma-réaction

Depuis 2000, la classification de Vienne modifiée (Dixon, 2002) a permis d'uniformiser la dénomination des lésions intestinales superficielles et de comparer les études internationales épidémiologiques et cliniques (**Tableau 4**).

Catégorie	Néoplasie	Correspondance pT
C1	absente	pT0
C2	indéfinie	pTX
C3	bas grade	pT0
C4	haut grade	pTis
C4.1	<i>adénome haut grade</i>	
C4.2	<i>suspicion de carcinome non invasif</i>	
C4.3	<i>carcinome intraglandulaire</i>	
C4.4	<i>suspicion de carcinome invasif</i>	
C4.5	<i>carcinome intra-muqueux</i>	
C5	carcinome invasif	≥ pT1

Tableau 4 : Classification de Vienne modifiée (Dixon, 2002).

Les facteurs pronostiques des adénomes de haut grade sont :

- La différenciation de la composante infiltrante du carcinome, comme celle déterminée au niveau de l'adénocarcinome commun.
- La présence d'embolies vasculaires
- Le niveau d'infiltration du carcinome invasif

Le niveau d'infiltration sera déterminé différemment selon que l'adénome est un polype pédiculé ou un polype sessile/ou un adénome plan.

- a. Lorsque l'adénome est un polype pédiculé, le niveau d'infiltration est défini par les niveaux de Haggitt (Haggitt, Glotzbach et al. 1985), selon que la lésion n'a pas dépassé la membrane basale, a dépassé la membrane basale au dessus du collet, au niveau du collet ou au-dessous du collet (**Figure 6**).

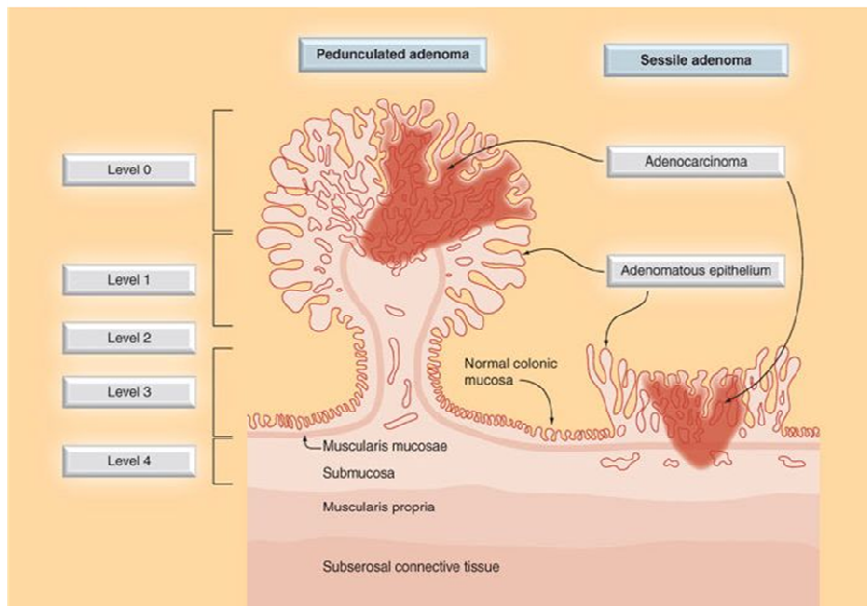


Figure 6 : Représentation des niveaux de Haggitt (Soreide, Nedrebo et al. 2009).

Niveau 0 : carcinome intra-muqueux.

Niveau 1 : carcinome envahissant la sous-muqueuse au niveau de la tête du polype.

Niveau 2 : carcinome envahissant la sous-muqueuse au niveau du collet du polype.

Niveau 3 : carcinome envahissant la sous-muqueuse au niveau du pied du polype.

Niveau 4 : carcinome envahissant la sous-muqueuse de la paroi colique, mais au dessus de la musculéuse.

- b. Lorsque l'adénome est un polype sessile ou un adénome plan, la classification japonaise du niveau d'invasion de la sous-muqueuse (Kikuchi, Takano et al. 1995) s'applique (**figure 7**).

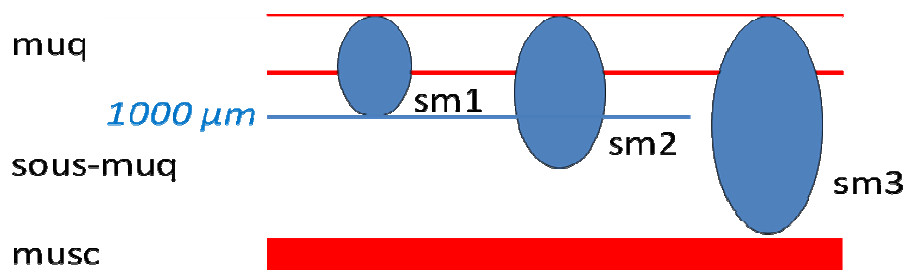


Figure 7 : Représentation des niveaux d'infiltration pour les lésions planes superficielles du côlon.

SM1 : carcinome envahissant la sous-muqueuse sur moins de 1000 μm .

SM2 : carcinome envahissant la sous-muqueuse sur plus de 1000 μm , mais sans atteindre le tiers inférieur.

SM3 : carcinome envahissant la sous-muqueuse sur plus des deux tiers de l'épaisseur.

- L'existence d'un budding

Le phénomène de budding a été récemment introduit, mais n'est pas encore intégré dans la routine et son caractère pronostique pas encore prouvé (Sy, Fung et al. 2010). Il correspond à une dédifférenciation de la composante invasive du carcinome sur le front d'invasion, sous la forme d'au moins 5 amas de moins de 5 cellules se détachant des boyaux ou glandes adénocarcinomeuses mêlés à la stroma-réaction (**Figure 8**), à un grossissement de x20 pour un champ de 0,785 mm² (Ueno, Price et al. 2004). Ce processus serait lié à un plus grand risque d'évolution et constituerait un critère majeur de reprise chirurgicale en cas d'exérèse par voie endoscopique d'un adénome dégénéré (Choi, Sohn et al. 2009).

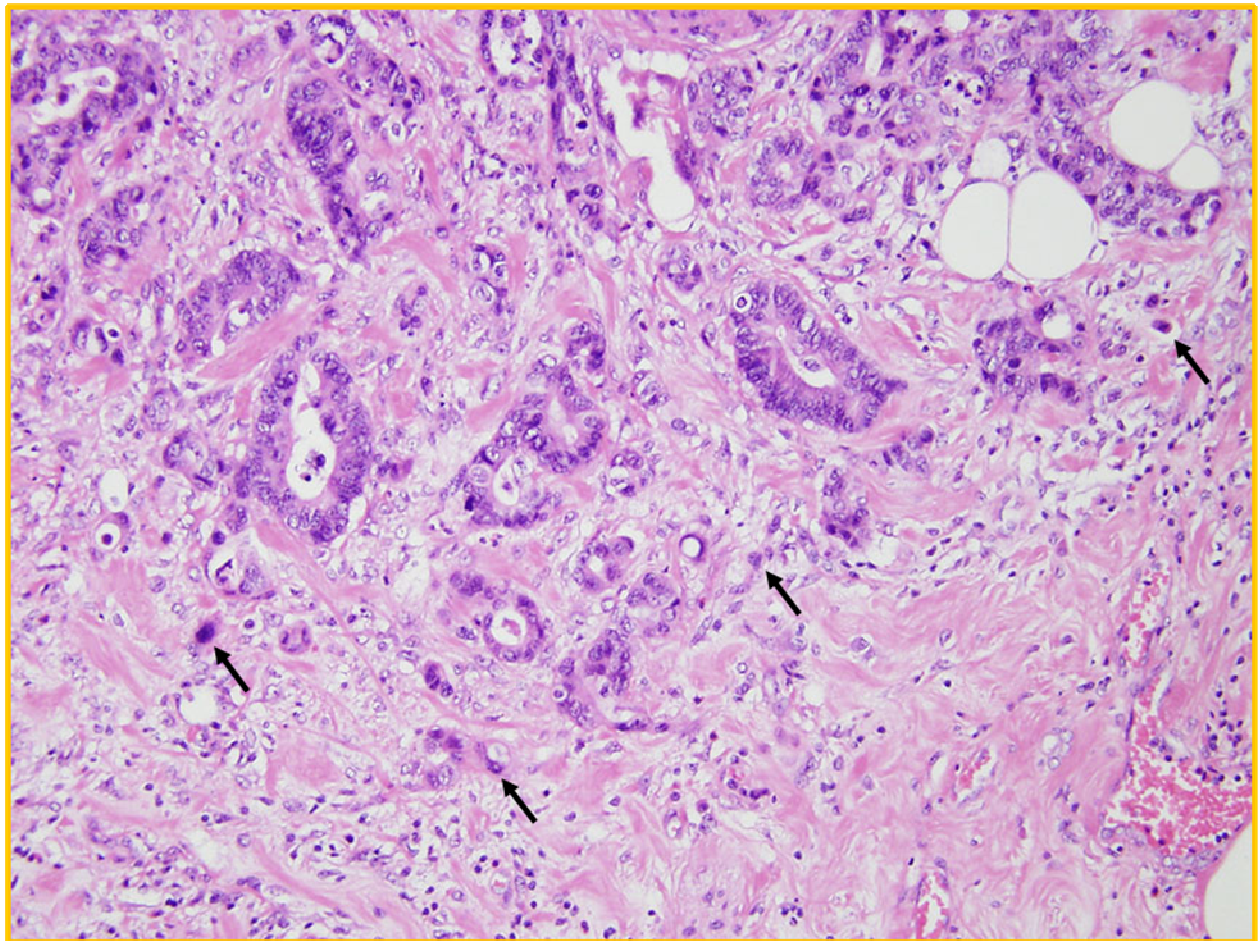


Figure 8 : coupe histologique du front d'invasion d'un adénocarcinome intestinal colique (HEx20). Des cellules isolées se détachent des glandes carcinomateuses (flèches) correspondant au phénomène de budding.

2. La carcinogenèse colique

2.1 Généralités

La carcinogenèse colique est un processus d'initiation et de transformation de cellules normales suite à des modifications génétiques et épigénétiques de l'ADN (Kinzler and Vogelstein, 1996, Smith, Carey, 2002, Jones and Baylin, 2002). L'instabilité génétique est une caractéristique des cellules cancéreuses. Les déterminants de l'instabilité génétique sont en majorité liés à des mécanismes cellulaires fondamentaux, comme le cycle cellulaire, la réplication ou la réparation de l'ADN. Ces mécanismes sont normalement reliés entre eux par de multiples interconnexions, qui forment autant de points de contrôle dont les cellules tumorales s'affranchissent. A ces mécanismes s'ajoutent les modifications épigénétiques qui sont transmises de cellules mères à cellules filles. Celles-ci, sans affecter directement l'intégrité du génome, peuvent en changer le

fonctionnement et la dynamique.

Deux grandes voies de cancérogenèse colique correspondent à deux types d'instabilité génomique :

- l'instabilité des microsatellites (MIN, MSI, encore appelé RER pour replication error) dans 15% des cas pour lesquels le contenu génétique reste diploïde et
- l'instabilité chromosomique (CIN) dans 85% des cas, caractérisée par une aneuploïdie (Lengauer, Kinzler et al. 1998).

Ces deux phénotypes peuvent être distingués par allélotypage, qui correspond à l'analyse par PCR de microsatellites, séquences di-, tri- ou tétra-nucléotidiques répétées et hautement polymorphes réparties dans les séquences codantes et non codantes du génome. La cartographie de ces séquences a été en partie réalisée par le Généthon (1996 Généthon Microsatellite Maps) (Dib, Faure et al. 1996).

Cette instabilité génomique s'accompagne de nombreuses mutations dans trois classes de gènes : les oncogènes, les gènes suppresseurs de tumeur (Gryfe, Swallow, et al. 1997) et les gènes mutateurs.

Les oncogènes sont des gènes codant pour des protéines impliquées dans la prolifération cellulaire. A l'état normal, ils influent positivement sur la croissance et la division cellulaire : ils sont alors appelés proto-oncogènes. Lorsqu'ils sont mutés par gain de fonction, ils prennent la

nomination d'oncogène. Leur activation exagérée, qui transforme par définition un proto-oncogène en oncogène, peut résulter d'une altération structurale directe du gène et/ou d'une expression exagérée de son produit. La mutation d'un seul allèle est suffisante pour entraîner une stimulation exagérée de la prolifération cellulaire. Parmi les oncogènes décrits, on trouve des gènes pro- et anti-apoptotique (*Bcl2*, *Bax*) et des gènes mutagènes codant pour des petites protéines G (*RAS*, *RAF*, *RAB*).

Les gènes suppresseurs de tumeurs (GST), ou anti-oncogènes, sont des gènes présents dans toutes les cellules normales qui agissent sur le contrôle du cycle cellulaire et la différenciation cellulaire, et certains sont impliqués dans l'induction de l'apoptose, comme la protéine très fréquemment impliquée p53. Leur mutation entraîne une perte de fonction. Les deux allèles doivent donc être mutés pour entraîner la perte de fonctionnalité qui aboutit à la transformation des cellules et à une prolifération cellulaire dérégulée.

Le troisième type de gène pouvant affecter la carcinogenèse colique sont des gènes de type mutateur. Ces gènes sont impliqués dans les mécanismes de réparation. Au cours de la synthèse de l'ADN, la plupart des erreurs de réplication sont immédiatement corrigées par l'activité exonucléase de l'ADN polymérase. Les erreurs échappant à ce système intrinsèque sont corrigées par l'activation du système de réparation des mésappariements de l'ADN (MMR pour *mismatch repair*). La mutation va entraîner une perte de fonction. Les mutations germinales de ces gènes sont responsables du syndrome de prédisposition au cancer colorectal non polyposique (HNPCC, pour *hereditary non polyposis colorectal cancer* - Lynch, Smyrk, et al. 1996). À ce jour, cinq gènes impliqués dans le système MMR ont été identifiés chez l'homme : *hMSH2*, *hMLH1*, *hPMS1*, *hPMS2* et *hMSH6* (Modrich and Lahue, 1996). D'autres gènes intervenant dans les systèmes de réparation de l'ADN endommagé, comme le système de réparation par excision de base (BER) ou le système de réparation par excision de nucléotides (NER) appartiennent à ce groupe de gènes mutateurs. Des gènes appartenant au système de correcteurs de modifications épigénétiques (*MGMT*) peuvent aussi être mutateurs.

2.2 Voies moléculaires impliquées

A l'heure actuelle, il a été montré que cinq voies de signalisation, mdm2/p53, Wnt/APC/ β caténine, TGF β /SMAD, RAS/MAPK et PI3K/AKT, sont très fréquemment impliquées dans la carcinogenèse colique, comme dans de nombreux autres cancers.

2.2.1 Le gène suppresseur de tumeur *TP53*

Le GST *TP53* est localisé sur le bras court du chromosome 17 en 17p13.1. Il comporte 11 exons et 10 introns. Il code pour une protéine nucléaire de 53 kDa et se compose de 5 régions hautement conservées et 4 domaines fonctionnels (**Figure 9**).

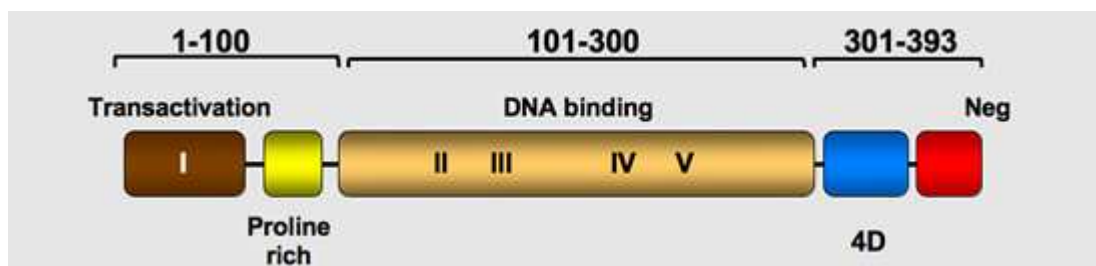


Figure 9 : représentation des domaines de la protéine p53 (d'après le *TP53* web site : <http://p53.free.fr>).

Dans une cellule normale non soumise à un stress, la protéine p53 est rapidement ubiquitinilée puis dégradée par le protéasome. En réponse aux différents stress cellulaires de types intrinsèque (activation de voies oncogéniques) ou extrinsèque (agents génotoxiques), l'augmentation de l'expression de la protéine p53 induit l'expression de très nombreux gènes cibles essentiels au maintien de l'intégrité du génome. Sa mutation bi-allélique va entraîner une dérégulation du cycle cellulaire et du contrôle de l'intégrité de l'ADN.

2.2.1.1 p53 et le cycle cellulaire

Dans une situation de stress, p53 participe à la régulation du cycle cellulaire en induisant l'expression de trois gènes essentiels, *p21* (el-Deiry, Tokino et al. 1993), *14-3-3 σ* (Hermeking, Lengauer et al. 1997) et *GADD45* (Kastan, Canman et al. 1995). L'induction de *p21* en phase G1, permet une inhibition des cyclin-dépendant kinase (CDK), CycA/CDK2, CycE/CDK2, CycD/CDK4 (Harper, Adami et al. 1993), et induit donc un arrêt prolongé du cycle pouvant conduire à la sénescence. Suite à des lésions de l'ADN, les protéines 14-3-3 σ et GADD45 permettent un arrêt du cycle en G2/M via soit une séquestration de CycB1 et CDC2 dans le cytoplasme par la protéine 14-3-3 σ , soit une déstabilisation des complexes CycB1/CDC2 par GADD45 (Chan, Hermeking et al. 1999; Jin, Tong et al. 2002).

2.2.1.2 p53 et l'apoptose

Suite à des dommages affectant l'ADN, à l'activation d'oncogènes, à l'hypoxie ou à des stress oxydatifs, la protéine p53 participe également à la régulation de l'apoptose en induisant l'expression de gènes impliqués dans la voie intrinsèque mitochondriale (*Bid*, *Puma*, *Noxa*, *AFAP-1*) (Michalak, Villunger et al. 2005; Willis and Adams 2005; Wiseman 2006), ou l'expression des récepteurs membranaires de mort *CD95/Fas/Apo-1*, *DR4*, *DR5*, impliqués dans le processus d'apoptose extrinsèque (Sheikh, Burns et al. 1998; Bouvard, Zaitchouk et al. 2000; Liu, Yue et al. 2004). *p53* est également capable d'induire la surexpression du GST *PTEN* afin d'inhiber les signaux de survie cellulaire (Stambolic, MacPherson et al. 2001). La protéine p53 fait partie des protéines les plus fréquemment mutées dans les cancers (30 à 40% tous cancers confondus) et a été un des premiers GST décrits.

2.2.2 Voie Wnt/APC/ β caténine

La voie Wnt est une des voies de signalisation majeure dans le processus d'embryogenèse et de cancérogenèse (Gregorieff and Clevers 2005). Dans le cas du processus normal de développement, l'activation des récepteurs frizzled par la fixation du ligand Wnt conduit à la phosphorylation de la protéine Dishevelled qui, par son association à l'Axine, empêche la GSK3 β de phosphoryler ses substrats dont la β -caténine, qui sera alors transloquée dans le noyau où elle s'associe aux facteurs de transcription TCF/LEF pour activer la transcription de gènes cibles (*c-myc*, *Cycline D1* ou *c-jun*, *MMP7* (ou matrilysine), *PPAR δ* ; **Figure 10**) (Behrens, von Kries et

al. 1996). En l'absence du ligand Wnt, la GSK3 β phosphoryle la β -caténine dans le complexe Axine/APC, qui sera ubiquitinilée et dégradée par le protéasome.

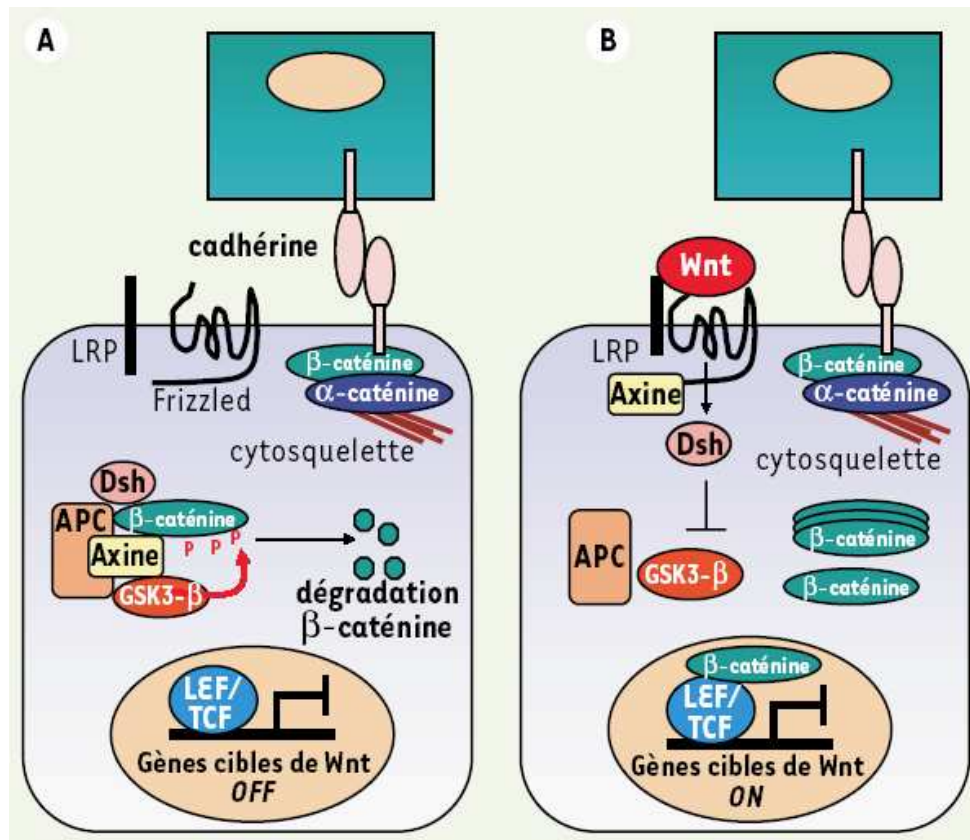


FIGURE 10 : Voie de signalisation Wnt. En l'absence de Wnt (**A**), la β -caténine est incluse dans un complexe protéique comprenant l'axine, APC et la kinase GSK3- β . La phosphorylation de la β -caténine par la GSK3- β entraîne sa dégradation par le protéasome. La β -caténine peut aussi être liée à la cadhérine et contrôler l'adhérence intercellulaire. En présence du ligand Wnt (**B**), β -caténine se dissocie du complexe APC/Axine/GSK, n'est plus phosphorylée et migre dans le noyau où elle se lie aux facteurs de transcription LEF/TCF pour déclencher la transcription de gènes cibles.

Dans les cancers, la voie Wnt peut être activée en absence du ligand par la mutation de gènes codant pour les protéines de cette voie. Le plus souvent, il s'agit de l'inactivation par mutation du gène *APC* induisant la synthèse d'une protéine APC tronquée incapable d'interagir et de neutraliser la β -caténine et/ou l'Axine. Mais dans 15% des tumeurs sporadiques ayant une protéine APC intacte, il peut aussi s'agir de mutations au niveau de séquences répétées du gène de l'Axine, celles activatrices, dans l'exon 3 du gène codant pour la β -caténine (*CTNNB1*) au niveau des sites de phosphorylation (Samowitz, Powers et al. 1999), ou de *TCF4* (Shimizu, Ikeda et al. 2002).

2.2.3 Voie TGF- β /SMAD

La superfamille du Tumor Growth Factor beta (*TGF- β*) comprend différents membres dont les BMPs (Bone Morphogenic Proteins), les *activines* et les *TGF- β* : ce sont des morphogènes dont le rôle est essentiel au cours de l'embryogenèse. Leurs effets dépendent de l'intensité et de la durée des signaux induits via leurs récepteurs (Schmierer and Hill 2007) et une de leurs fonctions essentielles chez l'adulte, en particulier au niveau de l'épithélium digestif, est une inhibition de la prolifération par ralentissement du cycle en G1 (Heldin, Miyazono et al. 1997).

La liaison du TGF- β à son récepteur TGFBRII permet le recrutement et l'activation par phosphorylation du récepteur de type I. En fonction du ligand initial, le récepteur activé phosphoryle des protéines SMAD (SMAD 2 et 3 / SMAD 1, 5 et 8) et les phospho-SMAD s'associent à SMAD4. Les complexes formés sont transloqués au noyau où ils induisent la transcription de gènes cibles (dont des régulateurs majeurs du cycle cellulaire comme les Cyclin Dependent Kinase Inhibitors (CKI) p15/p16 et p21 ; **Figure 11**).

Dans les CC, une mutation inactivatrice du *TGFBRII* est retrouvée dans la plupart des tumeurs de phénotype MIN (90%), et dans 58% des tumeurs, une mutation sur le récepteur de type II de l'activine (ACVR2). Ces mutations sont beaucoup plus rares dans les tumeurs CIN (Grady, Myeroff et al. 1999).

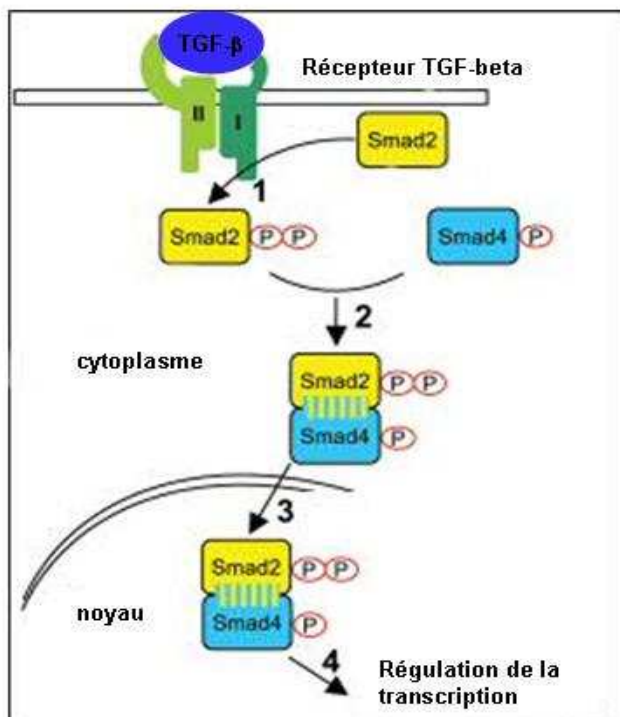


Figure 11 : La voie TGF- β /SMAD. L'activation des récepteurs sérine/thréonine kinases de type I et II du TGF- β phosphoryle les protéines Smad qui vont migrer dans le noyau et forment des complexes avec des facteurs transcriptionnels de type Fast-1 pour moduler la transcription de gènes contrôlant la prolifération et la différenciation.

2.2.4 Voie RAS/MAPK

La cascade Ras/Raf/MEK/Erk (ou Ras/MAPK) transmet des signaux mitogènes via des récepteurs à tyrosine kinase comme EGFR ou VEGFR (**Figure 12**). Cette voie est particulièrement importante pour la croissance et la prolifération des cellules tumorales mais également pour l'induction de signaux pro-angiogéniques (McCubrey, Steelman et al. 2006).

Dans les CC, les mutations de *KRAS* (codons 12, 13 et plus rarement 61) entraînent l'activation constitutive de la voie et sont retrouvées dans 50% des tumeurs de phénotype CIN, alors que les mutations de B-Raf spécifient plutôt les tumeurs de phénotype MIN. L'activation de la voie RAS/MAPK peut également être associée à une surexpression de récepteurs de facteurs de croissance, comme l'*EGFR*, surexprimé dans 30 à 70% des CC (Roberts and Der 2007).

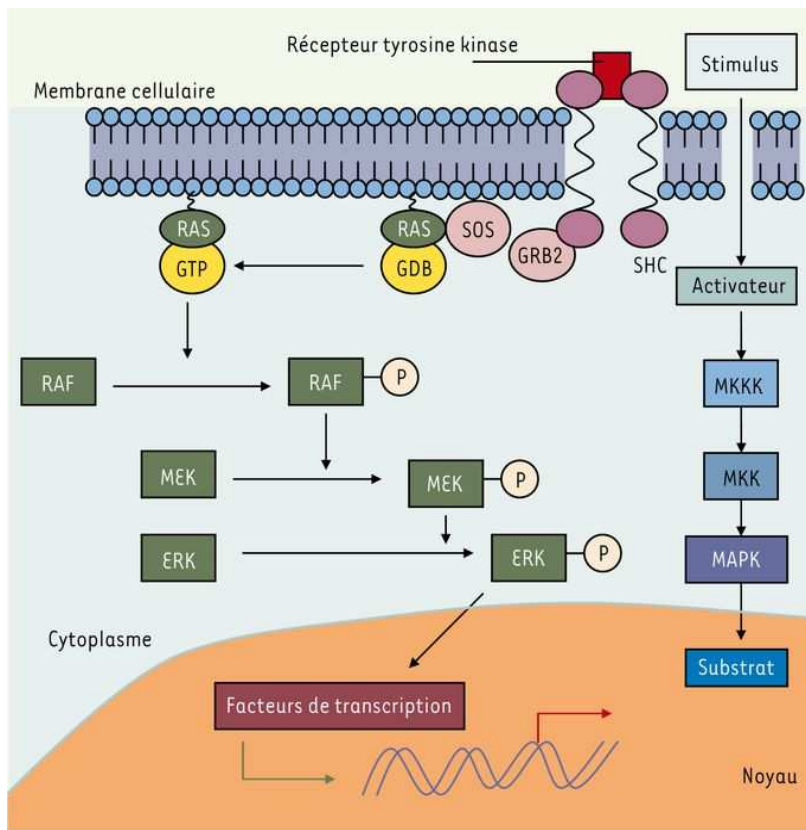


FIGURE 12 : La voie Ras/MapK. L'activation de cette voie conduit à des modifications transcriptionnelles qui favorisent la division cellulaire (Dalle, Martin-Denavit et al. 2006).

Parmi les cibles de la voie RAS/MAPK, il y a entre autres le gène *p14/ARF*, qui va inhiber *Mdm2*, stabiliser la protéine p53 et induire l'expression de *p21*, favorisant ainsi un arrêt prolongé du cycle cellulaire et la sénescence. Comme dans les CC, la protéine p53 est fréquemment inactive, l'activation de la voie induit une prolifération accélérée (Lowe, Cepero et al. 2004). La cascade Raf/MEK/ERK joue également un rôle important dans la survie cellulaire via l'induction de signaux anti-apoptotiques, en particulier par phosphorylation de la protéine pro-apoptotique

Bad, entraînant une séquestration de la molécule par les protéines de la famille 14-3-3 et permettant l'induction de la protéine anti-apoptotique BCL2. Plus récemment les travaux de Allan et al. ont montré que la kinase ERK pouvait phosphoryler/inactiver la caspase 9 et ainsi inhiber l'apoptose (Allan, Morrice et al. 2003).

2.2.5 Voie PI3K/Akt

La voie PI(3)K/Akt transmet, comme la voie Ras/MAPK, des signaux provenant de nombreux récepteurs à tyrosine kinase (**Figure 13**). Les mutations de différents composants de cette voie sont très fréquentes dans la majorité des cancers et leur caractère oncogénique (mutations gain de fonction) est démontré.

Dans les CC, la plupart des mutations affectent le gène PIK3CA codant pour la sous-unité catalytique alpha de la PI(3)Kinase (p110 α). 40% des CC sporadiques ont une mutation sur l'un des acteurs de la voie PI(3)K/Akt, dont PTEN, Akt2 ou p85 (Samuels and Ericson 2006). Ces mutations sont généralement exclusives, suggérant qu'une atteinte unique de la voie est suffisante pour procurer un avantage de sélection.

La kinase Akt possède de très nombreux substrats ayant un rôle mitogénique ou d'induction de survie cellulaire, direct ou indirect : - Phospho-Mdm2, transloquée au noyau pour induire la dégradation de p53 (Ogawara, Kishishita et al. 2002) - Phospho-p21, stabilisée pour induire la survie cellulaire (Zhou, Liao et al. 2001) - Phospho-Bad, séquestrée par 14-3-3, pour inhiber l'apoptose (Nomura, Shimizu et al. 2003).

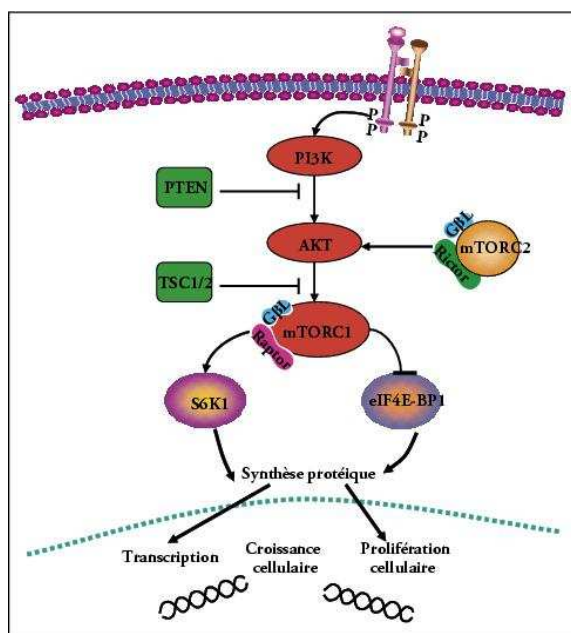


Figure 13 : La voie PI(3)K/Akt (Levy, Sauvin et al. 2008).

2.3 Génétique et cancer du côlon

Au cours des 15 dernières années, de très nombreux travaux ont permis de mieux définir la cancérogenèse des CC sporadiques et héréditaires.

On peut actuellement différencier deux catégories de CC :

- les cancers héréditaires, c'est à dire dus à des mutations germinales de forte pénétrance, qui représentent environ 5% des cas, avec la Polypose Adénomateuse Familiale (FAP, 0,1% des CC) et le syndrome HNPCC (Hereditary Non Polyposis Colorectal Cancer, 3 à 4% des CC) connu également sous le nom de syndrome de LYNCH,
- les cancers sporadiques, c'est à dire qui surviennent suite à des mutations somatiques spontanées (et non des mutations héritées), qui représentent environ 95% des cas (van Krieken and van de Velde 2002; Hamelin and Duval 2003).

2.3.1 Instabilité chromosomique

L'instabilité chromosomique est le type le plus commun d'instabilité génomique concernant des régions entières ou ciblées de chromosomes et qui résulte en la perte ou l'amplification d'un ou plusieurs allèle(s) ; en cas de perte, on parle de perte d'hétérozygotie ou LOH (Loss of Heterozygosity ; **Figure 14**). L'aneuploïdie qui en résulte caractérise de nombreux cancers dont le CC.

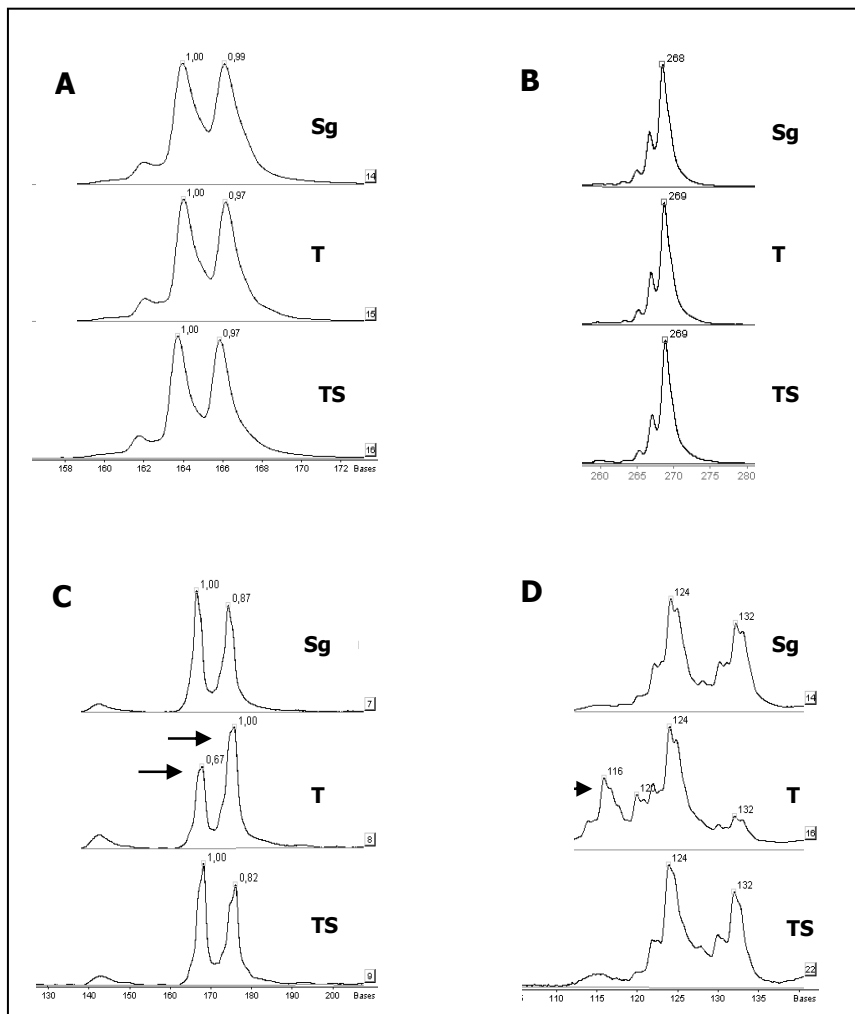


Figure 14 : Exemples de profils d'amplification de microsatellites.

Le profil obtenu pour les cellules de la tumeur (T) est comparé à celui du sang (Sg) et du tissu sain (TS) associés.

- A.** : Exemple de microsatellite informatif (hétérozygote) sans altération, la hauteur relative des pics des deux allèles est constante dans Sg, T et TS.
- B.** : Exemple de microsatellite non informatif (homozygote) : un seul pic dans Sg, T, TS.
- C.** : Déséquilibre allélique (AI), la hauteur relative des pics correspondant aux deux allèles est différente dans les cellules tumorales (T) par rapport aux cellules saines (Sg, TS). Ces profils caractérisent le phénotype CIN (chromosome instable ou LOH) : Ces profils caractérisent le phénotype MIN (instabilité de microsatellite) avec la présence de pics additionnels dans les cellules tumorales (T) (flèche).

2.3.1.1 Anomalies génomiques dans le CC

Plusieurs études par cytogénétique, CGH conventionnelle et FISH ont identifié des altérations chromosomiques spécifiques récurrentes présentes dans les adénomes et retrouvées dans les carcinomes, suggérant leur implication dans la transition adénome - carcinome (Ried, Knutzen et al. 1996; Leslie, Pratt et al. 2003; Leslie, Stewart et al. 2006, Tsafrir, Bacolod et al. 2006). Les résultats

concernant la transition carcinome – métastase ont été repris dans une méta-analyse de 30 études parue en 2006 et étudiant 720 tumeurs primaires et 139 métastases hépatiques (Diep, Kleivi et al. 2006). Cette étude a permis de souligner les altérations les plus fréquemment observées et d'identifier des altérations précoces associées à une plus grande fréquence et celles plus tardives (**Figure 15**). Ainsi, les gains du 8q, 13q, 20q et les pertes du 17p, 18, 22q seraient des événements précoces. Les gains du 7p, 17q et les pertes du 4p, 8p seraient plutôt impliqués dans la transition carcinome invasif – métastases précoces et les pertes du 14q, 17q et gains du 1q, 9p, 11, 12p, 19 seraient plutôt retrouvés dans des métastases bien établies.

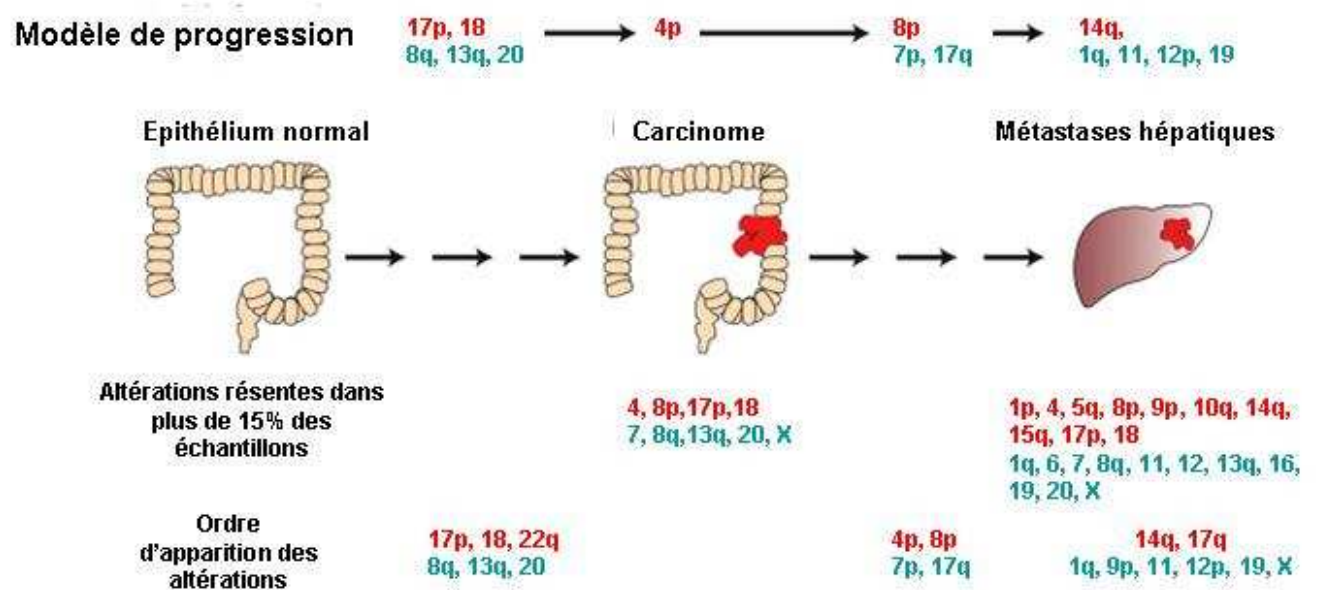


Figure 15 : Progression des cancers coliques de phénotype CIN proposé par Diep et al. suite à une méta-analyse de 30 articles basés sur des approches de cytogénétiques. Les gains sont indiqués en rouge et les pertes en vert (Diep, Kleivi et al. 2006).

2.3.1.2 Nature des gènes altérés au cours de la progression tumorale

a. Foyers de cryptes aberrantes – Adénomes

Il est largement accepté que la formation des carcinomes et le passage de FCA au cancer peut se faire par le biais d'une instabilité chromosomique. Selon le modèle d'accumulation séquentielle d'altérations spécifiques proposé par Vogelstein (Fearon and Vogelstein 1990), l'altération du GST *APC* localisé en 5q21 et régulateur négatif de la prolifération cellulaire, semble être un évènement précoce, sinon initiateur, de la carcinogenèse (Powell, Zilz et al. 1992). En effet, des altérations du gène *APC* sont déjà observées dans les FCA et les adénomes

avec dysplasie de bas grade (Smith, Stern et al. 1994; Pretlow and Pretlow 2005). Les mutations du gène *APC* généralement associées à une perte du second allèle par délétion (Spirio, Samowitz et al. 1998) ont pour conséquence, l'apparition d'un codon stop prématuré et la synthèse d'une protéine tronquée. Les mutations sont plus particulièrement localisées dans le « Mutation Cluster Region » à l'extrémité 5' de l'exon 15 (acides aminés 1263 à 1587). La fonction du gène *APC* est bien connue dans la voie de signalisation Wnt (voir page 26). Par ailleurs, grâce à son interaction avec les microtubules et les kinétochores des chromosomes, elle participe au bon déroulement de la mitose et son altération entraîne une anomalie du nombre de chromosomes (Dikovskaya, Li et al. 2010).

KRAS (*KRAS2*) est un proto-oncogène localisé en 12p12.1, ayant un rôle majeur dans la croissance et la différenciation cellulaire (voir § 2.1.4). Les mutations de *KRAS* produisent une protéine constitutivement active et surviennent durant les stades intermédiaires du développement des adénomes mais se retrouvent déjà dans les FCA (Smith, Stern et al. 1994). Il a été suggéré que la dysplasie ne se développe que lorsque les mutations de *KRAS* se produisent en association avec d'autres mutations telles que celles du gène *APC* (Phelps, Chidester et al. 2009).

b. Transition adénome – carcinome

Le GST *DCC* (Deleted in Colorectal Cancer) présente des homologies avec la famille des molécules d'adhésion NCAM qui ont un rôle dans l'adhésion et la reconnaissance cellulaire. La perte allélique du gène *DCC*, localisé en 18q21.2, survient dans 47% des adénomes avec dysplasie de haut grade et 73% des carcinomes et semble donc être un événement plus tardif dans la progression. L'injection d'une copie normale du chromosome 18q dans des cellules cancéreuses fait disparaître le phénotype cancéreux (Goyette, Cho et al. 1992). L'inactivation du gène *DCC* serait responsable des troubles de l'adhésion cellulaire, de l'invasion et des métastases.

Les gènes *SMAD2* et *SMAD4*, codant pour des protéines impliquées dans la cascade de signalisation TGF- β /SMAD (voir § 2.1.3), sont eux aussi localisés en 18q (18q21.1) ; ils induisent une inhibition de la prolifération cellulaire, et seraient impliqués dans l'induction de la transition épithélio-mésenchymateuse (Zavadil and Bottinger 2005). Des mutations des gènes *SMAD2* et *SMAD4* ciblent 5% et 15% des CC respectivement.

c. Altérations plus tardives

TP53, GST localisé en 17p13.1 (voir § 2.1.1), code pour un facteur de transcription dont les fonctions essentielles d'intégration de stress cellulaire et de régulation des programmes de mort cellulaire (apoptose/sénescence) ou d'arrêt du cycle l'ont désigné comme le « Gardien du génome » (Levine 1997). Les conséquences d'une perte d'activité de p53 sont donc multiples. Des mutations « perte de fonction » ont été mises en évidence dans plus de 40% des CC, tous stades confondus, avec une augmentation de leur fréquence aux stades les plus avancés. Ces mutations touchent principalement le site de fixation de la protéine à l'ADN, diminuant fortement son activité transcriptionnelle (Vousden and Lu 2002). La présence de ces mutations accroît l'instabilité génétique ainsi que la résistance à l'induction de la mort cellulaire et participe aux mécanismes de la tumorigenèse (**Figure 16**).

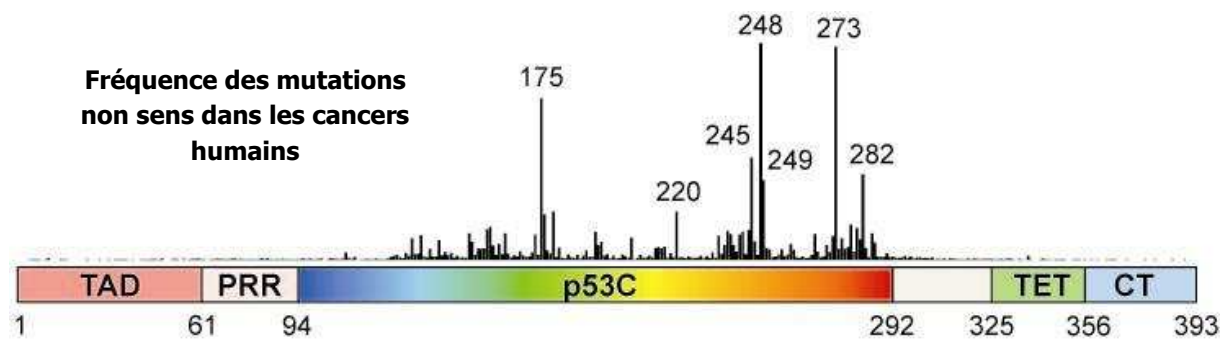


Figure 16 : Représentation de la protéine p53 (393 acides aminés) en alignement avec un histogramme de mutations représentant la fréquence de mutations et leur position sur la séquence codante (Joerger and Fersht 2008).

En résumé, l'instabilité chromosomique est le phénomène le plus fréquent des CC, car environ 85% des CC sont de phénotype CIN. Mais ces modèles reposent sur des fréquences d'altérations observées sur des cohortes relativement faibles

2.3.2 Instabilité des microsatellites

Ce mécanisme caractérise les cancers des syndromes HNPCC et environ 15% des carcinomes sporadiques. Il se caractérise par une réparation déficiente de l'ADN (déficient DNA mismatch repair) ou MMR. Jusqu'à ce jour, 6 gènes MMR ont été identifiés, mais seulement quatre d'entre eux ont une fréquence cliniquement significative. Il s'agit des gènes MLH1, MSH2, MSH6

(Peltomaki and Vasen 1997) et PMS2 (Peltomaki 2005). Ces gènes codent pour un complexe de protéines dont la fonction est de réparer les erreurs de réplication de la polymérase de l'ADN.

En effet de nombreuses erreurs surviennent lors des réplifications de l'ADN, particulièrement dans les séquences microsatellites. Un individu porteur d'une mutation d'un seul allèle d'un gène MMR ne présente pas d'instabilité des microsatellites. Mais, très tôt dans le processus de cancérogénèse, la fonction du gène va être perdue par l'altération du second allèle sain et l'apparition d'une cellule avec une réparation déficiente de l'ADN (appelée MMR-). En l'absence de réparation post-répllicative efficace dans la tumeur, les erreurs persistent et se transmettent lors de la réplication suivante entraînant l'émergence et la persistance d'allèles de tailles différentes. Ainsi, le complexe de réparation de l'ADN des cellules MMR- ne va plus contrôler les erreurs de réplication dans les séquences répétitives des microsatellites, et donc, pour un même microsatellite, la fréquence de ces répétitions va varier au sein du tissu tumoral.

Une instabilité des microsatellites ou MSI (Microsatellite Instability) s'observera après amplification d'un microsatellite, qui dans le tissu sain détectera 2 allèles distincts, alors que dans le tissu tumoral plus de deux allèles seront identifiés (**Figure 14**). Une cellule MMR- présente 1000 fois plus de mutations qu'une cellule normale (Eshleman, Lang et al. 1995).

Ces mutations ciblent essentiellement

- i) des gènes régulateurs de la croissance, comme le récepteur du TGF- β (*TGFBR2*) ou le récepteur de l'IGF-II, permettant aux cellules tumorales d'échapper aux effets anti-prolifératifs des acteurs la voie du TGF- β ;
- ii) des gènes régulateurs du cycle cellulaire (*E2F4*) ;
- iii) des gènes régulateurs de l'apoptose (*Bax*) et
- iv) des gènes de réparation de l'ADN eux-mêmes (*hMSH3* et *hMSH6*), accélérant l'accumulation des mutations. D'autres gènes ont été identifiés et sont résumés dans le **Tableau 5**.

Gene	Chromosome	Microsatellite tract	MSI colorectal cancer mutation frequency (%)	Normal function
<i>TGFBR2</i>	3p22	A ₁₀	90	TGF-β signaling
<i>ACVR2</i>	2q22-23	A ₈ and A ₈	86	Activin signaling
<i>IGFIR</i>	6q26-27	G ₈	10	Insulin-like growth factor and TGF-β signaling
<i>BAX</i>	19q13.3-q13.4	G ₈	50	Apoptosis
<i>hMSH3</i>	5q11-q12	A ₈	50	DNA MMR
<i>hMSH6</i>	2p16	C ₈	33	DNA MMR
<i>E2F-4</i>	16q21-q22	(CAG) ₁₃	65	Cell cycle control
<i>PTEN</i>	10q23.3	A ₆ and A ₆	19-34	Growth regulation
<i>MBD4 (MED1)</i>	3q21-q22	A ₁₀	40	DNA repair and binding to methylated DNA
<i>TCF4</i>	10q25.3	A ₉	39	Growth regulation
<i>CHK1</i>	11q22-q23	A ₉	10	G ₂ cell cycle checkpoint
<i>STK11 (LKB1)</i>	19p13.3	C ₈	<2	Signal transduction
<i>BLM</i>	15q26.1	A ₉	<18	Chromosome stability/DNA repair; helicase
<i>Caspase-5 (ICErel-III)</i>	11q22.2-22.3	A ₁₀	62	Apoptosis
<i>CDX2</i>	13q12.3	C ₆ and G ₆	<2	Homeobox protein
<i>TBP</i>	6q27	(CAG) ₁₃ and (CAG) ₁₆	83	TATA binding protein
<i>RIZ</i>	1p36	A ₈ and A ₉	26	Interacts with RB
<i>hRAD50</i>	5q31	A ₉ and A ₈	31	DNA repair
<i>SEC63</i>	6q21	A ₁₀	49	ER chaperone protein
<i>AIM2</i>	1q22	A ₁₀	48	Interferon-inducible protein

Tableau 5 : Fréquence de mutations des gènes cibles dans les cancers coliques de phénotype microsatellite instable (MIN ou MSI ou MMR-) (Grady and Carethers 2008).

Une étude de l'équipe de Boland et al. a proposé un panel minimal de 5 microsatellites pour le typage moléculaire des CC afin de les classer en MSI-H (MSI-High, avec au moins 2 microsatellites instables), MSI-L (MSI-Low, avec 1 microsatellite instable) et MSS (Microsatellite stable, aucun microsatellite instable) (Boland, Thibodeau et al. 1998). Si au moins 30% des microsatellites analysés ont un profil MSI, la tumeur est considérée MSI.

Du point de vue clinique, les tumeurs de type MSI ont surtout une localisation proximale, une morphologie de carcinome hypermucipare ou de carcinome peu indifférencié, un infiltrat lymphocytaire important, une meilleure survie en comparaison des CC de phénotype CIN et une meilleure réponse aux traitements (Popat, Hubner et al. 2005).

Actuellement une pentaplex de 4 microsatellites homozygotes et un microsatellite hétérozygote sont recommandés pour établir le statut MSI (Boland, Thibodeau et al. 1998), qui est défini par la présence d'au moins 2 microsatellites instables sur les 5 analysés.

2.3.3 Formes héréditaires

2.3.3.1 La Polypose Adénomateuse Familiale (FAP)

En 1986, Herrera et Sandberg démontrent la présence d'une délétion chromosomique en 5q21 chez les malades porteurs de polypose familiale multiple (Herrera, Kakati et al. 1986). Le gène *APC* fut ensuite isolé (Grodin, Thliveris et al. 1991). La FAP est une maladie autosomique dominante à forte pénétrance, liée à la mutation constitutionnelle délétère du gène *APC*, responsable de 1% des CC de phénotype Chromosome Instable (CIN).

Cette maladie se traduit par l'apparition de nombreux adénomes dans le côlon au moment de la puberté. Le gène *APC*, constitué de 15 exons, code une protéine de 2843 acides aminés. La position de la mutation sur le gène *APC* est corrélée au phénotype de la maladie à savoir le nombre de polypes (polypose classique ou atténuée) (Olschwang 2001).

Dans le cadre de la FAP, le risque de CC est quasi de 100% pour les sujets non dépistés avant 50 ans. Au moment du diagnostic, une colectomie préventive est programmée (Laurent-Puig, Blons et al. 1999; Fearnhead, Britton et al. 2001).

2.3.3.2 Le syndrome HNPCC

Le syndrome HNPCC est dû à une mutation délétère des gènes codant pour des protéines de la réparation des mésappariements des bases de l'ADN : hMSH2 ou hMLH1 et le risque de CC avoisine les 80% à 80 ans (Aarnio, Mecklin et al. 1995). Les CC sont de phénotype Microsatellite Instable (MIN). Le risque est constant au cours de la vie et est proche de 1% par an. D'autres gènes porteurs de mutations délétères ont été identifiés et notamment hMSH6, PMS1, PMS2, gènes également impliqués dans la réparation des mésappariements, et EXO1, plus récemment identifié, qui code pour un partenaire de hMSH2 (Wu, Berends et al. 2001).

2.3.3.3 La Polypose Associée à MUTYH (MAP)

L'implication du gène *MUTYH* dans la prédisposition au CC a été découverte plus récemment (Sampson, Dolwani et al. 2003). L'inactivation bi-allélique de ce gène appartenant au système BER (Base Excision Repair) entraîne une polypose cliniquement semblable à la FAP, mais à transmission autosomique récessive.

2.3.3.4 Le syndrome de Peutz-Jeghers (SPJ)

Le SPJ est caractérisé par des polypes hamartomateux gastro-intestinaux, des hyperpigmentations muco-cutanées typiques et le caractère familial de l'atteinte polypeuse. Les polypes du SPJ se manifestent avant tout au niveau de l'intestin grêle (70 à 90%), du colon (50%) ou au niveau de la muqueuse gastrique. Mais ils peuvent aussi être présents dans la vésicule biliaire, dans le système uro-génital ou même dans les voies respiratoires.

Au niveau du tractus gastro-intestinal, les polypes ont tendance à la dégénérescence maligne (van Lier, Wagner et al. 2010). De plus, le SPJ est associé à un risque élevé de tumeurs malignes extra-gastrointestinales aux phases précoces de la vie, comme le carcinome du pancréas, les tumeurs de la lignée germinale dans les ovaires, le carcinome du col utérin et le carcinome mammaire chez la femme, les cancers testiculaires chez l'homme.

Des mutations germinales du Gène LKB1/STK11 sont détectées dans plus de 50% des cas familiaux (Lim, Hearle et al. 2003).

2.3.3.5 La polypose juvénile familiale (PJF)

Cette polypose rare, liée à l'âge, est caractérisée par la présence de nombreux polypes du côlon et du rectum, mais aussi de l'estomac, du duodénum et du grêle. Les polypes sont majoritairement hamartomateux mais il existe tout de même une prédisposition au CC avec un risque évalué à 40%. C'est une maladie autosomique dominante pour laquelle des mutations des gènes *PTEN* et *SMAD4* ont été identifiées (Teniere, Songne et al. 2002).

2.3.3.6 La polypose hyperplasique (PH)

La polypose hyperplasique, décrite pour la première fois par Spjut et al (Spjut and Estrad, 1977) est définie selon les critères OMS par la présence d'au moins 5 polypes hyperplasiques en amont du côlon sigmoïde (dont 2 de plus de 1 cm) ou la présence de 30 polypes hyperplasiques sans localisation préférentielle ou la présence d'au moins un polype hyperplasique en amont du côlon sigmoïde dans un contexte de polypose hyperplasique familiale (Jass, Iino et al. 2000). La fréquence de ces polyposes dans la population générale n'est pas connue (Carvajal-Carmona, Howarth et al. 2007).

2.3.3.7 Le syndrome de Cowden (SC)

Le SC est caractérisé par la présence d'hamartomes dans différents organes, notamment dans le tractus gastro-intestinal. Les individus atteints ont un risque majoré de tumeurs de la thyroïde, de la glande mammaire, de l'endomètre et de la peau. Le risque de développement de carcinome colorectal n'est pas encore bien connu. Chez une proportion variable de patients atteints de SC, des mutations de la lignée germinative du gène *PTEN* ont été trouvées (Nelen, Padberg et al. 1996).

Les informations principales concernant ces formes héréditaires sont consignées dans le Tableau 6.

Forme clinique	Gènes impliqués	Transmission	Traduction histologique	Fréquence	Risque de cancer
FAP	<i>APC</i>	Autosomique dominante	adénome	1:10 000	100%
HNPCC	<i>hMSH2</i> <i>hMSH6</i> <i>hPMS1/2</i> <i>Exo1</i>	Autosomique dominante	adénome	1:1000/10 000	>80%
MAP	<i>MUTYH</i>	Autosomique récessive	adénome	1 :100 000	70%
SPJ	<i>LKB1 (STK11)</i>	Autosomique dominante	hamartome	1:200 000	40%
PJ	<i>SMAD4</i> <i>BMPR1A</i>	Autosomique dominante	hamartome	1:100 000	40%
PH	?	?	polype festonné	?	40%
SC	<i>PTEN</i>	Autosomique dominante	hamartome	1:200 000	faible

Tableau 6 : Récapitulatif des polypes gastro-intestinaux avec prédisposition génétique.

2.2.4 Hyperméthylation

Au sein du génome, il existe des régions enrichies en dinucléotides CpG, appelées îlots CpG, représentant 1-2% du génome, présentes dans la région 5' de 50%-60% des gènes et qui sont normalement maintenues non méthylées. Le processus de méthylation est réalisé par des DNA

méthyltransférases (DNMT) qui induisent l'addition enzymatique d'un groupement méthyl sur le carbone en 5' d'une cytosine.

Lorsque certains des îlots CpG deviennent méthylés, cela entraîne une répression transcriptionnelle du gène par le changement de structure et d'accessibilité de l'ADN (**Figure 17 et Tableau 7**).

Dans les cancers, l'hyperméthylation des GST est un phénomène fréquent, impliqué dans la cancérogenèse et la progression tumorale (Deltour, Chopin et al. 2005).

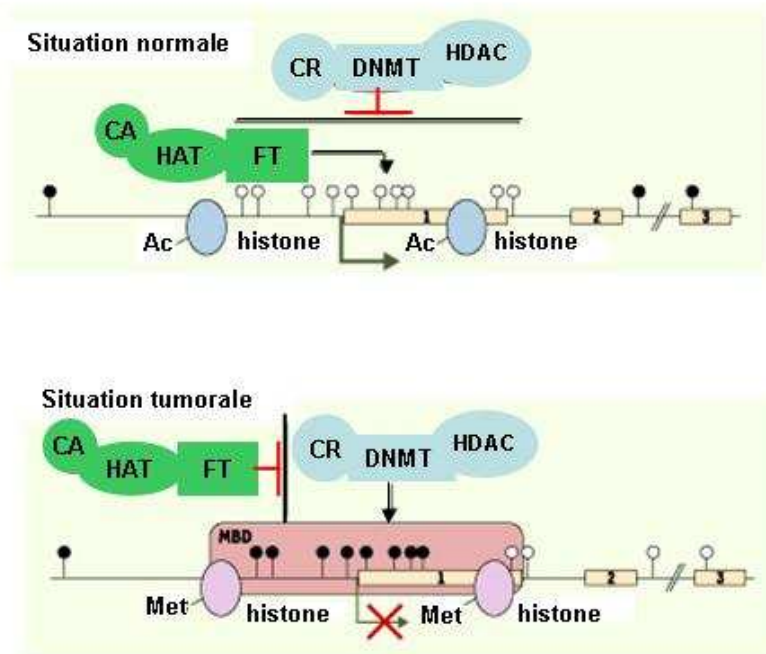


Figure 17 : Aspects de la chromatine au niveau de gènes suppresseurs de tumeur, dans une situation normale et tumorale avec méthylation. Dans une situation normale, la plupart des CpG en dehors des îlots sont méthylés (cercles noirs), tandis que les CpG des îlots ne sont pas méthylés (cercles blancs). Dans cette région promotrice, l'acétylation des histones permet de maintenir la chromatine dans un état « relâché », accessible aux complexes de transcription. Dans une situation tumorale, les profils de méthylation sont inversés, avec une hypométhylation des CpG répartis le long du génome et dans les régions codantes, et une hyperméthylation des îlots CpG au niveau des promoteurs (Deltour, Chopin et al. 2005).

Gènes	Fréquence (%)
MLH1	10 - 20
TIMP3	20 - 30
P14	10 - 20
RIZ1	10 - 20
SOCS1	20
RUNX3	> 20
CDKN2A	20 - 30
MEGALIN	10 - 20
NEUROG1	20
CACNA1	20
APC	10 - 20
CALCA	50
MiR342/EVL	70 - 80
TSP1	20 - 30
HIC1	50
HLTF	50
VIM	60 - 80
MGMT	30 - 40
SFRP2	60 - 80
SLC5A8	40 - 60

Tableau 7 : Gènes affectés par des modifications épigénétiques dans les CC

(adapté d'après Grady and Carethers 2008).

Le phénotype qui résulte de la méthylation d'îlots CpG dans les tumeurs a été identifié comme le phénotype méthylateur ou CIMP par Toyota et al. en 1999 (Toyota, Ahuja et al. 1999). Ce phénotype est caractérisé par un fort taux de méthylation des îlots CpG et par son association majoritaire au phénotype MSI, avec la méthylation du promoteur du gène de réparation de l'ADN, MLH1 (Alemayehu, Sebova et al. 2008).

Cependant, ce phénotype est aussi décrit dans les tumeurs de type CIN (Ostwald, Linnebacher et al. 2009, Walther, Johnstone et al. (2009). Ce phénotype est décrit dans différents types de cancers, dont le côlon, l'estomac (Toyota, Ahuja et al. 1999), le poumon (Marsit, Houseman et al. 2006), le foie (Shen, Ahuja et al. 2002), l'ovaire (Strathdee, Appleton et al. 2001), dans des leucémies (Roman-Gomez, Jimenez-Velasco et al. 2005), mais les gènes cibles semblent spécifiques à chaque tumeur.

Pour identifier le phénotype CIMP dans le côlon, Issa et al. ont suggéré l'analyse des promoteurs des 5 gènes MINT1, MINT2, MINT31, CDKN2A et MLH1 (Issa 2004), et en 2006, Weisenberger et al. ont analysé 195 loci et défini un autre panel de 5 gènes (CACNA1G, IGF2, NEUROG1, RUNX3 et SOCS1) (Weisenberger, Siegmund et al. 2006). En complément, une étude réalisée par le groupe d'Ogino et al. a abouti à une classification des CC en fonction de la fréquence de méthylation des îlots CpG dans le panel étudié : CACNA1G, IGF2, NEUROG1, RUNX3, SOCS1, CDKN2A/p16 et CRABP1 et à la distinction des tumeurs CIMP-high, CIMP-low et Non-CIMP (Ogino, Kawasaki et al. 2007, **Tableau 8**).

L'ensemble de ces études indiquent que les CC non-CIMP sont plutôt associés à des mutations du gène *TP53*, les CC CIMP-low à des mutations du gène *KRAS* et les CC CIMP-high à des mutations du gène B-Raf.

	Non-CIMP	CIMP-low	CIMP-high
Localisation	distal > proximal	-	proximal > distal
Sex ratio	hommes = femmes	hommes > femmes	femmes > hommes
<u>Mutations</u>			
BRAF	sauvage	sauvage	muté
KRAS	muté	muté	sauvage
p53	70%	30%	10%
Statut	CIN	CIN	MIN

Tableau 8 : Caractéristiques des tumeurs de phénotype CIMP (adapté d'après Ogino, Kawasaki et al. 2007).

Il est important de pouvoir identifier les CC CIMP puisque si aucune chimiothérapie n'est administrée après chirurgie, la survie globale est mauvaise. Au contraire, avec une chimiothérapie à base de 5-Fluorouracile (5FU), les patients CIMP ont une bien meilleure survie globale après chirurgie que les patients CIMP sans chimiothérapie post-opératoire (Van Rijnsoever, Elsleh et al. 2003).

Les mécanismes à l'origine de ce phénotype sont encore mal connus malgré le nombre important d'études récentes. Il existe actuellement deux hypothèses pour expliquer les mécanismes induisant une méthylation aberrante : elle résulterait d'une surexpression des DNMT (Bakin and Curran 1999, Nosh, Shima et al. 2009) ou d'une altération des facteurs protégeant normalement l'ADN d'une méthylation aberrante, comme les histones acétyltransférases (HAT) qui maintiennent l'ADN dans un état transcriptionnellement actif (Wong, Hawkins et al. 2007).

A l'heure actuelle, il est clairement admis que l'inactivation de GST par méthylation est une cause majeure d'instabilité génomique et que la coopération d'altérations génétiques et épigénétiques est à l'origine de l'initiation de la carcinogenèse. Comme l'enzyme responsable de la maintenance du statut méthylé et de l'activité méthyltransférase de novo est codé par le gène *DNMT1*, plusieurs études centrées sur les inhibiteurs de DNMT1 et plus précisément le 5-aza-2'-deoxycytidine (decitabine) et le 5-azacytidine (5-azaC) sont en cours (Goffin and Eisenhauer 2002).

L'hyperméthylation est un processus largement décrit dans les CC (Toyota, Ahuja et al. 1999), mais aussi dans les lésions précancéreuses comme les adénomes (Derks, Postma, et al. 2006, Judson, Stewart, et al. 2006), suggérant sa possible implication dans l'initiation de la transformation cancéreuse.

2.4 Modèles de carcinogenèse

2.4.1 Modèle de Vogelstein

En 1990, un premier modèle de progression des CC a été proposé par Fearon et Vogelstein (Fearon and Vogelstein 1990) montrant que l'évolution de l'épithélium normal vers la lésion cancéreuse impliquerait l'accumulation de mutations activatrices d'oncogènes et de mutations inactivatrices de GST selon une séquence précise (**Figure 18**). Ce modèle a été établi par l'étude moléculaire comparative d'épithélium normaux, d'adénomes et d'adénocarcinomes, en étudiant les fréquences d'altérations dans ces différents groupes.

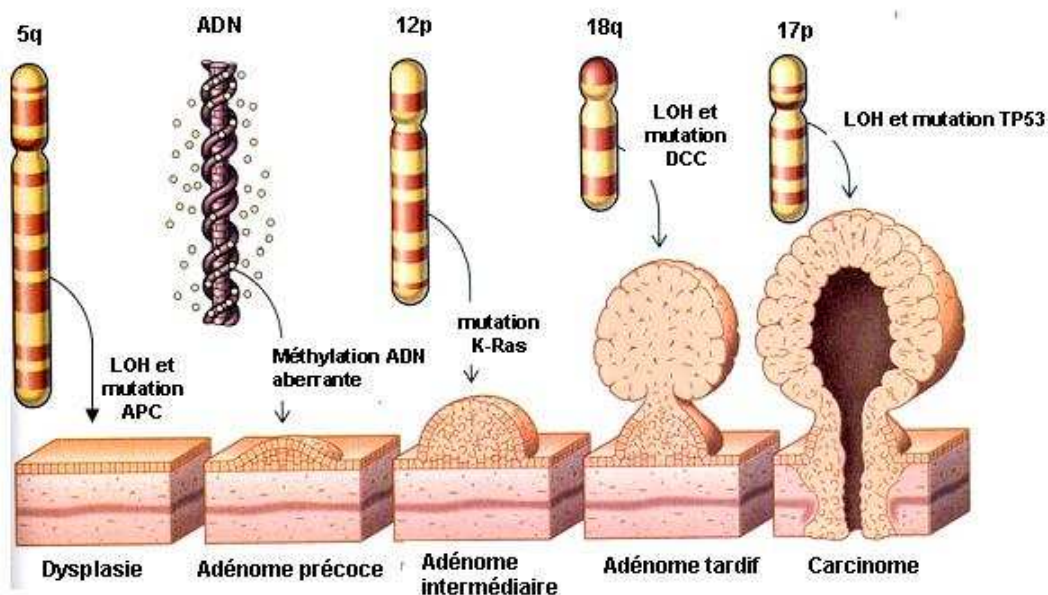


Figure 18 : Identification de gènes altérés pendant la progression tumorale (Fearon and Vogelstein 1990). Dans ce modèle, l'accumulation séquentielle dans le temps d'altérations génomiques de GST et d'oncogènes permet aux cellules tumorales d'acquies un potentiel métastatique. Ainsi, des altérations ont été identifiées avec dans un premier temps l'altération des GST APC et β -caténine, puis de l'oncogène KRAS et plus tardivement le GST TP53.

L'existence de deux grandes voies de cancérogenèse colique, correspondant à deux types d'instabilité génomique, l'instabilité de microsatellites (MIN) et l'instabilité chromosomique (CIN), et du phénotype CIMP a depuis conduit à revoir ces séquences tout en reconnaissant l'importance de l'accumulation séquentielle d'altérations (Jones, Chen et al. 2008) dans cette séquence de transformation épithélium – adénome - carcinome.

2.4.2 Voie festonnée

Cette voie de carcinogenèse décrit la progression des polypes festonnés vers le cancer colique. En 1990, Longacre and Fenoglio-Preiser proposent le terme d'adénome festonné pour des polypes comportant des aspects de polype hyperplasique et d'adénome (Longacre, Fenoglio-Preiser, 1990). Torlakovic and Snover en 1996, puis Torlakovic en 2003 (Torlakovic, Skovlund et al. 2003) commencent à démembrer ces polypes festonnés en distinguant les adénomes festonnés sessiles et les adénomes festonnés traditionnels.

Une nouvelle nomenclature est rapportée par Snover et al (Snover, Jass et al. 2005) qui montre une filiation entre polype hyperplasique, classiquement considéré comme un polype bénin sans potentiel de dégénérescence, adénome festonné et adénocarcinome.

Les bases moléculaires de cette voie festonnée ont été bien étudiées pour les adénomes sessiles festonnés. Une mutation activatrice de BRAF est un évènement précoce, déjà présente au sein d'ACF et de polypes hyperplasiques (Rosenberg, Yang et al. 2007, Spring, Zhao et al. 2006), responsable d'une prolifération cellulaire. Une hyperméthylation mise en évidence par l'existence d'un phénotype CIMP (Weisenberger, Siegmund et al. 2006), inhiberait alors des gènes responsables de contrôle cellulaire par sénescence comme p16INK4a or IGFBP7, conduisant à la formation de l'adénome festonné sessile. La dégénérescence maligne passerait par une inhibition du gène MLH1 par hyperméthylation de son promoteur. Les carcinomes associés aux lésions festonnées sont souvent de phénotype CIMP-High et MIN (O'Brien, Yang et al. 2006).

La progression vers l'adénome traditionnel festonné est plus confuse. Des mutations de BRAF et KRAS ont été mises en évidence (O'Brien, Yang et al. 2006). L'association à une dérégulation de la voie WNT et à un statut CIMP-Low avec perte d'expression de *MGMT* seraient des évènements de cette voie alternative (Jass, Baker et al. 2006).

2.4.3 MGMT

La réparation de l'ADN comprend un ensemble de processus mis en œuvre par une cellule pour identifier et corriger les dommages de l'ADN génomique soumis à des modifications métaboliques et à des facteurs environnementaux portant atteinte à son intégrité. La réparation de la mutation de la guanine en O₆-méthyl-guanine, qui peut alors s'apparier par erreur avec une thymine, induisant une mutation de transition G-C vers A-T, s'effectue par l'enzyme O⁶-Méthyl-Guanine DNA Méthyl Transférase (MGMT). Elle transfère le groupement méthyle de la base vers un de ses acides aminés, une cystéine (Esteller, Hamilton et al. 1999). L'enzyme de réparation est alors alkylée de façon irréversible, et donc à chaque réparation, une molécule d'enzyme est consommée.

L'hyperméthylation du promoteur du gène *MGMT* induisant une perte de son activité enzymatique est décrite dans de nombreux types tumoraux dont le CC (Esteller, Gaidano, et al. 2002, Shen, Kondo, et al. 2005) et peut être associée à des mutations de gènes (Kaina, Christmann, et al. 2007), dont certains fortement impliqués dans la carcinogenèse colique comme *KRAS* et *TP53* (Esteller, Herman, 2004). Son implication dans les stades précoces du CC et dans les adénomes coliques (Krtolica, Krajnovic, et al. 2007, Nagasaka T, Goel, et al. 2008) en font un candidat alternatif potentiel pour l'initiation de la carcinogenèse. Son rôle comme inducteur de cancer de phénotype MIN a été récemment décrit dans les CC survenant dans les maladies inflammatoires de l'intestin (Svreck, Buhard et al. 2010).

2.4.4 Sous-typage moléculaire du phénotype CIN

Actuellement, les seuls marqueurs pronostiques préconisés par l'EGMT (European Group on Tumour Markers) pour le diagnostic sont : la concentration en CEA (Carcinoembryonic Antigen), marqueur utilisé dans la surveillance post-chirurgicale, la présence de sang dans les selles et l'envahissement ganglionnaire (Duffy, van Dalen et al. 2007). Plusieurs études ont proposé des signatures transcriptionnelles spécifiques aux tumeurs évoluant vers des métastases (Croner, Peters et al. 2005, Komuro, Tada et al. 2005), mais pour le moment aucune de ces études n'a conduit à la validation de marqueurs pronostiques. Une hypothèse possible était que les CC de phénotype CIN, représentant 85% des CC, pouvaient relever de plusieurs sous-types moléculaires conduisant à des comportements cliniques différents.

Au laboratoire, un allélotypage ciblant les 18 chromosomes les plus fréquemment altérés dans les CC et réalisé sur 104 carcinomes de phénotypes CIN de tous stades cliniques, a permis de mettre en évidence 3 sous-types de tumeurs selon la fréquence globale d'altérations (**Figure 19**, Weber, Meyer et al. 2007) et les types d'altérations observées. En effet, une clusterisation hiérarchisée à deux entrées a permis d'identifier un sous-type très faiblement altéré ou CIN-Very Low (14±8,1%), un sous-type moyennement altéré ou CIN-Low (38,9±9,4%) et un sous-type fortement altéré ou CIN-High (65,2±16,7%). Chaque stade clinique étant représenté dans chaque sous-type, cela suggérerait que l'accumulation des altérations n'est pas un pré-requis pour le potentiel métastatique d'une tumeur.

Bien que cette étude n'ait pas permis de définir clairement un set de marqueurs pronostiques, elle a permis de proposer un panel minimal d'altérations composé du 8p, 17p, 17q, 18q qui seraient présentes précocement. Dans ce panel, on peut noter l'absence de l'altération du locus associé au gène *APC*, signifiant que l'altération simultanée des loci associés aux gènes *APC* et *TP53* ne serait pas systématique.

Ces observations permettraient de faire l'hypothèse de voies de carcinogenèse colique par accumulation séquentielle d'altérations spécifiques différentes de celles proposées par Vogelstein.

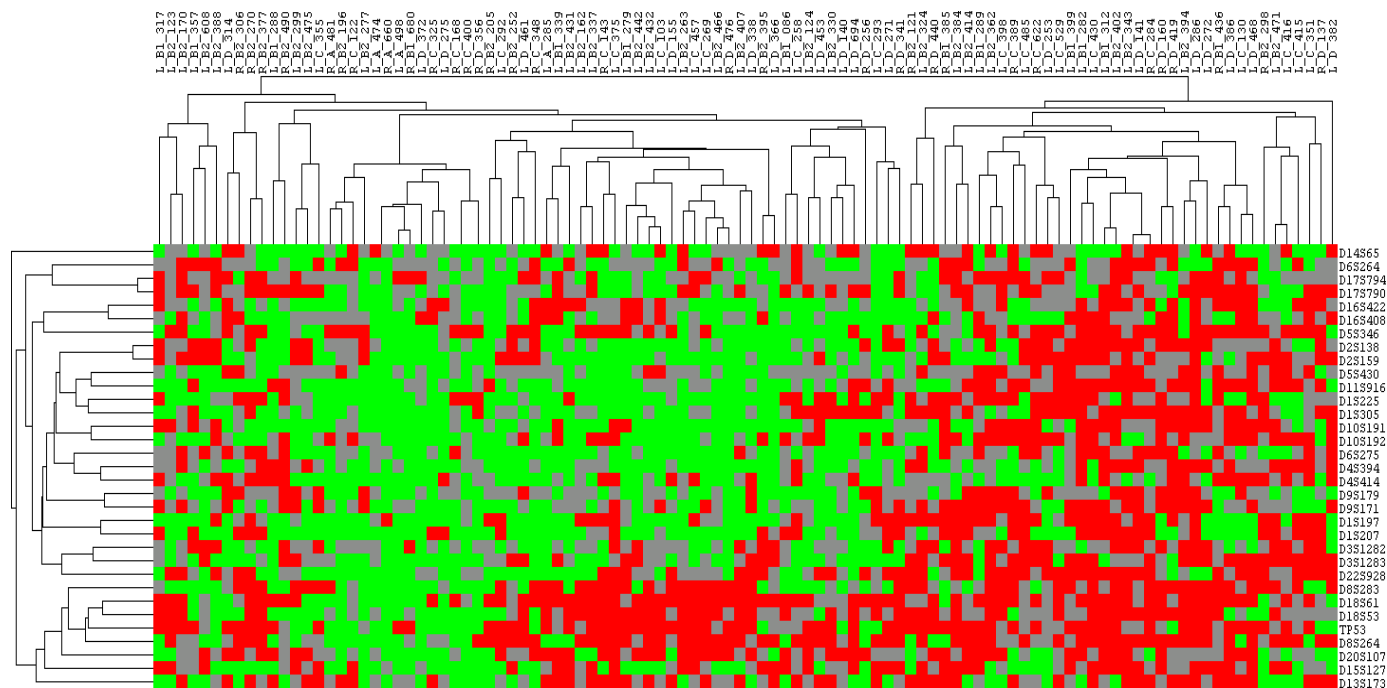


Figure 19 : Clusterisation hiérarchisée à deux entrées d'un allélotypage ciblant 18 chromosomes les plus fréquemment altérés dans le CC (Weber, Meyer et al. 2007). CIN-Very Low : de 0 à 22% ; CIN-Low : de 23 à 48% ; CIN-High : de 49 à 100%.

2. Conséquences cliniques du diagnostic d'adénome colique

3.1 Facteurs pronostiques de récurrence des adénomes coliques

Les facteurs pronostiques prédisposant à un risque plus élevé de récurrence d'adénome avancé (voir § 1.4.2.2) sont retrouvés dans la plupart des études, avec néanmoins quelques discordances.

Pour Bonithon et al (2004), la présence d'une dysplasie de haut grade, de plus de deux adénomes et la localisation colique proximale représentaient des facteurs pronostiques indépendants.

L'âge ≥ 60 ans, la localisation colique gauche, la présence d'un contingent villositaire, la taille ≥ 1 cm et la dysplasie de haut grade apparaissaient comme des facteurs prédictifs selon Gschwantler et al (Gschwantler, Kriwanek et al. 2002).

Winaver et al (Winawer, Zauber et al. 2006) retrouvait également le nombre d'adénomes, la dysplasie de haut grade, la présence d'un contingent villositaire et la taille de plus de 1 cm comme des éléments prédictifs.

Plus récemment une méta-analyse de 8 études prospectives montre que les facteurs les plus prédictifs sont un âge de plus de 60 ans, un nombre ≥ 5 adénomes et une taille d'adénome ≥ 2 cm (Martinez, Baron et al. 2009). Les autres facteurs comme le sexe masculin, la localisation colique proximale et la présence d'un contingent villositaire étaient également prédictifs, mais la dysplasie de haut grade n'était pas un facteur indépendant.

Il apparaît aussi que le nombre d'adénomes est le facteur le plus reproductible, suivi de la taille de l'adénome et de la dysplasie de haut grade.

3.2 Prise en charge des patients porteurs d'adénome

3.3.1 Indications chirurgicales

En cas de dysplasie de haut grade (carcinome intra-épithélial ou intra-muqueux sans extension à la sous-muqueuse = stade pTis), la polypectomie est suffisante si l'exérèse est complète, avec une marge de sécurité de plus de 1 mm (conférence de consensus, 1998).

En cas de tumeur envahissant la sous-muqueuse (stade pT1) :

- Si le polype est pédiculé, la polypectomie est suffisante si
 - l'exérèse est complète, avec une marge de sécurité de plus de 1 mm
 - le carcinome est bien ou modérément différencié
 - il n'existe pas d'emboles vasculaires
 - il n'existe pas de budding

En l'absence d'un de ces critères ou si le carcinome est peu différencié, une colectomie complémentaire est recommandée.

- Si le polype est sessile ou plan, il existe le plus souvent un critère de mauvais pronostic conduisant à un traitement complémentaire.

3.3.2 Périodicité de la surveillance

Les recommandations françaises pour la pratique clinique des indications de la coloscopie, en particulier pour la surveillance des patients porteurs de polypes colorectaux, ont été établies par l'ANAES en 2004 (**Tableau 9**).

Situation	Rythme de surveillance de la coloscopie
Polypes hyperplasiques de petite taille du rectosigmoïde	10 ans
Polypes hyperplasiques > 1 cm et/ou nombre > 5, de siège colique, et/ou dans un contexte familial de polypose hyperplasique	5 ans puis 10 ans
Exérèse complète d'1 ou 2 adénomes non avancés, et absence d'ATCD familial de CCR	5 ans – 5 ans – 10 ans
Exérèse complète d'1 ou 2 adénomes non avancés, et ATCD familial de CCR	/3 ans
Exérèse complète de > 10 adénomes non avancés, et absence d'ATCD familial de CCR	/3 ans
Exérèse complète d'un adénome transformé	/3 ans
Exérèse incomplète d'un adénome de bas grade ou transformé	A 3 mois puis /3 ou 5 ans

Tableau 9 : Indications de surveillance par coloscopie des patients porteurs de polypes coliques (selon les Recommandations pour la pratique clinique des indications de la coloscopie, ANAES 2004).

3.3.3 Consultation d'oncogénétique

Une consultation d'oncogénétique est recommandée (Olschwang, Bonaiti et al. 2004) :

- en cas d'adénome avancé survenant avant 40 ans
- en cas d'adénome avancé survenant entre 40 et 60 ans avec des antécédents familiaux d'adénomes ou de cancer du spectre HNPCC (côlon-rectum, utérus, grêle, urothélium, voies biliaires, estomac, ovaire) chez un apparenté au 1^{er} degré.

3.2 Variabilité inter-observateurs du diagnostic des adénomes coliques

Comme précisé précédemment, un traitement chirurgical complémentaire ou la périodicité de surveillance par coloscopie dépend essentiellement des facteurs pronostiques histologiques établis par le pathologiste.

La modification des types de grade de dysplasie et la classification de Vienne modifiée (Dixon, 2002) ont permis d'améliorer la variabilité inter-pathologistes du diagnostic de grade des adénomes coliques.

Il persiste néanmoins d'importantes variabilités inter-observateurs concernant le diagnostic du sous-type histologique d'adénome, en particulier l'existence d'un contingent vilieux et le diagnostic d'adénome festonné, et le stade d'extension du cancer (Farris, Misdraji et al. 2008, Denis, Peters et al. 2009).

La détermination de marqueurs moléculaires prédictifs utilisables en routine (profil génomique, expression protéique...) pourrait représenter une aide complémentaire à l'établissement des caractéristiques histologiques classiques ou permettre de s'affranchir de critères en partie subjectifs (comme le sous-type histologique des adénomes).

OBJECTIF DU TRAVAIL

Les CC ayant des instabilités chromosomiques (phénotype CIN) représentent la majorité des CC (85%) et malgré une recherche intensive, aucune combinaison d'altérations génomiques n'a pu être validée à visée pronostique. Cette difficulté à identifier des marqueurs pourrait être liée à la grande hétérogénéité inter- et intra-tumorale. De plus, la majorité des articles publiés dans ce domaine depuis 20 ans portent sur l'étude de cohortes hétérogènes de carcinomes incluant souvent des phénotypes MIN et CIN, des carcinomes proximaux, distaux et rectaux, ces derniers étant souvent irradiés avant la résection chirurgicale.

L'identification de tels marqueurs moléculaires reste un objectif prioritaire pour l'amélioration de la prise en charge des patients. Une meilleure connaissance de l'hétérogénéité moléculaire à l'origine des CC pourraient permettre l'identification de marqueurs pronostiques de récurrence d'adénome ou d'évolution vers le carcinome invasif.

Ainsi l'objectif de notre équipe est d'identifier plus précisément, en étudiant des populations clairement définies de CC sporadiques de phénotype CIN localisés dans le côlon proximal et distal,

- les anomalies et leur combinatoire à l'origine de l'initiation des adénomes et de la transition vers un carcinome
- des marqueurs ayant une valeur pronostique.

Différentes études publiées ont identifié des événements tels que des altérations chromosomiques avec les gains du 13q, 20q et les pertes du 5q, 17p, 18, des mutations récurrentes telles que la mutation activatrice de l'oncogène *KRAS*, des *GST APC* et *TP53*, et des modifications épigénétiques (voir § 2.1). En 1990, Vogelstein a proposé un modèle dans lequel l'accumulation séquentielle d'altérations spécifiques conduirait au développement d'un carcinome puis de métastases (**Figure 18**).

Une étude réalisée dans notre laboratoire et basée sur un allélotypage a montré sur une cohorte de 104 carcinomes de phénotype CIN (**Figure 19**), confirmé sur une cohorte de 363 carcinomes de phénotype CIN (**Figure 20**), que ce modèle, bien que clairement admis, ne peut pas expliquer le développement de l'ensemble des carcinomes. De plus, les carcinomes peuvent être classés selon la fréquence globale d'altération, avec des carcinomes « CIN-Very Low » présentant très peu d'altérations, « CIN-Low » moyennement altérés et « CIN-High » fortement altérés, et ce indépendamment du stade clinique du carcinome, suggérant que l'accumulation d'altérations ne peut pas, à elle seule, expliquer la dissémination métastatique. Il semblerait

que ce soit l'association d'altérations particulières qui confère un potentiel métastatique plutôt qu'une quantité d'altérations.

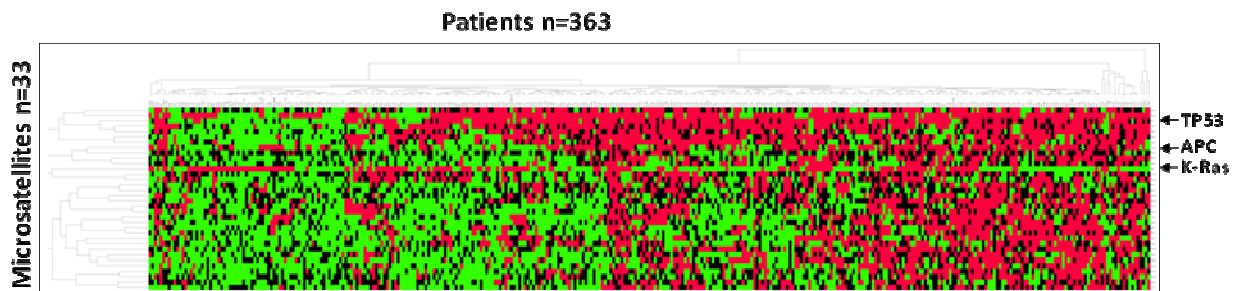


Figure 20 : Clusterisation hiérarchisée à deux entrées d'un allélotypage ciblant 18 chromosomes (Neuville et Guénot, data personnelles).

De plus, les loci associés aux gènes *APC* et *TP53* ne sont pas toujours altérés simultanément, comme le suggérerait le modèle de Vogelstein, puisque certaines tumeurs présentent l'une ou l'autre, ou aucune de ces altérations.

Dans ce contexte, l'objectif de mon travail de thèse était de déterminer :

- **si le sous-typage moléculaire observé dans les carcinomes de phénotype CIN existe à un stade plus précoce que celui d'adénocarcinome invasif ;**
- **si la méthylation du promoteur du gène *MGMT* a un rôle dans l'initiation des adénomes de type CIN, comme cela a été montré dans les CC de type MIN ;**
- **quelles sont les associations d'altérations génomiques possibles au stade d'adénomes coliques ;**
- **si les altérations spécifiques trouvées sont similaires pour les localisations proximale et distale ;**
- **s'il existe un marqueur ou un panel de marqueurs moléculaires prédictif du type histologique ou du grade de dysplasie des adénomes.**

Nous avons évalué l'incidence d'altérations génomiques, en particulier celles proposées par Vogelstein, dans une cohorte de 159 polypes coliques de phénotypes non-MIN et localisés dans le côlon proximal et distal, en excluant le rectum. Seule la perte allélique des altérations ciblant *APC* et *TP53* a été recherchée, sans recherche de la mutation potentiellement associée. En

effet, il a été montré que cette anomalie mutation/perte allélique serait concordante dans la plupart des cas. De plus, le statut mutationnel de *KRAS* et la méthylation du promoteur du gène *MGMT* ont été analysés pour compléter la caractérisation des anomalies génétiques ou épigénétiques participant de façon fréquente à la carcinogenèse colique.

Par la suite, des puces pangénomiques ont été utilisées pour une analyse par CGH (CGH array, Comparative Genomic Hybridization) sur 110 des polypes afin, dans un premier temps :

- d'étendre l'analyse à l'ensemble des chromosomes
- de vérifier l'absence d'altérations dans les polypes sans AI.
- de préciser la nature des altérations observées par allélotypage (gain ou perte)

La réalisation d'un « tissu Micro Array » (TMA) est en cours. Il permettra d'analyser l'expression des protéines d'intérêt identifiées par CGH.

Au cours de ce travail, des collaborations ont conduit à plusieurs publications présentées en Annexes sur :

- La caractérisation d'une des altérations chromosomiques les plus fréquemment retrouvées dans les cancers coliques CIN : l'altération du bras long du chromosome 20 (20q).
- L'utilisation de la spectroscopie infra-rouge comme un nouvel outil de caractérisation histopathologique (analyse préliminaire sur des xénogreffes de carcinomes coliques humains).
- La caractérisation métabolique des cancers coliques en utilisant la résonance magnétique nucléaire à haute résolution à angle magique.

RESULTATS

1. Mise en place de la cohorte de polypes coliques

La cohorte de polypes coliques utilisée pour mon travail de thèse a été établie prospectivement depuis 1996 lors de la mise en place de la collection cancer du côlon initiée en 1994, date de mon arrivée au sein du Département de Pathologie de Strasbourg en tant qu'interne d'anatomie pathologique et de collaboration avec le Service de Biologie Moléculaire. J'ai participé activement à l'échantillonnage des pièces opératoires issues du Service de Chirurgie Générale. Il nous est apparu rapidement important de pouvoir disposer de lésions pré-cancéreuses ou cancéreuses précoces, ce qui m'a amené à congeler les polypes présents sur ces pièces opératoires, seuls ou associés à un cancer invasif. La gestion de cette collection s'est par la suite effectuée par le Centre de Ressources Biologiques dont je suis le responsable depuis 2006.

Entre 1996 et 2006, 300 polypes ont ainsi été congelés et stockés, associés à de la muqueuse colique saine située à plus de 5 cm de la lésion et à un prélèvement sanguin. Un premier allélotypage ciblant 12 microsatellites, dont l'objectif était de détecter en routine les cas de phénotype MIN, a permis de détecter 40 polypes MIN qui ont été écartés de la série. J'ai ensuite revu du point de vue histologique et de la localisation les 260 polypes restants. 50 polypes rectaux (polypes hyperplasiques ou adénomes) et 60 polypes coliques ou rectaux non épithéliaux (hyperplasies lymphoïdes, polypes inflammatoires) ont été éliminés de la cohorte.

La série est donc constituée de 159 polypes coliques, dont 77 de localisation droite et 82 de localisation gauche, avec une répartition assez homogène entre les sous-types histologiques (**Figure 21**). Ces polypes ont été réséqués chez 120 patients dont 5 au cours d'une endoscopie et 154 à l'occasion d'une résection chirurgicale programmée (dont 132 provenant de 101 patients étaient associés à un carcinome invasif synchrone pour lequel nous disposons également du matériel congelé). 27 patients comportaient plus d'un polype (19 cas avec 2 polypes, 9 cas avec 3 polypes et 1 cas avec 6 polypes). Chaque polype a été examiné en histologie par deux pathologistes habitués à la pathologie digestive (AN et AO), sur une coupe paraffine et sur une coupe à congélation à partir l'échantillon congelé. Les discordances sur les sous-types tubuleux et tubulo-villeux et sur les grades ont été facilement résolues après une relecture commune.

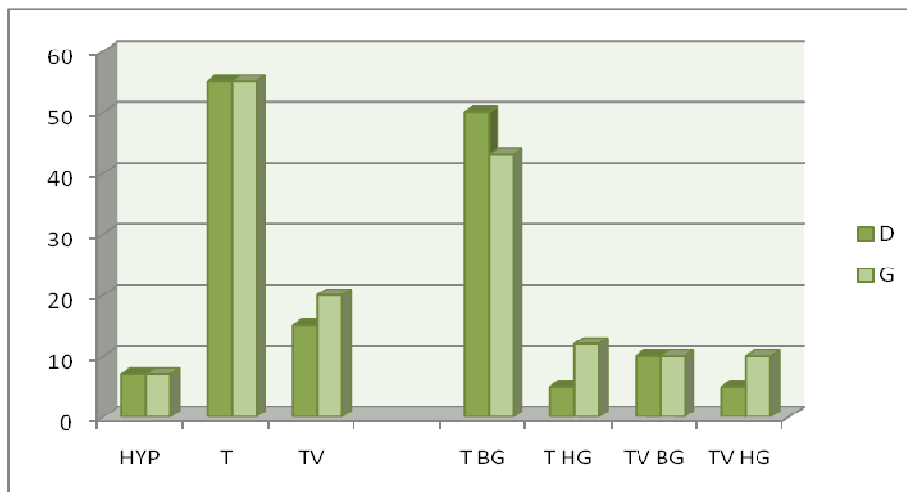


Figure 21 : Répartition des 159 polypes coliques selon leur localisation et leur type histologique. D=droite ; G=gauche ; HYP=polype hyperplasique ; T=adénome tubuleux ; TV=adénome tubulo-villeux ; BG=bas grade ; HG=haut grade.

Le pourcentage de cellules « pathologiques » a été établi sur coupe à congélation et correspond au nombre de cellules épithéliales hyperplasiques ou adénomateuses par rapport au nombre total de cellules (= cellules épithéliales hyperplasiques ou adénomateuses + cellules épithéliales normales + cellules non épithéliales). Ce chiffre varie de 30 à 80%, a parfois nécessiter une macrodissection de l'échantillon congelé pour augmenter la cellularité (**Figure 22**).

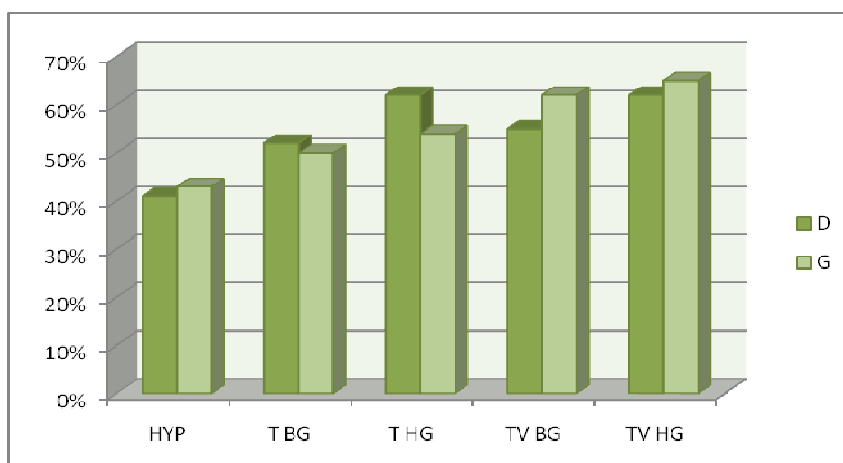


Figure 22 : Moyenne des pourcentages de cellules pathologiques des 159 polypes coliques selon leur localisation et leur type histologique. D=droite ; G=gauche ; HYP=polype hyperplasique ; T=adénome tubuleux ; TV=adénome tubulo-villeux ; BG=bas grade ; HG=haut grade.

2. Sous-typage moléculaire des polypes coliques non-MIN

La recherche de sous-types moléculaires au sein d'une cohorte de 159 polypes coliques non MSI a été faite par allélotypage, analyse du statut de méthylation du promoteur du gène *MGMT* et du statut mutationnel du gène *KRAS*. L'allélotypage a ciblé les régions les plus fréquemment réarrangées dans les CC, dont les régions contenant les gènes *APC* et *TP53*. Le panel utilisé en routine hospitalière permet de détecter au moins une altération dans 90 à 95% des tumeurs coliques, dont l'échantillon analysé renferme plus de 30% de cellules tumorales après évaluation par le pathologiste.

Ces analyses montrent que 62% des adénomes (n=99) de tous grades et types histologiques présentent des déséquilibres alléliques (AI) aux loci analysés et, que 38% (n=60) ne comportent aucun AI pour ces loci et sont considérés comme non-CIN non-MIN. Une méthylation de *MGMT* et une mutation de *KRAS* sont trouvées dans respectivement 39% et 32% des polypes, et dans respectivement 38% et 26% des polypes non-CIN non-MIN.

Aucune de ces altérations analysées n'est détectée dans 14% des polypes (n=23), avec une confirmation histologique qu'il s'agit bien de polypes hyperplasiques ou d'adénomes dont les échantillons comportent plus de 30% de cellules pathologiques. Les seuils de sensibilité des techniques employées dans cette étude varient de 20%, 5% à 1% pour l'allélotypage, la méthylation de *MGMT* et la mutation de *KRAS* respectivement.

La fréquence globale d'AI est significativement corrélée aux adénomes tubulo-villeux (p=0,005) et aux adénomes de haut grade (p=0,0001), également détectés en utilisant un panel restreint de microsatellites ciblant des régions des chromosomes 1, 6 et 9.

La clusterisation hiérarchisée des données d'allélotypage a classé la cohorte de 159 polypes en 4 sous-types selon la fréquence globale d'altération (non-CIN non-MIN, CIN-Very Low, CIN-Low et CIN-High).

3. Identification de sous-types moléculaires par association d'altérations

3.1. Polypes avec AI

Dans les 99 adénomes ayant des AI, l'étude de l'incidence des altérations des loci ciblant *APC* (D5S346) et *TP53* montre que l'altération la plus fréquente est celle du locus associé à *APC*, altéré dans près de la moitié des polypes (46%, 35/74 informatifs) avec 17% d'entre eux présentant l'altération d'*APC* seul (16/74). Le locus ciblant *TP53* est altéré dans 15% des cas (11/74 cas informatifs) et 4% présentent l'altération de *TP53* seule. Comme présenté dans la publication, la mutation de *KRAS* est retrouvée dans 35% des cas (35/99), avec 3% des cas présentant la mutation seule, et la méthylation du promoteur du gène *MGMT* est observée dans 40% des cas (39/99), avec 7% des cas n'ayant que la méthylation de *MGMT*.

Par conséquent, l'altération majeure cible le locus d'*APC*, suivie de la mutation de *KRAS*, de la méthylation de *MGMT*, puis de l'altération de *TP53*, peu fréquente (**Tableau 10**).

n=74	1	2	3	4
	<i>APC</i>	<i>MGMT</i>	<i>KRAS</i>	<i>TP53</i>
Total	54%	40%	35%	16%
Seule	17%	7%	3%	4%

Tableau 10 : Altérations ciblant *APC*, *TP53* (allélotypage), *KRAS* (mutation), *MGMT* (méthylation) dans la cohorte des 99 adénomes avec AI. Total : incidence des altérations indépendamment du statut des autres altérations ; Seule : altération seule prise en compte mais les 4 altérations sont informatives simultanément (74/99 adénomes, 74%).

Les associations d'altérations sont multiples (**Figure 23**). Si on observe les combinaisons au sein d'un sous-groupe altéré pour *APC*, l'association avec la méthylation de *MGMT* est la plus fréquente (37%), suivie par l'association avec la mutation de *KRAS* (26%) puis avec l'altération pour *TP53* (14%). Mais quand *TP53* est altéré, il existe une co-altération pour *APC* dans presque la moitié des cas (45%).

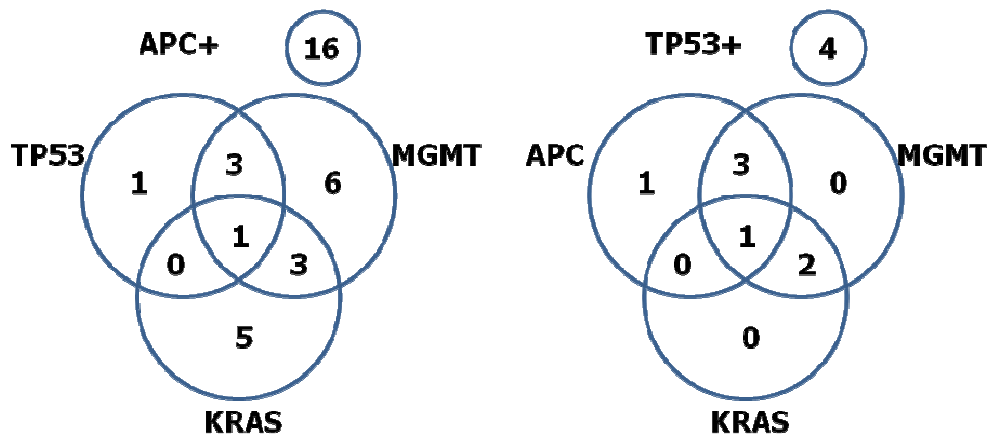


Figure 23 : Répartition des cas d'altérations ciblant *APC*, *TP53*, *KRAS* et *MGMT* pour les 74 cas informatifs simultanément, au sein d'un sous-groupe altéré pour *APC* (*APC+* $n=35$) et d'un sous-groupe altéré pour *TP53* (*TP53+* $n=11$).

Ainsi, dans le groupe des polypes présentant des AI, cette analyse montre une hétérogénéité d'anomalies importantes, où *APC* est fréquemment associé à *MGMT* seul ou à *KRAS* seul, contrairement à *TP53* qui ne semble pas être associé seul à *MGMT* ou à *KRAS*.

3.2. Polypes sans AI

Parmi les 60 polypes ne présentant pas d'AI aux loci examinés, trois sous-groupes peuvent essentiellement être distingués : un groupe caractérisé par une mutation de *KRAS* (23%), un groupe comportant le gène *MGMT* méthylé (35%) et un groupe sans altérations (38%) dans nos conditions (**Figure 24**).

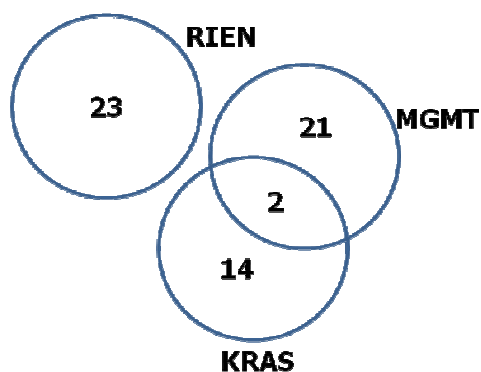


Figure 24 : Caractéristiques moléculaires du groupe des 60 polypes sans AI.

Ce groupe sans AI est aussi hétérogène. Mais contrairement au groupe avec AI, l'association mutation de *KRAS* / méthylation de *MGMT* est rare (3% versus 20%), pouvant suggérer que ces altérations ont un rôle différent dans l'initiation de la transformation adénomateuse.

Publication 1

**Histological characteristics of non-MSI colon adenomas correlate
with distinct molecular patterns**

Neuville A, Nicolet C, Meyer N, Schneider A, Legrain M, Brigand C, Duclos
B, Bachellier P, Oudet P, Bellocq JP, Kedinger M, Gaub MP, Guenot D.

Human Pathology

Sous presse



Original contribution

Histologic characteristics of non-microsatellite-instable colon adenomas correlate with distinct molecular patterns[☆]

Agnès Neuville MD^{a,b,c,*}, Céline Nicolet PhD^a, Nicolas Meyer PhD^d,
 Anne Schneider PharmD^e, Michèle Legrain PhD^e, Cécile Brigand MD^f,
 Bernard Duclos MD^f, Philippe Bachellier MD^f, Pierre Oudet MD^e,
 Jean-Pierre Bellocq MD^b, Michèle Kedingler PhD^g,
 Marie-Pierre Gaub PhD^{a,e,1}, Dominique Guenet MD^{a,1}

^aEA 4438, Université de Strasbourg, F-67200 Bâtiment Inserm, France

^bDépartement de Pathologie, CHRU, 67098 Strasbourg Cedex, France

^cCentre de Ressources Biologiques, CHRU, 67098 Strasbourg Cedex, France

^dDépartement de Santé Publique, CHRU, 67091 Strasbourg Cedex, France

^eLaboratoire de Biochimie et de Biologie Moléculaire, CHRU, 67098 Strasbourg Cedex, France

^fPôle de Pathologies Digestives, Hépatiques et de Transplantation, CHRU, 67098 Strasbourg Cedex, France

^gInserm, UMR-S682, Université de Strasbourg, F-67200 Strasbourg, France

Received 25 February 2010; revised 17 June 2010; accepted 21 July 2010

Keywords:

Colon carcinogenesis;
 Tumor initiation;
 Chromosome instability;
 Methylation;
K-RAS gene mutation;
 Allelotyping

Summary Colon carcinogenesis encompasses the stepwise accumulation of genomic aberrations correlated with the transition of aberrant crypt-adenoma-carcinoma. Recent data have revealed that, in addition to the microsatellite-instable phenotype, the chromosome instability pathway, representing four fifth of the colon carcinoma, could be involved in heterogeneous molecular alterations. Our project was aimed at determining the existence of distinct molecular subtypes in 159 non-microsatellite-instable colon polyps and their correlation with histology and dysplasia, using allelotyping, *MGMT* promoter gene methylation status, and *K-RAS* mutation analyses. Allelic imbalance, *MGMT* methylation, and *K-RAS* mutations arise in 62%, 39%, and 32% of polyps, respectively. Only 14% of polyps had no alterations. A 2-way hierarchical clustering analysis of the allelic imbalances identified subgroups of polyps according to their allelic imbalance frequency and distribution. Not only tubulovillous adenoma but also high-grade adenomas were correlated with high global allelic imbalance frequency ($P = .005$ and $P = .003$), with allelic imbalance at microsatellites targeting chromosomes 1, 6, and 9. In conclusion, the data presented in this study show that a large heterogeneity exists in the molecular patterns of alterations in precancerous colon lesions, favoring different modes of tumor initiation. Therefore, molecular alterations correlated with tubulovillous-type and high-grade dysplasia could represent targets identifying predictive factors of progression.

© 2011 Elsevier Inc. All rights reserved.

[☆] No conflict of interest.

* Corresponding author. Département de Pathologie, Hôpital de Hautepierre, CHRU de Strasbourg, 67098 Strasbourg Cedex, France.
 E-mail address: agnes.mechine@chru-strasbourg.fr (A. Neuville).

¹ Contributed equally to this work.

1. Introduction

The prevailing model in colon carcinogenesis implies mutations and/or loss of heterozygosity of tumor suppressor genes, such as TP53 and adenomatous polyposis coli (*APC*), and mutations of oncogenes such as *K-RAS* [1]. These alterations lead to 2 classic pathways, one with chromosome instability associated with loss of heterozygosity or gene copy number gain [2], defined as the chromosomal instability (CIN) phenotype, and the other defined as the microsatellite instability (MSI) phenotype, in which DNA repair gene mutations lead to MSI and a strong index of mutations in the genome [3]. The idea that inactivation of the tumor suppressor gene *APC* serves to initiate colorectal cancers may also have important tumorigenic and pathogenic effects. The early evolution of colorectal neoplasia could occur by mechanisms other than inactivation of *APC* or related alterations that would disturb the Wnt signaling pathway [4]. Oncogenic mutations involving both *BRAF* and *K-RAS* are likely alternative initiating steps that could synergize with DNA methylation and occur within the context of serrated polyps [5,6]. Chromosome instability resulting from gene mutations involved in chromosome stability could precede *APC* mutations and be the first event initiating malignancy [7].

Among the other mechanisms that can occur in tumor initiation, alterations in DNA methylation have been reported as an early event in carcinogenesis. Inappropriate promoter hypermethylation, associated with gene silencing [8], defines the CpG island methylator phenotype, which takes part in the microsatellite and chromosome instability phenotypes [9].

Recent studies indicate that methylation abnormalities are common in precancerous colorectal lesions [10,11], in particular, methylation of the *MGMT* promoter gene [12,13]. Moreover, *MGMT* is considered a mutator gene because loss of *MGMT* expression could lead to an accumulation of mutations [14,15].

Genomic and epigenetic alterations can arise before the appearance of mutations in genes known to be at the origin of colon carcinogenesis, thus generating a growing interest in the identification of new alteration pathways responsible for the initiation of colon cancer [16,17].

Among the criteria used to stratify the risk of advanced colorectal adenoma recurrence, the number and size of adenomas are the most useful for surveillance colonoscopy [18]. However, several studies have reported the impact in risk stratification of histologic features, such as the tubulovillous or villous component and high-grade dysplasia [19-21], but to our knowledge, genomic abnormalities have not been associated with the recurrence of colorectal adenomas.

Previously, we reported that allelotyping CIN colon adenocarcinomas and their synchronous metastases revealed 3 subtypes without correlation to the allelic

imbalance (AI) frequency and the evolutionary stage [22]. These results suggested that a minimal set of alterations would be sufficient to promote tumor progression and metastasis without further alterations. These data also suggested the existence of different pathways of carcinogenesis depending on the frequency and the type of alterations. Therefore, to answer whether such molecular subtypes are observed early in carcinogenesis and to identify the molecular basis of colon cancer, a study that focused on the precancerous stage with a large number of non-MSI colon polyps was conducted. Using allelotyping, *MGMT* promoter gene methylation, and *K-RAS* mutation (K-RASm) analyses, we aimed to show the potential links between the types of molecular alterations and the histologic features of the polyps.

2. Materials and methods

2.1. Patients and tumor samples

Frozen and paraffin-embedded samples of 249 colorectal polyps, with macrodissected paired normal tissue and/or blood, were collected from 1996 to 2006 at the Biological Resources Center at the University Hospital of Strasbourg, using protocols approved by the institutional review board. Once the pathological diagnosis was established, the remaining tissue was frozen (polyps, >0.5 cm). Most of the collected polyps were obtained from synchronous colon carcinoma resections (132 of 159 polyps, resected from 101 patients). Forty polyps of MSI phenotype were excluded by allelotyping using the Bethesda microsatellite panel [23]. Fifty polyps of rectal location were also excluded. One hundred fifty-nine non-MSI colon polyps (115 males and 44 females; mean age, 68 years) from 120 patients (85 males and 35 females) were then available for further allelotyping, *MGMT* promoter gene methylation analysis, and K-RASm detection. Of the 120 patients, 93 had 1 polyp (5 resected by endoscopy and 88 during colon carcinoma resection), whereas the remaining 27 had between 2 and 6 polyps, which were obtained during carcinoma resection.

2.2. Histopathological analyses

For all cases, the histopathological type of the polyps (tubular, tubulovillous or villous adenomas, and hyperplastic polyp), sessile or conventional serrated adenomas, the degree of dysplasia (low and high grade), and the evaluation of the epithelial cell number were performed on frozen and on paraffin-embedded sections and confirmed by 2 independent experienced pathologists as described previously [24]. Table 1 summarizes the pathological features of the polyps included in this study. Twenty-nine percent of the polyps had between 50% and 30% of hyperplastic or

Table 1 Anatomic characteristics of the 159 colon polyps

	HP, n = 14	LG T, n = 93	HG T, n = 17	LG TV, n = 20	HG TV, n = 15	All polyps
Age (y)						
<60	7 (50%)	14 (15%)	3 (12%)	3 (15%)	3 (20%)	30 (19%)
≥60	7 (50%)	79 (85%)	14 (88%)	17 (85%)	12 (80%)	129 (81%)
Sex						
Men	14 (100%)	64 (69%)	12 (71%)	13 (65%)	12 (80%)	115 ^a (72%)
Women	0	29 (31%)	5 (29%)	7 (35%)	3 (20%)	44 ^a (28%)
Site						
Proximal	7 (50%)	50 (54%)	5 (29%)	10 (50%)	5 (33%)	77 (48%)
Distal	7 (50%)	43 (46%)	12 (71%)	10 (50%)	10 (67%)	82 (52%)
Size (cm)						
<1	6 (43%)	42 (45%)	2 (12%)	2 (10%)	0	49 (31%)
≥1	8 (57%)	51 (55%)	15 (88%)	18 (90%)	15 (100%)	110 (69%)
Epithelial cell count (%)						
<50	8 (57%)	30 (32%)	3 (18%)	4 (20%)	1 (7%)	46 (29%)
≥50	6 (43%)	63 (68%)	14 (82%)	16 (80%)	14 (93%)	113 (71%)
Invasive synchronous carcinoma						
Yes	12 (86%)	81 (87%)	13 (76%)	14 (70%)	12 (80%)	132 (83%)
No	2 (14%)	12 (13%)	4 (24%)	6 (30%)	3 (20%)	27 (17%)

Abbreviations: HP, hyperplastic polyp; LG T, low-grade tubular adenoma; HG T, high-grade tubular adenoma; LG TV, low-grade tubulovillous adenoma; HG TV, high-grade tubulovillous adenoma.

^a These numbers were issued from 85 male and 35 female patients.

adenomatous cells, and 71% of the polyps had greater than 50% of hyperplastic or adenomatous cells.

2.3. DNA extraction and allelotyping

DNA was extracted from paired blood, frozen polyp, and normal mucosa using the QIAamp DNA purification kit (Qiagen, Hilden, Germany). Extracted DNA was amplified by fluorescent polymerase chain reaction (PCR) as described previously [25] (see [Supporting information](#)). To detect events with potential biological impact, we used 33 microsatellites targeting 18 chromosomes corresponding to the most frequently rearranged loci found in colon cancers [26,27].

2.4. Analysis of the CpG island methylation status of the *MGMT* gene

The methylation pattern of the CpG islands in the *MGMT* gene was determined using methylation-specific PCR after the DNA (100 ng) was modified with sodium bisulfite and purified using the Epitect Bisulfite kit (Qiagen; see [Supporting information](#)).

2.5. Detection of K-RASm

To analyze K-RASm, we used the peptide nucleic acid clamping method (see [Supporting information](#)). PCR reactions were performed in the presence or absence of peptide nucleic acid using the Light cycler Fast Start DNA Master Hyb Probe kit (Roche Diagnostics). Fluorescence data were analyzed using light cycler software.

2.6. Statistical analysis

Data were analyzed using the χ^2 test and Fisher exact test.

The *P* value was considered significant when below .05. For clustering analysis, data were coded in binary form using “1” for AI and “-1” for a normal informative locus. Homozygotes were treated as missing data. Data were then clustered using a 2-way hierarchical clustering method, with uncentered correlation as similarity metrics for both genes and subjects value vectors. The average linkage was chosen as the aggregation method. Computations were run with Gene Cluster 3.0. Cluster trees were produced using Java TreeView 1.0.4 (Eisen’s Softwares, Berkeley, CA, USA).

3. Results

To identify molecular subtypes and new carcinogenesis pathways in CIN colon polyps, MSI and rectal polyps were excluded in this study.

3.1. Pathological findings

Among the 159 non-MSI consecutive colon polyps eligible for this study, 110 (69%) were tubular adenomas, and 35 (22%) were tubulovillous adenomas ([Table 1](#)). High grade represented 20% and 43% in tubular and tubulovillous adenomas, respectively, and 14 (9%) polyps were hyperplastic ([Fig. 1](#)). Villous adenomas, sessile serrated polyps, and traditional serrated adenoma were not identified. Overall, the polyp incidence was higher in men older than

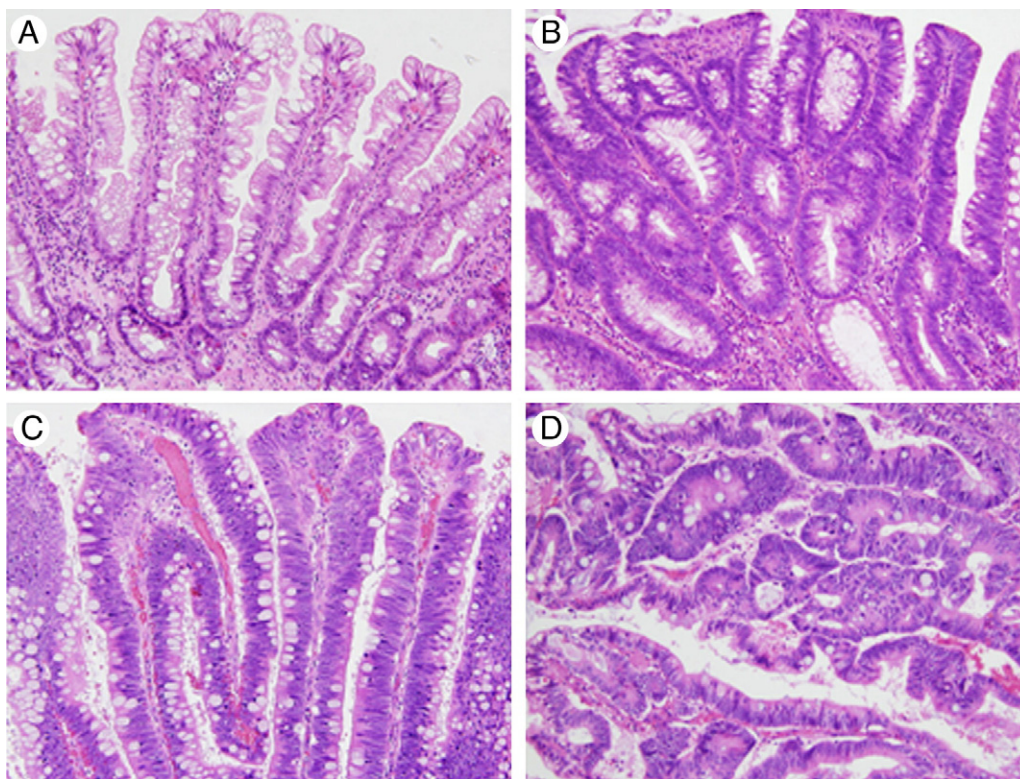


Fig. 1 Histologic examples of polyps. A, Typical morphology of hyperplastic polyp with serration at the surface and nonbranching crypts lined with proliferative cells at the base. B and C, Conventional adenoma with tubular and villous architecture, respectively. D, Nuclear atypia and architectural disorganization corresponding to high-grade adenoma (hematoxylin-eosin, original magnification $\times 20$).

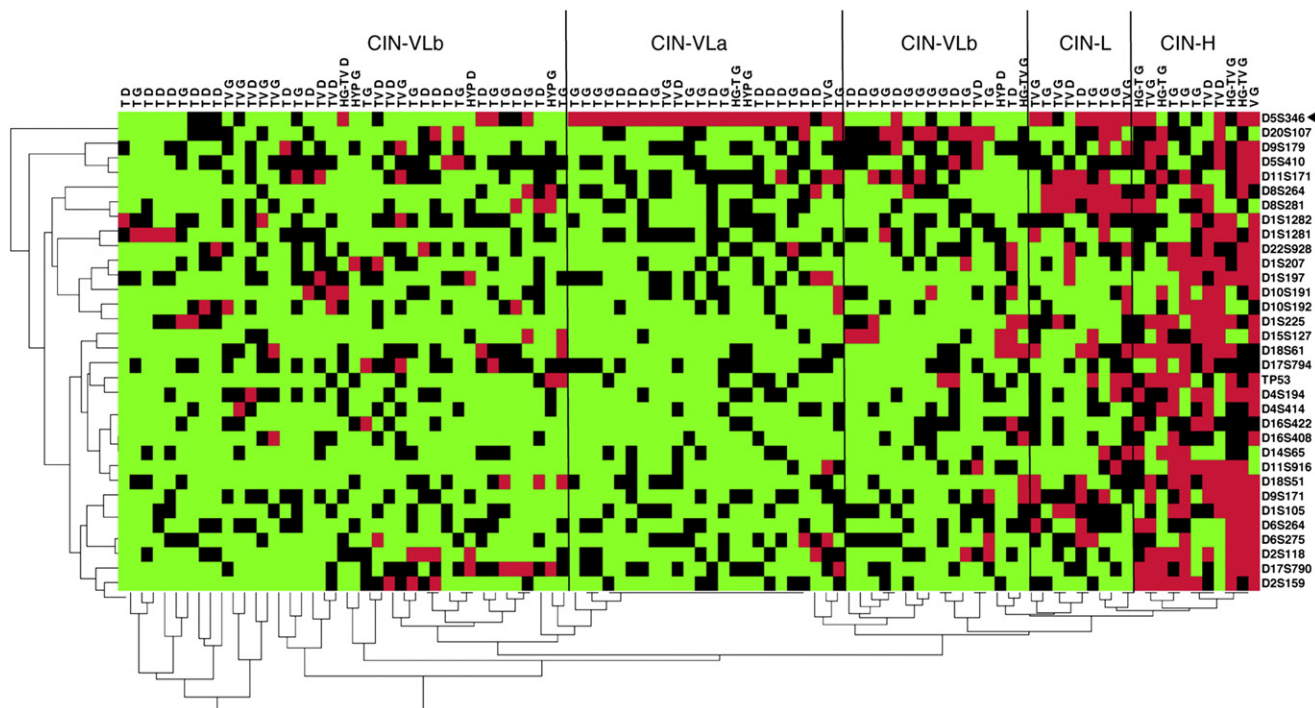


Fig. 2 Two-way hierarchical clustering of allelotyping 99 colon polyps with at least 1 AI. Red rectangles indicate AI; green rectangles, normal locus; and black rectangles, locus noninformative. Horizontal lines correspond to the microsatellites; vertical columns correspond to patients. CIN-H, CIN-L, CIN-VLa, and CIN-VLb correspond to patients with CIN with high (H), low (L), or very low (VL) frequency of alteration. Arrow head (\blacktriangleleft) indicates the *APC* locus (*D5S346*).

60 years, and no hyperplastic polyps were detected in women. High-grade adenomas of tubular or tubulovillous type were predominant in the left part of the colon, whereas hyperplastic polyps and low-grade adenomas were almost equally distributed between the right and left colon. Most polyps (83%) were synchronous with an invasive carcinoma. Among the 27 polyps removed without an invasive carcinoma, 21 (78%) were 1 cm or larger in size. Larger polyps were correlated with tubulovillous and high-grade groups ($P = .003$).

3.2. Allelotyping

Overall, 99 polyps (62%) showed at least 1 AI at an informative locus (Fig. 2). Polyps without AI ($n = 60$, or 38%) were observed in all histologic groups and grades but more frequently in hyperplastic polyps and low-grade tubular adenomas (Table 2). Among polyps larger than 1 cm, those without AI were generally right sided (34 versus 26), whereas polyps with AI were located mainly in left colon (56 versus 43), but this was not considered significant. Although the global AI frequency in the different groups of polyps showed variations (Fig. 3A), the mean global AI frequency was 3% for hyperplastic polyps (0%-15% AI), 6% for low-grade tubular (0%-52% AI), 18% for high-grade tubular (0%-56% AI), 13% for low-grade tubulovillous (0%-58% AI), and 30% for high-grade tubulovillous adenomas (0%-86% AI). High global AI frequency (46%-100% AI) was statistically associated with the tubulovillous-type ($P = .005$) and high-grade dysplasia ($P = .003$; Fig. 3B). In patients with multiple polyps (27 patients), the genomic alterations were different between the polyps for 24 of 27 patients. Interestingly, only 25% of the polyps showed an alteration at *D5S346* (microsatellite informative for *APC*

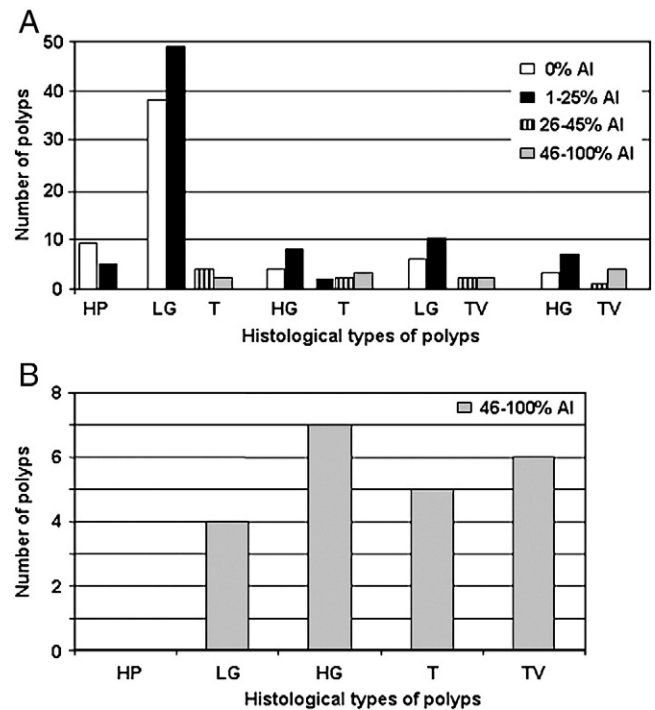


Fig. 3 Distribution of global AI frequency among the polyps. A, Distribution within the 159 polyps classified according to the different histopathological types. B, Distribution within the polyps with a high global AI frequency (46%-100% AI) classified according to the different histopathological types. HP indicates hyperplastic polyp; LG T, low-grade tubular adenoma; HG T, high-grade tubular adenoma; LG TV, low-grade tubulovillous adenoma; HG TV, high-grade tubulovillous adenoma. The global AI frequency corresponds to the number of altered loci among the microsatellites analyzed per patient.

locus). Alteration at locus *D2S159* was observed mostly in women, and alterations at loci *D14S65*, *D17S790*, and *TP53* were significantly associated to distal location (Table 3). In

Table 2 Molecular characteristics for the 159 colon polyps

	HP, n = 14	LG T, n = 93	HG T, n = 17	LG TV, n = 20	HG TV, n = 15	All polyps	Size, <1 cm	Size, ≥1 cm
AI								
Yes	5 (36%)	55 (59%)	13 (76%)	14 (70%)	12 (80%)	99 (62%)	25 (51%)	74 (67%)
No	9 (64%)	38 (41%)	4 (24%)	6 (30%)	3 (20%)	60 (38%)	24 (49%)	36 (33%)
AI only	2 (14%)	31 (33%)	6 (35%)	4 (20%)	2 (3%)	45 (28%)		
<i>MGMT</i> gene methylation								
Yes	5 (36%)	33 (36%)	8 (47%)	10 (50%)	6 (40%)	62 (39%)	16 (33%)	46 (42%)
No	9 (64%)	60 (64%)	9 (53%)	10 (50%)	9 (60%)	97 (61%)	33 (67%)	64 (58%)
Methy only	2 (14%)	15 (16%)	1 (6%)	2 (10%)	1 (7%)	21 (13%)		
K-RASm								
Yes	2 (14%)	22 (24%)	9 (53%)	8 (40%)	10 (67%)	51 (32%)	7 (14%)	44 (40%)
No	12 (86%)	71 (76%)	8 (47%)	12 (60%)	5 (33%)	108 (68%)	42 (86%)	66 (60%)
K-RASm only	0	9 (10%)	2 (12%)	2 (10%)	1 (7%)	14 (9%)		
AI at <i>D5S346</i>								
Yes	1 (8%)	23 (30%)	6 (50%)	4 (22%)	6 (46%)	40 (30%)	15 (36%)	25 (27%)
No	12 (92%)	55 (70%)	6 (50%)	14 (78%)	7 (54%)	94 (70%)	26 (64%)	68 (73%)

Abbreviations: HP, hyperplastic polyp; LG T, low-grade tubular adenoma; HG T, high-grade tubular adenoma; LG TV, low-grade tubulovillous adenoma; HG TV, high-grade tubulovillous adenoma; K-RASm, K-RAS mutation.

NOTE. The locus *D5S346* was informative for 13 polyps and indicates the *APC* gene status.

addition, alterations at *D1S225*, *D4S414*, and *D10S191* were strongly associated to be histologically tubulovillous, and alterations at *D2S138*, *D11S916*, and *D17S790* were associated with high-grade dysplasia (Table 3), whereas alterations at loci *D1S305*, *D4S394*, *D6S264*, *D6S275*, and *D9S179* showed the strongest correlation for both tubulovillous and high-grade features (Table 3).

3.3. MGMT gene methylation

To better assess the molecular status of the polyps, promoter methylation analysis of the *MGMT* gene was performed because it has been shown previously to be a potential mutator gene [14]. *MGMT* was methylated in 62

(39%) of the 159 polyps (Table 2) and was associated with alterations at loci *D3S1282* ($P = .03$) and *D9S179* ($P = .02$; Table 3). Among the 99 polyps with at least 1 AI, 39 (39%) were also found to be *MGMT* methylated. The distribution of *MGMT* methylation was homogeneous among the different histologic groups of polyps (Table 2).

3.4. K-RASms

Because the *K-RAS* gene has been suggested to initiate the adenoma-carcinoma transition, mutations at exon 1 (codons 12 and 13) were analyzed. Among the 159 polyps, 51 (32%) had K-RASm with significant correlations with the size (44 K-RASm of 110 polyps >1 cm versus 7 K-RASm of 49

Table 3 Correlations of alterations at the 33 microsatellites with clinicopathological and molecular features

MS	Chrom location	AI frequency (%)	Female vs male	Distal vs proximal	TV ^b vs T	HG ^b vs LG	MGMT methyl	K-RAS mutation
D1S197 ^a	1p33	9.6	NS	NS	$P = .005$	NS	NS	$P = .008$
D1S207 ^a	1p31.1	9.5	$P = .027$	NS	$P = .045$	NS	NS	NS
D1S225 ^a	1q42.2	9.9	NS	NS	$P = .022$	NS	NS	NS
D1S305 ^a	1q21.3	7.3	NS	NS	$P = .001$	$P = .015$	NS	NS
D2S138	2q31.1	12.3	NS	NS	NS	$P = .033$	NS	NS
D2S159	2q36.3	8.6	$P = .005$	NS	NS	NS	NS	NS
D3S1282	3q26.2	8.8	NS	NS	$P = .013$	NS	$P = .03$	$P = .007$
D3S1283	3p22-24	10.8	NS	NS	NS	NS	NS	NS
D4S394	4p16.1	7.4	NS	NS	$P = .008$	$P = .001$	NS	NS
D4S414	4q22.1	4.3	NS	NS	$P = .001$	NS	NS	NS
D5S346	5q22.2	30	NS	NS	NS	NS	NS	NS
D5S430	5p13.1	8.9	NS	NS	NS	NS	NS	NS
D6S264 ^a	6q27	6.8	NS	NS	$P = .01$	$P = .002$	NS	NS
D6S275 ^a	6q16.1	9	NS	NS	$P = .031$	$P = .005$	NS	NS
D8S264	8p23.3	12.3	NS	NS	NS	NS	NS	NS
D8S283	8p12	10.3	NS	NS	NS	NS	NS	NS
D9S171	9p21	10.7	NS	NS	NS	NS	NS	NS
D9S179 ^a	9q34.13	11.7	NS	NS	$P = .001$	$P = .002$	$P = .02$	NS
D10S191	10p13	9.8	NS	NS	$P = .034$	NS	NS	NS
D10S192	10q24.31	8	NS	NS	NS	NS	NS	NS
D11S916	11q13.4	7.6	NS	NS	NS	$P = .04$	NS	NS
D13S173	13q33.3	16.7	NS	NS	NS	NS	NS	NS
D14S65	14q32.2	4.7	NS	$P = .025$	NS	NS	NS	NS
D15S127	15q26.1	9.1	NS	NS	NS	NS	NS	NS
D16S408	16q13	3.3	NS	NS	$P = .001$	NS	NS	$P = .02$
D16S422	16q23.3	4.1	NS	NS	NS	NS	NS	NS
D17S790	17q22	11.4	NS	$P = .027$	NS	$P = .006$	NS	NS
D17S794	17q23.2	6.6	NS	NS	NS	$P = .045$	NS	NS
TP53	17p13.1	9.7	NS	$P = .04$	NS	$P = .044$	NS	NS
D18S53	18p11.21	11.7	NS	NS	$P = .015$	$P = .039$	NS	NS
D18S61	18q22.2	13.7	NS	NS	$P = .03$	$P = .034$	NS	NS
D20S107	20q12	13.8	NS	NS	NS	NS	NS	NS
D22S928	22q13.31	10.2	NS	NS	NS	NS	NS	NS

Abbreviations: MS, microsatellites; Chrom Location, chromosomal location (according to Ensembl 52); NS, not significant; TV, tubulovillous adenoma; T, tubular adenoma; HG, high-grade adenoma; LG, low-grade adenoma; MGMT methyl, *MGMT* gene methylation.

NOTE. The locus frequency corresponds to the number of altered loci among the heterozygous MS analyzed. Proximal and distal correspond to the anatomic location of the polyps. All the P values indicated are significant ($P < .05$).

^a Loci used for the hierarchical clustering as shown in Fig. 5.

^b Significant correlation in favor of TV or HG histologic types.

Table 4 Chromosomal instability, *MGMT* methylation, and K-RASm phenotypes

	A + B	A + C	B + C	A + B + C	None
HP	1 (7%)	0	0	2 (10%)	7 (30%)
LG T	12 (13%)	7 (47%)	1 (50%)	5 (25%)	13 (56.5%)
HG T	1 (6%)	1 (6.5%)	1 (50%)	5 (25%)	0
LG TV	4 (20%)	2 (13%)	0	4 (20%)	2 (8.5%)
HG TV	1 (7%)	5 (33%)	0	4 (20%)	1 (4%)
All polyps	19/159 (12%)	15/159 (9.5%)	2/159 (1.5%)	20/159 (13%)	23/159 (14.5%)

Abbreviations: HP, hyperplastic polyp; LG T, low-grade tubular adenoma; HG T, high-grade tubular adenoma; LG TV, low-grade tubulovillous adenoma; HG TV, high-grade tubulovillous adenoma. A + B, polyps with AI and *MGMT* gene methylation; A + C, polyps with AI and *K-RAS* mutation; B + C, polyps with *MGMT* gene methylation and *K-RAS* mutation; A + B + C, polyps with AI, *MGMT* gene methylation, and *K-RAS* mutation.

NOTE. Number of polyps is indicated in each clan.

polyps ≤ 1 cm, $P = .05$; Table 2), tubulovillous-type (18 K-RASm of 35 tubulovillous adenomas versus 33 K-RASm of 124 hyperplastic polyps and tubulous adenomas, $P = .01$; Table 2), and high-grade dysplasia (19 K-RASm of 32 high-grade adenomas versus 30 K-RASm of 113 low-grade adenomas, $P = .001$; Table 2). In addition, K-RASm was associated with alterations at *DIS197* ($P = .008$), *D3S1282* ($P = .007$), and *D16S408* ($P = .02$; Table 3).

3.5. Allelotyping, *MGMT* methylation, and K-RASm

We analyzed the different patterns of associations between the existence of AI, *MGMT* methylation, and K-RASm (Tables 2 and 4). Markedly, 50% of the polyps had only 1 of the 3 types of alterations (Table 2). The largest group corresponded to polyps with AI only (28%) (Table 2), integrating a subgroup with alterations at the *APC* locus (14/45, or 31%). Among the 60 polyps with no AI at the loci analyzed, 21 (13%) polyps had only *MGMT* methylation and 14 (9%) had only K-RASm (Table 2). When considering combinations of alterations (Table 4), 2 polyps had both *MGMT* methylation and K-RASm (1.5%), 15 (9.5%) had AI and K-RASms, and 20 (12.5%) had all 3 alterations. Finally, 23 polyps (14%) had no detectable alteration and were mostly represented as hyperplastic polyps (7/23, or 30%) and low-grade adenomas (13/23, or 57%), without a predominance of location (Table 4).

3.6. Clustering analyses

3.6.1. Allelotyping data

A 2-way hierarchical clustering analysis of the allelotyping data suggested the existence of 3 subgroups, according to

the AI frequency and loci distribution (Fig. 2). Most of the polyps (79/159, or 49.5%) had a global AI frequency of less than 25%, corresponding to a CIN-very low (VL) subgroup, which could be divided into 2 parts: a CIN-VLa part of 24 polyps with a mean AI frequency of 7% (3%-20%), all polyps having AI at informative locus *APC* and 11 having simultaneous K-RASm or *MGMT* methylation; and a CIN-VLb part of 55 polyps with a mean AI frequency of 8% (3%-21%), with a heterozygous distribution of AI among the different informative loci, and only 1 of these 55 polyps had an altered locus *APC*. Among these 55 polyps, only 8 had K-RASm (14.5%), 8 had only *MGMT* methylation (14.5%), and 14 had both K-RASm and *MGMT* methylation (25.5%). A second group, CIN-low (L), including 9 polyps (5.5%), had a global AI frequency between 27% and 43% (mean, 30%), with alterations mainly on chromosomes 18p, 8p, and 5p. Finally, a CIN-high (H) subgroup of 11 polyps (7%) showed strong chromosome instability, with a mean AI frequency higher than 45% (52%-86%; mean, 66%).

Proximal and distal locations were distributed equally in the CIN-VL subgroup (38 versus 41), whereas a distal location was predominant in the CIN-L (6 versus 3) and CIN-H groups (9 versus 2).

3.6.2. *MGMT* methylation, K-RASm, and *APC* locus

Clustering analysis was performed to analyze the distribution of alterations at the *APC* locus, AI, *MGMT* methylation, and K-RASm in the 99 polyps with at least 1 AI. Interestingly, the analysis of K-RASm and altered *APC* locus showed an exclusive alteration in 43 (80%) of the 54 polyps (Fig. 4). A similar exclusive alteration was observed for *MGMT* methylation and AI at the *APC* locus, with an

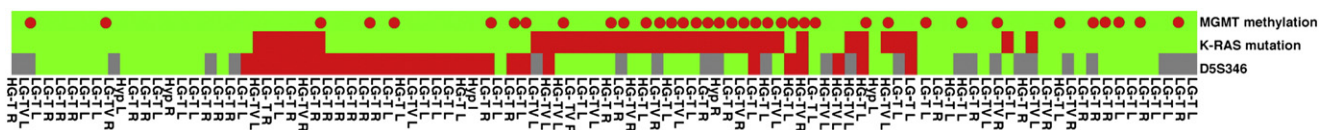


Fig. 4 *MGMT* methylation, K-RASm, and *D5S346* locus among 99 AI polyps. A 2-way hierarchical clustering was performed for the 99 polyps with at least 1 AI at an informative locus. Red rectangles indicate AI; green rectangles, normal locus; black rectangles, locus noninformative. Horizontal columns correspond to *MGMT* methylation, K-RASm, and AI for *D5S346* (informative for *APC*). Vertical columns correspond to the polyps. L indicates left adenoma; R, right adenoma.

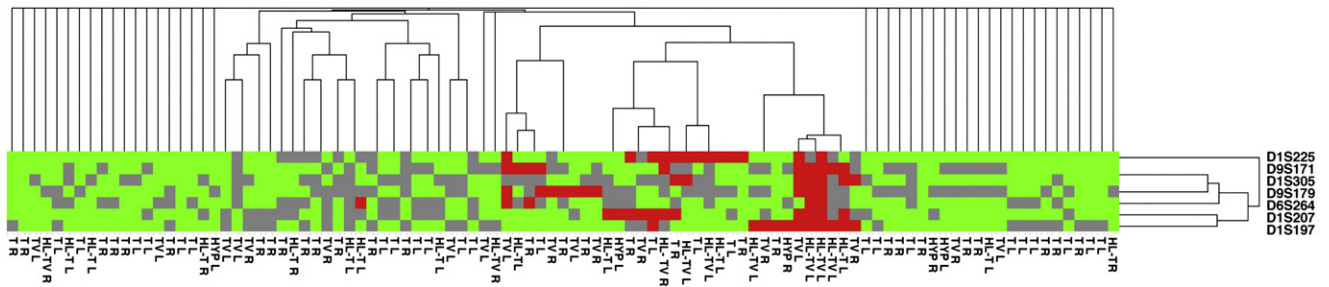


Fig. 5 Two-way hierarchical clustering of allelotyping 159 colon polyps with microsatellites targeting chromosomes 1, 6, and 9. Red rectangles indicate AI; green rectangles, normal locus; gray rectangles, locus noninformative. Vertical columns correspond to the polyps; horizontal lines correspond to the microsatellites *DIS197*, *DIS207*, *DIS225*, *DIS305*, *D6S264*, *D6S275*, and *D9S179*. HP indicates hyperplastic polyp; LG T, low-grade tubular adenoma; HG T, high-grade tubular adenoma; LG TV, low-grade tubulovillous adenoma; HG TV, high-grade tubulovillous adenoma; L, left adenoma; R, right adenoma.

exclusive alteration in 49 (85%) of the 58 polyps (Fig. 3). Finally, *MGMT* methylation and K-RASm were exclusive in 55 (74%) of the 74 polyps (Fig. 4).

3.6.3. Loci targeting chromosomes 1, 6, and 9

To show that a subset of microsatellites could represent targets for predictive factors of progression, we performed a 2-way hierarchical clustering analysis using a set of loci associated with tubulovillous adenomas and/or high-grade adenomas, because these loci are considered of higher risk for progression. Of interest, the clustering of the allelotyping data, targeting loci *DIS197*, *DIS207*, *DIS225*, *DIS305*, *D6S264*, *D6S275*, and *D9S179* for the 99 polyps, separated 60% of polyps with tubulovillous and/or high-grade features (Fig. 5).

4. Discussion

In this study, we investigated genomic alterations in a large cohort of colon polyps of a non-MSI phenotype, assessing the allelotypes of 33 microsatellites, *MGMT* promoter gene methylation, and K-RASm. We correlated these results with clinical and histologic features to identify the molecular basis of CIN colon carcinogenesis initiation. In our cohort, several distinct patterns of molecular alterations involving AI, K-RASm, and *MGMT* methylation were identified (Fig. 6), suggesting distinct pathways for the development of carcinogenesis. Interestingly, molecular alterations located on chromosomes 1, 6, and 9 were associated with the histologic tubulovillous component and high-grade dysplasia.

Our previous study focused on invasive colon cancer of the CIN phenotype [22]. Carcinomas were found to be separated into 3 subgroups, according to the global AI frequency and specific loci distribution without correlation to the clinical stage. This suggests that a subset of alterations, present at early stage, could be sufficient for metastasis. This observation is in agreement with other studies [17,28]. We therefore questioned whether such molecular subtypes could

be detected in precancerous colon lesions. For this purpose, a collection of 159 colon polyps, with established CIN status and histologic features, was analyzed. Clustering analysis identified 4 subgroups of polyps according to their AI frequency and altered loci distribution. One set (CIN-VLa) contained polyps with an altered *APC* locus and few simultaneous K-RASm, possibly corresponding to the canonical pathway [1], where the inactivation of the gene *APC* initiates the adenoma-carcinoma sequence. In subgroup CIN-VLb, polyps showed a heterogeneous distribution of AI, *MGMT* methylation, and K-RASm with missed alterations at locus *APC*, suggesting an additional pathway for tumor initiation linked to CIN. Finally, a subgroup of polyps without AI at the studied loci is characterized by *MGMT* methylation or K-RASm. In our cohort of a small number of cases without any genomic alteration, the 33 analyzed loci or *MGMT* methylation status or K-RASm or the comparative genomic hybridization array analysis available on 8 of these 23 polyps did not reveal any further specific genomic alterations (data not shown). Therefore, we hypothesize that these molecular subtypes underlie different modes of initiation of colon carcinogenesis into the non-MSI pathway,

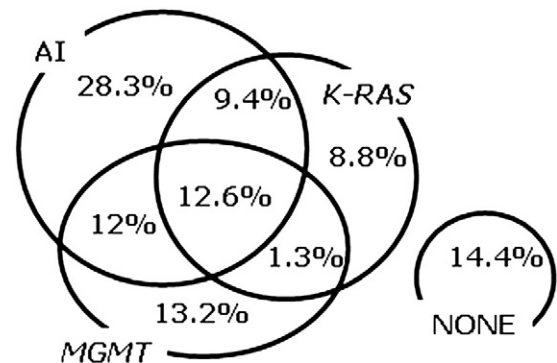


Fig. 6 Different modes of initiation of non-MSI colon carcinogenesis. There are at least 4 modes of initiation in the non-MSI pathway. *MGMT* indicates promoter methylation of *MGMT* gene; *K-RAS*, K-RASm; NONE, group free of any of the studied alterations.

in addition to the Fearon-Vogelstein model involving *APC* alteration as a primary initial event [1]. This is in agreement with the recent concept of colorectal carcinogenesis [9,29]. Apparently, colorectal cancer evolves through multiple pathways composed of many individual genetic and epigenetic alterations [9]. This multiplicity of pathways could explain the variations in tumor behavior and responsiveness to therapy [29]. Our results suggest that a potential pathway could associate K-RASm with alterations at chromosome arms 1p and 3q, but it remains to be confirmed in another cohort. In addition, the existence of a small group free of alterations reinforces other studies on colon cancer [4] in which other genes of the Wnt [30] or Notch signaling pathways [31] could initiate an adenoma.

Most colorectal carcinomas develop from preformed adenomas, but only a minority of adenomas undergo malignant transformation. As described in the literature [21,32], the risk factors for evolution of adenoma are mostly size, location, age, tubulovillous or villous histology, and the presence of high-grade dysplasia. Interestingly, in our cohort of non-MSI polyps, no villous or serrated adenoma was detected because, as expected, serrated adenomas were related to the MSI phenotype [33]. In our study, whatever alterations were considered, no significant differences between right and left location were observed. To our knowledge, studies involving colon carcinoma, which distinguished right and left locations either by large-scale transcriptome analyses or by epigenetic and genetic alterations, mostly included the MSI or CIN phenotypes [34] or addressed a small cohort [35]. By these genomic analyses, we did not observe any significant differences between right and left colon, which could be linked to the absence of MSI phenotype because the MSI phenotype is found preferentially in right-sided colon tumors [36].

We were not able to identify any independent specific locus alterations associated with size, highlighting the absence of specific genomic signatures involved in proliferation. A significant association between the sizes of the polyps, the main clinical parameter for a usual predictive factor, and the global AI frequency or K-RASm was observed in our cohort. Although K-RASm has been linked to aggressive carcinoma [37,38], no meta-analysis has confirmed this alteration as a valid prognostic factor in colorectal cancer [39].

A strong correlation was identified between the tubulovillous-type, high-grade dysplasia, and the global AI frequency. The impact of histologic features for the risk of advanced adenoma has already been demonstrated by epidemiologic studies [21,32]. To our knowledge, our study is the first that links histology and molecular alterations within a cohort of colon polyps of non-MSI status. Indeed, clustering the allelotype data on loci targeting chromosomes 1, 6, and 9 has allowed us to associate adenomas with tubulovillous histology and/or high-grade dysplasia. In the future, techniques such as comparative genomic hybridization array, single nucleotide polymorphism array, or high-

throughput sequencing will be required to identify specific genes associated with these loci.

In conclusion, genomic molecular analyses of 159 non-MSI colon polyps demonstrate a large variety of molecular alteration associations, supporting the idea that several distinct sets of alterations can be present at the precancerous stage. In addition to the subgroup related to the canonical Wnt/APC pathway, other molecular pathways for tumor initiation in precancerous colon lesions must exist. Finally, alterations in loci located on chromosomes 1, 6, and 9 are strongly correlated with tubulovillous and high-grade features. The genes that can be identified and associated with altered loci may be used to predict risk of recurrence/progression in colorectal carcinogenesis.

Acknowledgment

We thank Mrs Sylvie Delacour, Florence Guenard, Isabelle Keller, Veronique Kussaibi, Laetitia Ruck, Tuy-Thien Tong, and Mr Fabien Grüner for their excellent technical assistance and Mr Etienne Bergmann for the data management of the Centre de Ressources Biologiques. This work was supported by the University Hospital of Strasbourg, the University of Strasbourg, Inserm, Ligue contre le Cancer Comités Départementaux du Bas-Rhin et du Haut-Rhin, Institut National du Cancer, and Cancéropôle Grand-Est.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.humpath.2010.07.003](https://doi.org/10.1016/j.humpath.2010.07.003).

References

- [1] Fearon E, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61:759-67.
- [2] Gisselsson D. Chromosome instability in cancer: how, when, and why? *Adv Cancer Res* 2003;87:1-29.
- [3] Kolodner RD, Marsischky GT. Eukaryotic DNA mismatch repair. *Curr Opin Genet Dev* 1999;9:89-96.
- [4] Jass JR. Colorectal cancer: a multipathway disease. *Crit Rev Oncog* 2006;12:273-87.
- [5] Harada K, Hiraoka S, Kato J, et al. Genetic and epigenetic alterations of Ras signaling pathway in colorectal neoplasia: analysis based on tumor clinicopathological features. *Br J Cancer* 2007;97:1425-31.
- [6] O'Brien MJ, Yang S, Mack C, et al. Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. *Am J Surg Pathol* 2006;30:1491-501.
- [7] Nowak MA, Komarova NL, Sengupta A, et al. The role of chromosomal instability in tumor initiation. *Proc Natl Acad Sci U S A* 2002;99:16226-31.

- [8] Kondo Y, Issa JP. Epigenetic changes in colorectal cancer. *Cancer Metastasis Rev* 2004;23:29-39.
- [9] Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 2007;50:113-30.
- [10] Derks S, Postma C, Moerkert PT, et al. Promoter methylation precedes chromosomal alterations in colorectal cancer development. *Cell Oncol* 2006;28:247-57.
- [11] Judson H, Stewart A, Leslie A, et al. Relationship between point gene mutation, chromosomal abnormality, and tumor suppressor gene methylation status in colorectal adenomas. *J Pathol* 2006;210:344-50.
- [12] Krtolica K, Krajnovic M, Usaj-Knezevic S, Babic D, Jovanovic D, Dimitrijevic B. Comethylation of p16 and *MGMT* genes in colorectal carcinoma: correlation with clinicopathological features and prognostic value. *World J Gastroenterol* 2007;13:1187-94.
- [13] Nagasaka T, Goel A, Notohara K, et al. Methylation pattern of the O6-methylguanine-DNA methyltransferase gene in colon during progressive colorectal tumorigenesis. *Int J Cancer* 2008;122:2429-36.
- [14] Esteller M, Herman JG. Generating mutations but providing chemosensitivity: the role of O6-methylguanine DNA methyltransferase in human cancer. *Oncogene* 2004;8:1-8.
- [15] Kaina B, Christmann M, Naumann S, Roos WP. M.G.M.T: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA repair* 2007;6:1079-99.
- [16] Diep CB, Kleivi K, Ribeiro FR, Teixeira MR, Lindgjaerde OC, Lothe RA. The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. *Genes Chromosomes Cancer* 2006;45:31-41.
- [17] Hermesen M, Postma C, Baak J, et al. Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. *Gastroenterology* 2002;123:1109-19.
- [18] Saini SD, Kim HM, Schoenfeld P. Incidence of advanced adenomas at surveillance colonoscopy in patients with a personal history of colon adenomas: a meta-analysis and systematic review. *Gastrointest Endosc* 2006;64:614-26.
- [19] Bonithon-Kopp C, Piard F, Fenger C, et al. Colorectal adenoma characteristics as predictors of recurrence. *Dis Colon Rectum* 2004;47:323-33.
- [20] Gupta S. Colorectal polyps: the scope and management of the problem. *Am J Med Sci* 2008;336:407-17.
- [21] Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US multi-society task force on colorectal cancer and the American cancer society. *Gastroenterology* 2006;130:1872-85.
- [22] Weber JC, Meyer N, Pencreach E, et al. Allelotyping analyses of synchronous primary and metastasis CIN colon cancers identified different subtypes. *Int J Cancer* 2007;120:524-32.
- [23] Boland CR, Thibodeau SN, Hamilton SR, et al. National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998;58:5248-57.
- [24] Snover DC, Jass JR, Fenoglio-Preiser C, Batts KP. Serrated polyps of the large intestine. *Am J Clin Pathol* 2005;124:380-91.
- [25] Schneider A, Rohr S, Kelly MD, et al. Microsatellite instability and allelic imbalance in primary and secondary colorectal cancer. *Aust N Z J Surg* 2000;70:587-92.
- [26] Delattre O, Olschwang S, Law DJ, et al. Multiple genetic alterations in distal and proximal colorectal cancer. *Lancet* 1989;2:353-6.
- [27] Olschwang S, Hamelin R, Laurent-Puig P, et al. Alternative genetic pathways in colorectal carcinogenesis. *Proc Natl Acad Sci U S A* 1997;94:12122-7.
- [28] Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet* 2003;33:49-54.
- [29] Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science* 2007;318:1108-13.
- [30] Thorstensen L, Lind GE, Løvig T, et al. Genetic and epigenetic changes of components affecting the WNT pathway in colorectal carcinomas stratified by microsatellite instability. *Neoplasia* 2005;7:99-108.
- [31] Chu D, Wang W, Xie H, et al. Notch1 expression in colorectal carcinoma determines tumor differentiation status. *J Gastrointest Surg* 2009;13:253-60.
- [32] Gschwantler M, Kriwanek S, Langner E, et al. High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics. *Eur J Gastroenterol Hepatol* 2002;14:183-8.
- [33] East JE, Saunders BP, Jass JR. Sporadic and syndromic hyperplastic polyps and serrated adenomas of the colon: classification, molecular genetics, natural history, and clinical management. *Gastroenterol Clin North Am* 2008;37:25-46.
- [34] Komuro K, Tada M, Tamoto E, et al. Right- and left-sided colorectal cancers display distinct expression profiles and the anatomical stratification allows a high accuracy prediction of lymph node metastasis. *J Surg Res* 2005;124:216-24.
- [35] Deng G, Kakar S, Tanaka H, et al. Proximal and distal colorectal cancers show distinct gene-specific methylation profiles and clinical and molecular characteristics. *Eur J Cancer* 2008;44:1290-301.
- [36] Chang SC, Lin JK, Yang SH, Wang HS, Li AF, Chi CW. Relationship between genetic alterations and prognosis in sporadic colorectal cancer. *Int J Cancer* 2006;118:1721-7.
- [37] Conlin A, Smith G, Carey FA, Wolf CR, Steele RJ. The prognostic significance of K-ras, p53, and APC mutations in colorectal carcinoma. *Gut* 2005;54:1283-6.
- [38] Hsieh JS, Lin SR, Chang MY, et al. APC, K-ras, and p53 gene mutations in colorectal cancer patients: correlation to clinicopathologic features and postoperative surveillance. *Am Surg* 2005;7:336-43.
- [39] Locker GY, Hamilton S, Harris J, et al. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol* 2006;24:5313-27.

4. Proposition de séquences d'initiation adénomateuse

L'ensemble de nos travaux permet de proposer quatre schémas d'initiation d'un adénome à partir de la muqueuse colique saine (**Figure 25**) :

Cas 1) La voie classique de carcinogenèse (d'après B. Vogelstein) avec les altérations d'*APC*, de *TP53* ainsi que d'autres AI, et/ou la mutation de *KRAS*, et/ou la méthylation

Cas 2) Une voie avec AI (au niveau des sites étudiés), mais sans perte allélique d'*APC*, de *TP53*, sans mutation du gène *KRAS* et sans méthylation du gène *MGMT*

Cas 3) Une voie sans AI (au niveau des sites étudiés) avec/sans mutation du gène *KRAS* et des modifications épigénétiques comme la méthylation du gène *MGMT*

Cas 4) Une voie ne présentant aucune des altérations étudiées et ciblant 14% des adénomes.

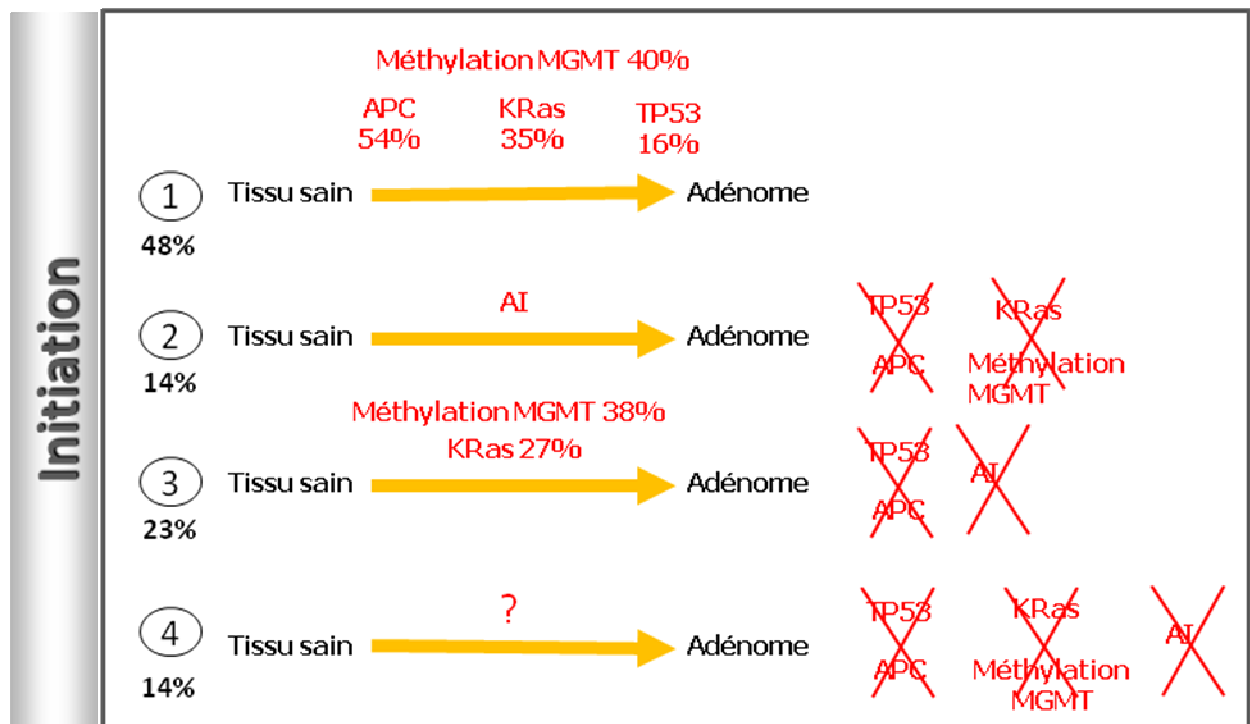


Figure 25 : Proposition de séquences d'initiation adénomateuse pour les adénomes de phénotype non-MIN.

5. Altérations moléculaires selon la localisation

5.1. Polypes avec AI

Dans le groupe des polypes présentant au moins un AI (n=99), il existe des différences d'incidence des altérations ciblant *APC* et *TP53* selon la localisation proximale ou distale, (**Figure 26**). En effet, *APC* et *TP53* sont plus fréquemment altérés dans le côlon distal comparé au côlon proximal, avec 55% vs 38% pour *APC* respectivement, 22% vs 3% pour *TP53* respectivement (p=0,04).

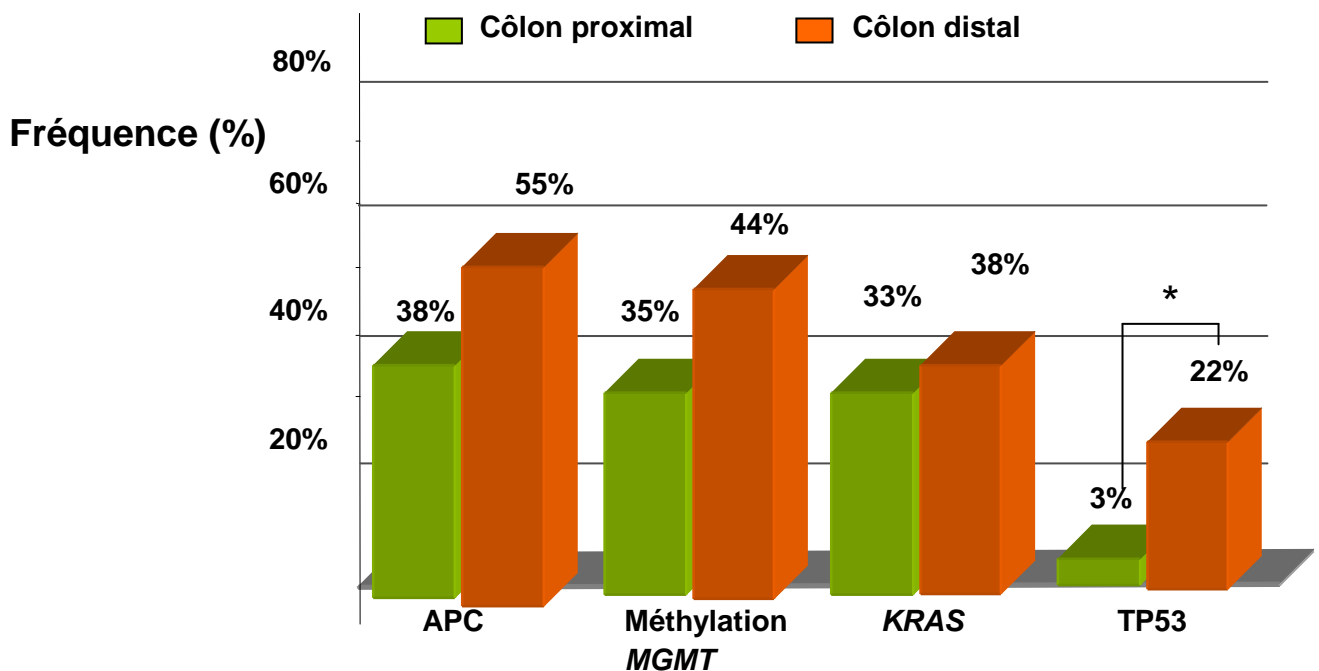


Figure 26 : Fréquence d'altérations de *APC*, *MGMT*, *KRAS* et *TP53* en fonction de la localisation dans les polypes avec AI.

5.2 Polypes sans AI

L'existence du sous-type non-CIN non-MIN sur les loci analysés et mis en évidence dans 60 adénomes de tous grades et types histologiques, mais avec une majorité de polypes hyperplasiques et d'adénomes tubuleux de bas grade, montre que des adénomes peuvent se développer sans altération chromosomique majeure dans les sites étudiés.

Dans ce sous-groupe, il n'a pas été mis en évidence de différence entre les altérations recherchées et la localisation proximale ou distale des polypes (**Figure 27**), ni d'altérations associées spécifiquement à un sous-type histologique ou à un degré de dysplasie.

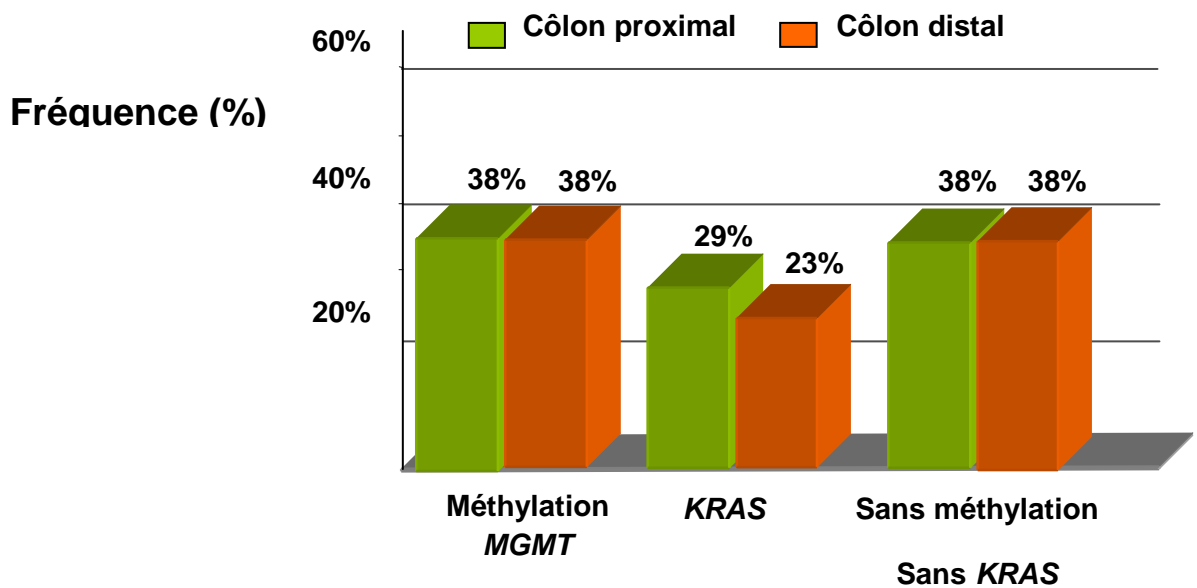


Figure 27 : Fréquence d'altérations de MGMT et KRAS, et d'absence d'altérations en fonction de la localisation dans les polypes sans AI.

6. Altérations moléculaires sur l'ensemble des chromosomes

Afin de préciser la nature des altérations observées par allélotypage et d'étendre l'analyse à l'ensemble des chromosomes, une hybridation génomique comparative à l'aide de puces à oligonucléotides (CGH array ou CGHa) a été réalisée sur 102 adénomes et 8 polypes hyperplasiques, dont 83 adénomes et 6 polypes hyperplasiques appartiennent à notre cohorte, associés à 58 carcinomes invasifs de phénotype CIN.

6.1 Polypes de type non-CIN non-MIN

Parmi les 23 polypes ne comportant aucune altération aux loci examinés par allélotypage, ni mutation de *KRAS*, ni méthylation de *MGMT*, 8 ont pu être examinés par CGHa (3 polypes hyperplasiques, 3 adénomes tubuleux de bas grade, 1 adénome tubulo-villeux de bas grade et 1 adénome tubulo-villeux de haut grade). Il n'a pas été détecté d'altération chromosomique majeure qui n'aurait pas été relevée par allélotypage, en particulier au niveau des chromosomes 7 et 19 non ciblés par notre panel de 33 microsatellites (**Figure 28**).

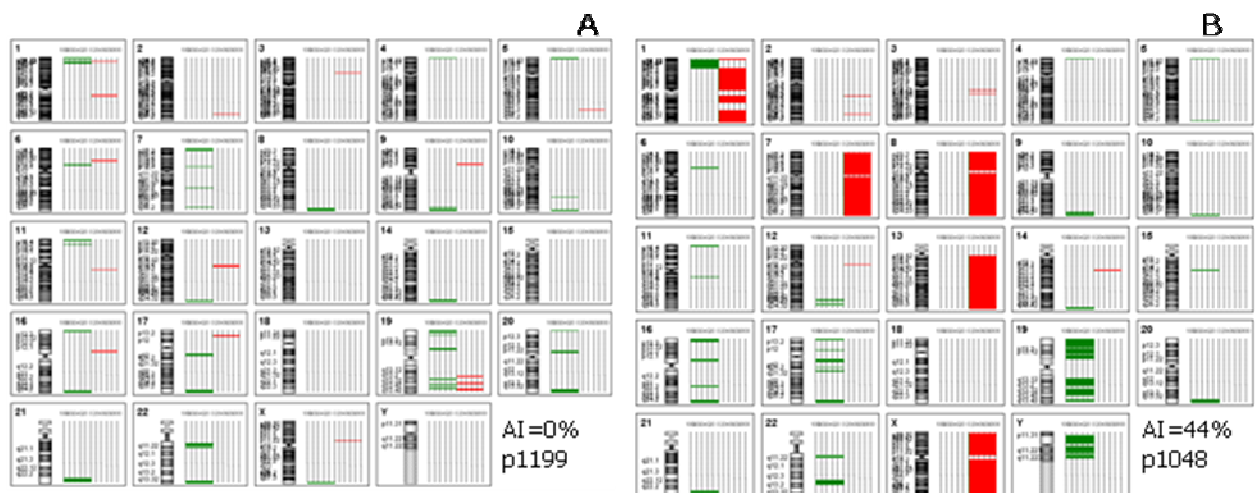


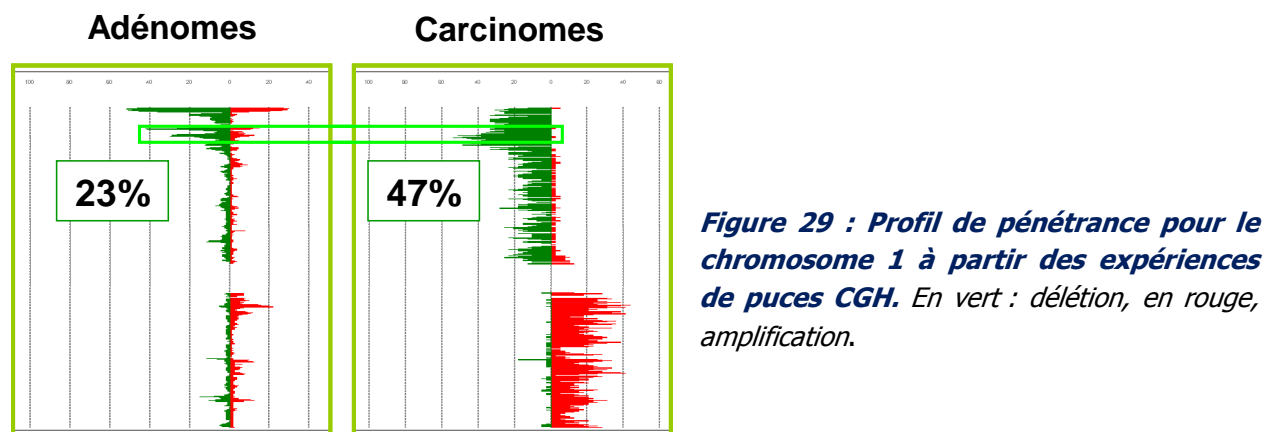
Figure 28 : Résultat de l'analyse par CGH pour un cas d'adénome non-CIN non-MIN (A) et un cas d'adénome fortement altéré (B). En vert : délétion, en rouge, amplification.

L'analyse des données de la CGHa est en cours, mais a permis de montrer une bonne corrélation entre la fréquence d'AI par allélotypage et le taux d'altérations observés par CGHa.

6.2 Identification d'une micro-délétion

L'analyse statistique préliminaire de la CGHa a permis d'identifier, parmi de nombreux gains et pertes, une région génomique localisée en 1p36.11-12 et délétée dans 23% des adénomes de phénotype non-MIN et 47% des carcinomes CIN (**Figure 29**). L'altération complète du bras chromosomique 1p a déjà été décrite dans le CC mais la micro-délétion en 1p36.11-12 n'a jamais été observée avec une telle pénétrance.

Ce travail, auquel j'ai collaboré, essentiellement porté par Céline Nicolet doctorante dans notre équipe, avait pour but d'identifier, au sein de cette région chromosomique, des gènes potentiellement impliqués dans le processus de carcinogenèse colique et d'étudier l'impact pronostique de cette délétion.



Parmi les gènes localisés en 1p36.11-12, le facteur de transcription *E2F2* peut être un gène intéressant de par son rôle de régulateur du cycle cellulaire. En effet, la QPCR ciblant le gène *E2F2* montre que pour 149 carcinomes CIN, 62 sont délétés pour *E2F2* (42%), avec 35/84 carcinomes distaux (42%) et 27/65 carcinomes proximaux (42%). Cependant, dans les carcinomes distaux, la fréquence de délétion diminue avec l'évolution des stades cliniques (58% au stade I et 34% au stade IV), alors que dans les carcinomes proximaux, la fréquence reste stable.

Cette observation suggère une potentielle valeur pronostique de la délétion génomique du gène *E2F2*, hypothèse confirmée par l'analyse de Kaplan-Meier des courbes de survie sans événement et par le modèle de régression de Cox : la délétion serait de bon pronostic dans les

carcinomes distaux ($p=0.027$) mais pas dans les carcinomes proximaux ($p=0.626$). Cette observation conforte l'hypothèse que les mécanismes de carcinogénèse seraient différents dans les deux divisions anatomiques du côlon et qu'il est important de les étudier séparément.

***E2F2* deletion as a genomic marker for prognosis
in distal CIN colon carcinomas**

Nicolet C, Voegelin M, Guérin E, Marisa L, Neuville A,
Flaman JM, Brigand C, Gaub MP and Guenot D

Human Molecular Genetic

Soumis

7. Proposition de distribution des altérations en fonction de la localisation

Les données d'allélotypage et de CGHa ont permis de mettre en évidence des différences génomiques dans les adénomes de localisation proximale et distale. La distinction de la localisation proximale est souvent liée à profil génomique particulier comme dans la voie festonnée ou lors du statut CIMP-high.

Dans notre étude, nous pouvons proposer des altérations génomiques impliquées au stade précoce de développement des adénomes (**Figure 30**).

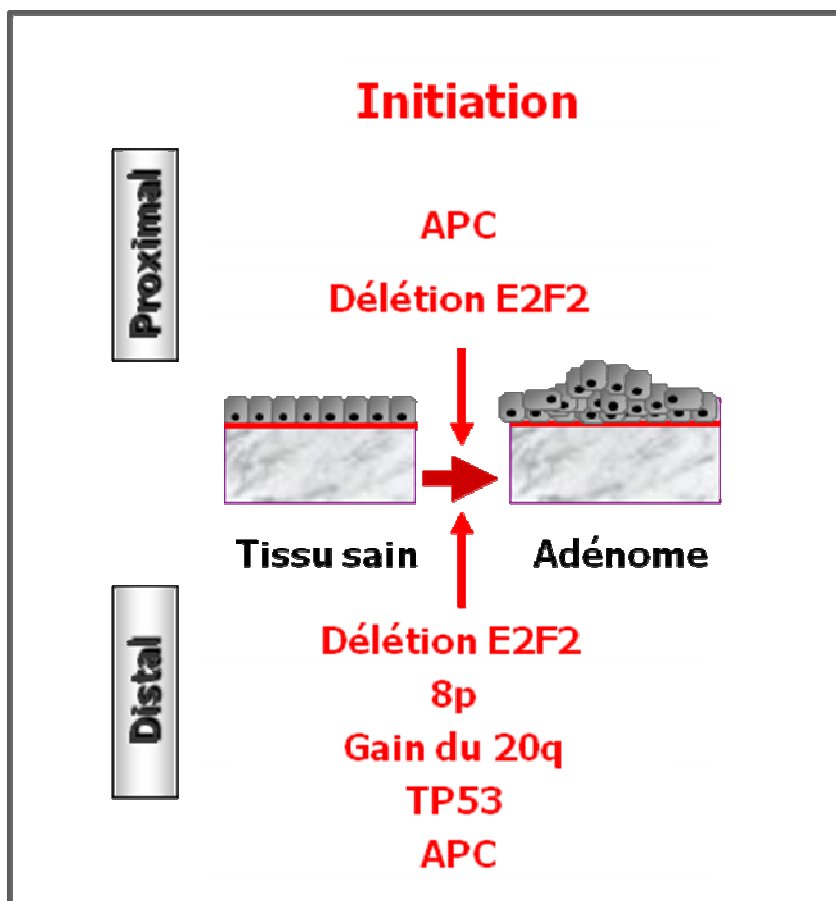


Figure 30 : Altérations génomiques pouvant avoir un rôle prépondérant dans l'initiation de la transformation adénomateuse en fonction de la localisation.

DISCUSSION

Mon travail a porté sur la recherche d'altérations génomiques impliquées dans les mécanismes d'initiation d'une série de polypes coliques de phénotype non-MIN, pour tester l'hypothèse de l'existence de différentes associations d'évènements sous-tendant la carcinogenèse colique humaine, en comparaison avec les données accumulées dans les carcinomes invasifs, et expliquant son hétérogénéité.

Les résultats de ce travail ont permis de mieux définir, entre autres, les altérations ciblant les loci *APC* et *TP53*, la mutation de l'oncogène *KRAS* et la méthylation du promoteur du gène *MGMT*, et leur répartition dans les localisations proximale et distale.

Toute la difficulté à proposer une séquence d'altérations initiatrice du processus de carcinogenèse colique réside dans le fait qu'il faut pouvoir dissocier les anomalies minimales indispensables au pouvoir carcinogène de celles qui ne seraient que des conséquences des premières et/ou de l'adaptation au micro-environnement.

Ce travail montre :

- l'existence d'autres voies de carcinogenèse dans les stades précoces, confirmant nos observations dans les carcinomes invasifs (Weber, Meyer et al. 2007) ;
- que des anomalies moléculaires pourraient être associées à un sous-type histologique et au grade histologique des adénomes ;
- que les adénomes proximaux et distaux se formeraient par des voies de carcinogenèse distinctes.

1. Sous-types moléculaires des adénomes de phénotype CIN

1.1 Corrélation avec le modèle de B. Vogelstein

L'étude des altérations génomiques par allélotypage ciblant des polypes non-MIN de tous grades et types histologiques a montré qu'ils peuvent être classés selon la fréquence globale d'altération. Dans la cohorte d'adénomes présentant des AI, nous avons identifié un sous groupe qui reproduit le modèle d'accumulation séquentielle d'altérations de B. Vogelstein (Fearon, Vogelstein, 1990), mais ne représente qu'environ 50% de notre cohorte. Dans les adénomes, le locus associé à *APC* est le plus fréquemment altéré avec une incidence moyenne de 54%. Il peut être altéré seul (dans 24% des cas) ou être associé à des altérations de *KRAS* (dans 25% des cas) et/ou de *TP53* (dans 14% des cas).

Concernant la perte allélique du locus *TP53*, il faut noter que celle-ci est quasi inexistante dans les polypes proximaux, alors que pour le côlon distal, la perte existe au stade d'adénome dans 22% des cas.

L'existence de sous-groupes moléculaires selon la localisation droite et gauche a déjà été suggérée, en particulier dans les CC (Gervaz, Bouzourene et al. 2001, Russo, Bazan et al. 2005). Cependant la plupart de ces études, par CGH, CGHa ou transcriptome, portaient sur des carcinomes de phénotype CIN et MIN, incluant parfois des carcinomes de localisation rectale (Birnkamp-Demtroder, Christensen et al. 2002, Williams, Gaynor et al. 2003).

Nos travaux ne portent que sur des adénomes coliques non-MIN. Dans notre série, la séquence d'apparition de l'altération dans les stades précoces serait très différente pour les deux localisations et le modèle de Vogelstein s'appliquerait principalement au côlon distal.

1.2 Discordances concernant les marqueurs précoces par rapport à ce modèle

Cependant, dans 50% des autres cas, ce modèle ne semble pas pouvoir s'appliquer.

Notre allélotypage a montré que l'altération de *TP53* pouvait exister sans altération d'*APC* dans la moitié des cas altérés pour ce loci, suggérant une possible implication précoce de TP53 et non plus tardive comme suggéré par le modèle de B. Vogelstein. Comme nous l'avons précisé précédemment, cette altération est préférentiellement liée à la localisation distale.

Ceci amène à proposer que dans le côlon proximal, la perte de p53 serait un évènement tardif alors que dans le côlon distal, cette perte pourrait être initiale.

L'altération du gène *KRAS* est classiquement décrite dans la cascade d'accumulation d'altération selon B. Vogelstein et serait impliquée plus tardivement qu'*APC*. Notre étude a montré l'existence de cette mutation dans 32% des cas, sans distinction selon la localisation proximale et distale. Cependant, des mutations de *KRAS* existent de façon isolée, sans altérations d'*APC* et de *TP53*, dans 8,8% des polypes qui sont des adénomes avec une prédominance de bas grade (11) versus 3 de haut grade, ou associées à des AI sans perte allélique d'*APC*, ou associées à la méthylation de *MGMT* sans AI.

Ainsi *KRAS* pourrait avoir un rôle initiateur, qui devra être précisé dans une cohorte plus importante intégrant des FCA. La mutation de *KRAS* ne semble pas être associée à une localisation particulière et semble donc un évènement majeur dans la transformation cancéreuse. Ce résultat est en accord avec nos travaux portant sur les CC où il n'est pas observé de distribution préférentielle des mutations de *KRAS* entre les localisations proximale et distale (Weber, Meyer et al. 2007). Peu d'observations ont rapporté une telle corrélation entre la mutation de *KRAS* et la localisation, et seulement dans les carcinomes (Bleeker, Hayes et al. 2000).

Une des altérations génomiques les plus fréquentes dans le CC est le gain du chromosome 20q. Dans la littérature, ce gain a été décrit comme intervenant dans la transition adénome - carcinome (Ried, Knutzen et al. 1996).

Notre travail a permis de montrer l'existence d'un gain du chromosome 20q dans 29% des adénomes de localisation distale et 11,6% des adénomes de localisation proximale (**Figure 30**). Cette proportion s'accroît dans les CC avec un gain du 20q dans 81% de localisation distale et 66% de localisation proximale.

Il a été déterminé deux mécanismes de gain de l'intégralité du bras chromosomique 20q : la tétrasomie, indiquant un gain simultané des deux chromosomes parentaux, et l'aneusomie. De façon intéressante, l'incidence du gain du 20q par tétrasomie est plus élevée dans l'adénome distal et elle ne varie plus avec la progression du CC ; par contre dans l'adénome proximal, l'incidence augmente plus tardivement (publication 1 en Annexes).

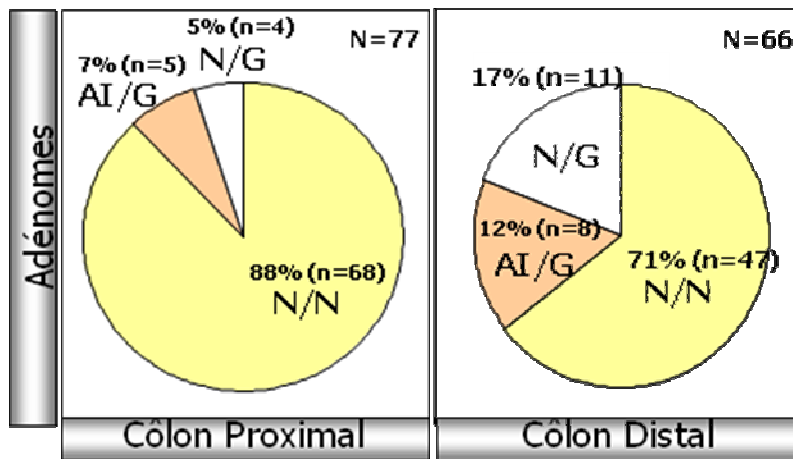


Figure 30 : Incidence de l'altération du chromosome 20q dans les adénomes selon le mécanisme du gain [aneusomie (AI/G) ou tétrasomie (N/G)] et la localisation.

Par conséquent, le gain du 20q par tétrasomie participerait avec d'autres anomalies comme *APC* et *TP53*, à l'initiation de l'adénome dans le côlon distal. Dans la littérature, la tétraploïdie est décrite comme un état ciblant tous les chromosomes et pouvant être soit un intermédiaire avant l'aneuploïdie, soit un état en soi.

Dans notre cohorte, en particulier dans le sous-groupe des 23 polypes sans altérations par allélotypage, parmi les 8 polypes sans prédominance de localisation particulière, ayant pu être analysés par CGHa, aucun ne montre de gain du 20q.

En conclusion, nos travaux permettent de confirmer l'existence de voies de carcinogénèse différentes. Ces voies mettent en jeu les acteurs principaux déjà décrits dans la littérature comme *APC*, *TP53* et *KRAS*. Mais ceux-ci peuvent être tour à tour initiateur de la transformation adénomateuse dans des combinaisons différentes de celles proposées par B. Vogelstein, selon leur localisation ou d'autres mécanismes non encore connus.

2. Autres altérations génomiques impliquées dans l'initiation

Les séquences d'initiation des adénomes de phénotype non MIN proposés dans nos travaux (**Figure 25**, page 64) indiquent à côté du modèle classique de Vogelstein d'autres voies moléculaires ne faisant pas intervenir *APC* comme évènement initiateur de la cancérogenèse. Ce nouveau concept, récemment admis (Jass, 2006, Wood, Parsons, et al. 2007), a peu été exploré dans les adénomes coliques (Hermsen, Postma, et al. 2002). Cependant, d'autres

altérations initiatrices sont encore à rechercher, puisque dans notre cohorte, 60 polypes ne présentent aucun remaniement aux 33 loci examinés, dont 23 sans aucune altération au niveau du gène *KRAS*.

Nous avons pu, dans 8 des 23 polypes de notre sous-groupe sans altérations, confirmer l'absence d'altérations majeures sur l'ensemble des chromosomes par une analyse CGHa. Néanmoins l'étude finale de la CGHa sur les polypes est en cours. Elle a déjà permis de détecter des anomalies plus fines comme les micro-délétions du locus 1p36.11-12, d'abord analysées dans les CC, puis observées dans les polypes. Enfin, nous avons pu observer que la méthylation de *MGMT* pouvait aussi être fréquente dans les adénomes non-MIN, suggérant, comme dans les CC de type MIN, un rôle initiateur dans la carcinogenèse.

2.1 L'altération du locus 1p36.11-12

La région 1p36.11-12 a été identifiée en analysant les données des puces CGHa d'adénomes et de carcinomes. Cette région est délétée dans 26% des adénomes et 8% ont une amplification, alors qu'elle est délétée dans 47% des CC.

Les adénomes présentant l'amplification en 1p36.11-12 sont en majorité tubuleux (75%) de bas grade (63%) et appartiennent au sous-type CIN-Very Low (63%). Or les foyers carcinomateux se développent préférentiellement à partir d'adénomes vilieux ou tubulo-vilieux et/ou de haut grade (Gschwantler, Kriwanek et al. 2002, Winawer, Zauber et al. 2006). Comme l'amplification de cette région est inexistante dans les carcinomes, cette altération ne serait pas favorable à la transition adénome – carcinome. Il serait intéressant d'analyser un plus grand nombre d'adénomes pour évaluer l'incidence de cette amplification afin de définir si elle est associée à d'autres anomalies génomiques défavorables à la transformation des adénomes.

La délétion de la région 1p36.11-12 est caractéristique des adénomes proximaux, où elle pourrait être un événement initiateur puisque son incidence passe de 21% dans les adénomes à 33% dans les carcinomes pour ensuite ne plus varier avec la progression tumorale ; dans le côlon distal par contre, cet événement est plus tardif et intervient plutôt au cours de la transition adénome - carcinome. Par conséquent, la délétion en 1p36.11-12 participerait, avec d'autres anomalies, à l'émergence d'adénomes proximaux.

A l'heure actuelle, la région 1p a encore été peu impliquée dans la carcinogenèse colique (Thorstensen, Qvist et al. 2000, Di Vinci, Infusini et al. 1996), vraisemblablement du fait que la région étudiée ne concerne qu'une micro-région de 0,7 Mbp non observable par CGH classique

ou par cytogénétique. L'analyse par CGHa permettra de définir les associations avec d'autres altérations ou mutations. Comme cette altération est plus fréquemment associée aux CC proximaux, il serait intéressant de rechercher la présence de cette micro-délétion dans les carcinomes de phénotype MIN plus fréquemment proximaux, ce qui permettrait de mieux comprendre son importance dans la cancérogenèse colique.

2.2 Rôle de *MGMT*

Dans notre travail, la méthylation du gène *MGMT* est retrouvée dans 39% de notre cohorte (62/159) et dans près de 40% des polypes avec AI (39/99). De façon intéressante, cette altération est présente dans 23 des 60 polypes sans AI (38%), dont 21 de façon isolée (35%), sans mutation de *KRAS*.

La méthylation de *MGMT* est un évènement fréquent dans le CC, mais les études publiées analysent le plus souvent des séries de CC invasifs ou d'adénomes sans connaître leur statut MSI (Krtolica, Krajnovic, et al. 2007, Nagasaka T, Goel, et al. 2008). Or il est bien connu maintenant que le gène *MGMT* a un rôle de gène mutateur impliqué dans la voie de cancérogenèse colique MIN (voir § 2.4.3), mais qu'il intervient également dans les voies de cancérogenèse CIN (Jass, 2007). Dans l'étude de Svrcek et al (2010), la méthylation de *MGMT* est plus fréquemment corrélée aux carcinomes de phénotype MIN, mais cette altération est aussi détectée dans 21 des 49 CC sporadiques de phénotype CIN.

Le rôle de *MGMT* dans les voies CIN et non-MIN non-CIN, en particulier les mécanismes par lesquels *MGMT* pourrait influencer l'initiation de la cancérogenèse (gènes cibles, cancérisation en champ, gènes associés, origine de la méthylation...) restent encore à définir. Les adénomes non MIN non CIN seraient-ils des lésions pré-MIN ? Mais il n'existe pas, dans notre groupe de polypes non MIN non CIN, de corrélation avec un type histologique, alors que les polypes associés à un phénotype MIN peuvent être des adénomes festonnés (Jass, 2007), ce qui laisse supposer que les adénomes non MIN non CIN seraient une entité à part, en accord avec l'étude par CGHa portant sur les carcinomes coliques non MIN non CIN (Jones, Douglas et al. 2005).

3. Corrélation entre combinaison d'altérations et caractéristiques histologiques

Nous avons montré que l'association des altérations moléculaires ciblant les chromosomes 1q, 6q et 9q était corrélée avec le sous-type histologique tubulo-villeux et avec le haut grade des adénomes (de type tubuleux ou tubulo-villeux). A l'heure actuelle, aucun de ces marqueurs seuls ne sont pronostiques dans les CC. L'association histologique tubulo-villeux et haut grade est cohérente avec les données de la littérature. En effet, la présence d'un contingent vilieux au sein d'un adénome est souvent associée à un risque plus élevé d'adénome avancé (Winawer, Zauber et al. 2006, Gschwantler, Kriwanek et al. 2002, Gupta, 2008). De la même façon, la présence d'une dysplasie de haut grade est associée à un risque plus élevé de récurrence d'adénome (Bonithon-Kopp, Piard et al. 2004), entraînant dans les recommandations de surveillance une périodicité d'examen par coloscopie de trois ans. En France, la présence du sous-type histologique tubulo-villeux ou vilieux ne modifie pas le schéma de surveillance mais dans les recommandations de surveillance post-polypectomie américaines (Winawer, Zauber et al. 2006), ce critère est utilisé pour fixer la périodicité de surveillance.

Or ces critères basés sur l'appréciation histologique des adénomes sont soumis à une variabilité inter-observateurs pouvant conduire à un sous ou sur-diagnostic, et donc à un sous ou sur-traitement (Denis, Peters et al. 2009). La détection de la dysplasie de haut grade est plus homogène parmi les pathologistes que le diagnostic d'adénome tubulo-villeux. Le sous-type histologique d'adénome vilieux pose moins de problème car les critères diagnostiques (plus de 80% d'architecture vilieuse) sont plus stringents. La recherche de cette association d'altérations serait aussi une aide au diagnostic.

La CGHa en cours permettra de limiter de façon précise les régions des chromosomes impliqués et d'identifier des gènes d'intérêt présents dans ces régions. Ainsi, l'identification de marqueurs moléculaires plus objectifs pourrait aider à la classification des adénomes. Au préalable, cette association 1q/6q/9q et types histologiques d'adénomes devra être confirmée sur une cohorte plus importante d'adénomes, avec une étude en double aveugle entre le pathologiste et le biologiste moléculaire.

PERSPECTIVES

Nos travaux indiquent l'existence de sous-types moléculaires associés à la localisation et à l'histologie des adénomes.

Pour mieux comprendre l'hétérogénéité des mécanismes fondateurs de la carcinogenèse colique, les prochaines étapes seront d'exploiter les données de la CGHa et de les confronter aux données d'allélotypage, d'évaluer le statut de la méthylation CIMP+ et CIMP- des adénomes, d'établir des données d'expression protéique par immunohistochimie et de comprendre les mécanismes d'initiation des adénomes non-CIN non-MIN et sans altération de *MGMT* et de *KRAS* en explorant, entre autres, la voie de signalisation Notch.

1. Etablir le statut CIMP des adénomes

Dans cette étude, seule la méthylation du gène *MGMT* a été recherchée. Or des données de la littérature décrivent la méthylation comme un événement initiateur prépondérant pour certains sous-types de carcinomes coliques (Derks, Postma et al. 2006; Judson, Stewart et al. 2006).

Le panel de gènes utilisé pour déterminer le phénotype CIMP n'est pas toujours le même (Nagasaka, Koi et al. 2008, Shen, Toyota et al. 2007, Kawasaki, Ohnishi et al. 2008), si bien qu'il pourrait être difficile d'établir un impact clinique ou pronostique des sous-groupes de carcinomes ainsi détectés. Cependant, une multiplex évaluant la méthylation des promoteurs de 5 autres gènes a été proposé comme étant la plus performante (*CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* et *SOCS1* - Weisenberger, Siegmund et al. 2006).

La détermination du phénotype CIMP de notre cohorte d'adénomes permettra de préciser l'importance de ces mécanismes épigénétiques dans les différents sous-types définis et leur association éventuelle avec un sous-type histologique.

2. Comprendre les mécanismes d'initiation des adénomes non-CIN non-MIN, sans altération de *MGMT* et de *KRAS*

L'absence de perte allélique d'*APC* et de *TP53*, bien que souvent corrélée au statut mutationnel, devra néanmoins être complétée par la recherche de mutations de ces gènes qui peuvent exister sans perte d'hétérozygotie. Leur rôle serait alors différent, possiblement lié à des altérations des voies de réparation autres que la déalkylation, comme l'excision de base (BER) ou l'excision de nucléotides (NER), que l'on sait être des mécanismes importants dans la

cancérogène (Ali, Meza et al. 2008, Joshi, Corral et al. 2009). D'autres gènes impliqués dans la voie Wnt pourraient également être mutés et ainsi moduler l'action d'APC comme l'Axine (Thorstensen, Lind et al. 2005).

Parmi les voies de signalisation non explorées, la voie Notch joue un rôle essentiel dans la différenciation, l'apoptose et la prolifération cellulaire. Notch est devenu une cible potentielle importante dans la carcinogénèse, en particulier dans la carcinogénèse colique (Reedijk, Odorcic et al. 2008, Chu, Wang et al. 2009). Son interaction avec la voie Wnt (Fre, Pallavi et al. 2009) en fait un candidat pour l'initiation du processus conduisant à la formation d'adénomes coliques. L'exploration de cette voie dans notre cohorte de polypes coliques non-MIN fera l'objet du sujet de master 2 de Noëlle Weingertner en 2010-2011.

Plus récemment, d'autres voies de signalisation ont été décrites comme impliquées dans la carcinogénèse. Ces autres voies potentiellement intéressantes à explorer sont la voie HGF/MET impliquée dans de nombreux cancers (Cecchi, Rabe et al. 2010), plus particulièrement une amplification du gène *c-MET* dans le CC (Zeng, Weiser et al. 2008), et la voie TGF β /SMAD dont des mutations de son récepteur TGF β -II seraient impliquées dans le CC (Bellam and Pasche, 2010, Valle, Serena-Acedo et al. 2008).

3. Analyser l'expression protéique des gènes altérés

L'impact réel des anomalies génomiques détectées dépend surtout du niveau d'expression des protéines produites par les gènes altérés. Pour cela, un TMA est en cours de réalisation sur la cohorte des 159 polypes.

Ce type de technique présente l'intérêt de pouvoir étudier simultanément et rapidement un grand nombre de patients dans le but d'identifier des marqueurs protéiques pronostiques et thérapeutiques facilement utilisables en routine hospitalière, et d'évaluer l'activation de différentes voies de signalisation et de certains de leurs acteurs comme EGFR, KRAS, BRAF, PI3K, PTEN, AKT, mTOR, HIF-1 α , VEGF, représentant des cibles thérapeutiques nouvelles.

Il permettra aussi de valider l'expression des gènes d'intérêts identifiés par la CGHa.

4. Approcher la problématique de l'hétérogénéité inter-tumorale

Au cours de mon travail de thèse, j'ai pu collaborer avec deux équipes utilisant des technologies différentes et innovantes, la spectrométrie par technique infra-rouge et la résonance magnétique nucléaire par spectrométrie à haute résolution de l'angle magique (HRMAS).

Ces deux techniques ont mis en évidence des spectres infra-rouge (publication 2 en Annexes) et des profils métaboliques (publication 3 en Annexes) dans des carcinomes coliques humains primitifs ou après xénogreffe.

Il serait intéressant de pouvoir tester, avec ces techniques, notre cohorte de polypes en vue de compléter les résultats des altérations génomiques, d'approcher les interactions cancer-microenvironnement et leur potentiel pronostique déjà décrit dans les CC (Ogino, Nosho et al. 2009) et d'ainsi ouvrir des pistes nouvelles pour la compréhension de la carcinogenèse colique.

CONCLUSION

L'objectif principal de ce travail était de tester l'hypothèse de l'existence de différentes voies d'initiation du cancer du côlon de phénotype non MSI et d'apporter des explications sur son hétérogénéité, à partir d'une série de 159 polypes, en comparaison avec les données accumulées dans les carcinomes invasifs.

A l'aide d'outils recherchant des altérations moléculaires génomiques, cette étude a montré, qu'en parallèle au modèle d'accumulation séquentielle d'altérations de B. Vogelstein, il existe d'autres voies de carcinogenèse dans les stades précoces, faisant intervenir de façon isolée ou associée une instabilité chromosomique sans altération du loci ciblant *APC*, en particulier celui ciblant *TP53*, un gain du chromosome 20q, une micro-délétion du chromosome 1p, la mutation de *KRAS* et la méthylation de *MGMT*. Ces données confirment celles obtenues sur notre série de 104 carcinomes invasifs de phénotype non MSI et seront complétées sur cette série qui compte maintenant 343 cas de carcinomes invasifs correctement annotés en terme d'anomalies moléculaires et de suivi clinique.

Dans notre travail, certaines de ces anomalies moléculaires ont une distribution différente selon la localisation proximale et distale, suggérant, comme dans plusieurs études rapportées dans la littérature, des voies de carcinogenèse distinctes entre les adénomes proximaux et distaux.

Enfin de façon tout-à-fait intéressante, nous avons pu observer une corrélation entre un panel d'anomalies moléculaires et des caractéristiques histologiques des adénomes, nécessitant des investigations complémentaires en vue de leur confirmation.

Ces données permettent de proposer des pistes d'explication concernant le concept d'hétérogénéité tumorale du cancer colique de phénotype non MSI.

L'absence de facteurs moléculaires prédictifs utilisables à l'heure actuelle en routine dans les carcinomes coliques en est certainement une conséquence. Cela montre l'importance de ces travaux pour la compréhension des voies de carcinogenèse colique et la nécessité de déterminer, dans le future et pour chaque patient, en parallèle aux données histologiques classiques déterminant le type d'adénomes et le stade pTNM des tumeurs, un profil moléculaire spécifique qui pourrait constituer une aide au diagnostic et à l'établissement du pronostic des adénomes et qui permettrait pour les carcinomes invasifs d'adapter le traitement des patients en fonction des cibles thérapeutiques disponibles.

REFERENCES BIBLIOGRAPHIQUES

A

Aarnio M, Mecklin JP, et al. (1995). Life-time risk of different cancers in hereditary non-polyposis colorectal cancer (HNPCC) syndrome. Int J Cancer 64(6): 430-3.

Alemayehu A, Sebova K, et al. (2008). Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. Genes Chromosomes Cancer 47(10):906-14.

Ali MF, Meza JL, et al. (2008). Prevalence of BER gene polymorphisms in sporadic breast cancer. Oncol Rep 19(4):1033-8.

Allan LA, Morrice N, et al. (2003). Inhibition of caspase-9 through phosphorylation at Thr 125 by ERK MAPK. Nat Cell Biol 5(7): 647-54.

B

Bakin AV and Curran T (1999). Role of DNA 5-methylcytosine transferase in cell transformation by fos. Science 283(5400): 387-90.

Belot A, Grosclaude P, et al. (2008). Cancer incidence and mortality in France over the period 1980-2005. Rev Epidemiol Sante Publique 56(3): 159-75.

Behrens J, von Kries JP, et al. (1996). Functional interaction of beta-catenin with the transcription factor LEF-1. Nature 382(6592): 638-42.

Bellam N, Pasche B. (2010). Tgf-beta signaling alterations and colon cancer. Cancer Treat Res 155:85-103.

Birkenkamp-Demtroder K, Christensen LL et al. (2002). Gene expression in colorectal cancer. Cancer Res 1;62(15):4352-63.

Bleeker WA, Hayes VM, et al. (2000). Impact of KRAS and TP53 mutations on survival in patients with left- and right-sided Dukes' C colon cancer. Am J Gastroenterol 95(10):2953-7.

Boland CR, Thibodeau SN, et al. (1998). A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 58(22): 5248-57.

Bonithon-Kopp C, Piard F, et al. (2004). Colorectal adenoma characteristics as predictors of recurrence. Dis Colon Rectum 47:323-33.

Bouvard V, Zaitchouk T, et al. (2000). Tissue and cell-specific expression of the p53-target genes: bax, fas, mdm2 and waf1/p21, before and following ionising irradiation in mice. Oncogene 19(5): 649-60.

C

Carvajal-Carmona LG, Howarth KM, et al. (2007). Molecular classification and genetic pathways in hyperplastic polyposis syndrome. J Pathol 7;212:378-85.

Cecchi F, Rabe DC, et al. (2010). Targeting the HGF/Met signalling pathway in cancer. Eur J Cancer 46(7):1260-70.

Chan TA, Hermeking H, et al. (1999). "14-3-3Sigma is required to prevent mitotic catastrophe after DNA damage." Nature 401(6753):616-20.

Choi DH, Sohn DK, et al. (2009). Indications for subsequent surgery after endoscopic resection of submucosally invasive colorectal carcinomas: a prospective cohort study. Dis Colon Rectum 52(3):438-45.

Chu D, Wang W, et al. (2009). Notch1 expression in colorectal carcinoma determines tumor differentiation status. J Gastrointest Surg 13:253-60.

Consensus conference. (1998). Prevention, diagnosis and treatment of colon cancer. Gastroenterol Clin Biol 22(2):205-26.

Croner RS, Peters A, et al. (2005). Microarray versus conventional prediction of lymph node metastasis in colorectal carcinoma. Cancer 15;104(2):395-404.

D

Dalle S, Martin-Denavit T, et al. (2006). [Genotypic hypervariability of melanoma: a therapeutic challenge]. Med Sci (Paris) 22(2): 178-82.

Deltour S, Chopin V, et al. (2005). [Epigenetics and cancer]. Med Sci (Paris) 21(4): 405-11.

Denis B, Peters C, et al. (2009). Diagnostic accuracy of community pathologists in the interpretation of colorectal polyps. Eur J Gastroenterol Hepatol 21(10):1153-60.

Derks S, Postma C, et al. (2006). Promoter methylation precedes chromosomal alterations in colorectal cancer development. Cell Oncol 28(5-6): 247-57.

Di Vinci A, Infusini E, et al. (1996). Deletions at chromosome 1p by fluorescence in situ hybridization are an early event in human colorectal tumorigenesis. Gastroenterology 111(1):102-7.

Dib C, Faure S, et al. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380(6570):152-4.

Diep CB, Kleivi K, et al. (2006). The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. Genes Chromosomes Cancer 45(1): 31-41.

Dikovskaya D, Li Z, et al. (2010). Microtubule assembly by the Apc protein is regulated by importin-beta--RanGTP. J Cell Sci 1;123(Pt 5):736-46.

Dixon MF. (2002). Gastrointestinal epithelial neoplasia: Vienna revisited. Gut 51(1):130-1.

Duffy MJ, van Dalen A, et al. (2007). Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use. Eur J Cancer 43(9): 1348-60.

E

el-Deiry WS, Tokino T, et al. (1993). WAF1, a potential mediator of p53 tumor suppression. Cell 75(4): 817-25.

Eshleman JR, Lang EZ, et al. (1995). Increased mutation rate at the hprt locus accompanies microsatellite instability in colon cancer. Oncogene 10(1):33-7.

Esteller M, Hamilton SR, et al. (1999). Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res 59(4):793-797.

Esteller M, Gaidano G, et al. (2002). Hypermethylation of the DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients with diffuse large B-cell lymphoma. J Natl Cancer Inst 2;94(1):26-32.

Esteller M, Herman JG. (2004). Generating mutations but providing chemosensitivity: the role of O6-methylguanine DNA methyltransferase in human cancer. Oncogene 8:1-8.

F

Farris AB, Misdraji J, et al. (2008). Sessile serrated adenoma: challenging discrimination from other serrated colonic polyps. Am J Surg Pathol 32(1):30-5.

Fearnhead NS, Britton MP, et al. (2001). The ABC of *APC*. Hum Mol Genet 10(7):721-33.

Fearon E. R. and Vogelstein B. (1990). A genetic model for colorectal tumorigenesis. Cell 61(5):759-67.

Fre S, Pallavi SK, et al. (2009). Notch and Wnt signals cooperatively control cell proliferation and tumorigenesis in the intestine. Proc Natl Acad Sci U S A 4;106(15):6309-14.

G

Gatta G, Capocaccia R, et al. (2000). Understanding variations in survival for colorectal cancer in Europe: a EURO CARE high resolution study. Gut 47(4): 533-8.

Gervaz P, Bouzourene H, et al (2001). Dukes B colorectal cancer: distinct genetic categories and clinical outcome based on proximal or distal tumor location. Dis Colon Rectum 44(3):364-72.

Giardiello FM, Welsh SB, et al. (1987). Increased risk of cancer in Peutz-Jeghers syndrome. N Engl J Med 316:1511-4.

Goffin J and Eisenhauer E (2002). DNA methyltransferase inhibitors-state of the art. Ann Oncol 13(11): 1699-716.

Goyette MC, Cho K, et al. (1992). Progression of colorectal cancer is associated with multiple tumor suppressor gene defects but inhibition of tumorigenicity is accomplished by correction of any single defect via chromosome transfer. Mol Cell Biol 12(3): 1387-95.

Grady WM, Myeroff LL, et al. (1999). Mutational inactivation of transforming growth factor beta receptor type II in microsatellite stable colon cancers. Cancer Res 59(2): 320-4.

Grady WM and Carethers JM (2008). Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 135(4): 1079-99.

Gregorieff A and Clevers H (2005). Wnt signaling in the intestinal epithelium: from endoderm to cancer. Genes Dev 19(8): 877-90.

Groden J, Thliveris A, et al. (1991). Identification and characterization of the familial adenomatous polyposis coli gene. Cell 66(3): 589-600.

Gross CP, Andersen MS, et al. (2006). Relation between Medicare screening reimbursement and stage at diagnosis for older patients with colon cancer. Jama 296(23): 2815-22.

Gryfe R, Swallow C, et al. (1997). Molecular biology of colorectal cancer. Curr Probl Cancer 21(5):233-300.

Gschwantler M, Kriwanek S, et al. (2002). High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics. Eur J Gastroenterol Hepatol 14(2): 183-8.

Gupta S. (2008). Colorectal polyps: the scope and management of the problem. Am J Med Sci 336:407-17.

H

Haggitt RC, Glotzbach RE, et al. (1985). Prognostic factors in colorectal carcinomas arising in adenomas: implications for lesions removed by endoscopic polypectomy. Gastroenterology 89(2):328-36.

Hamelin R and Duval A (2003). [Beta-catenin and cell proliferation in normal and transformed intestinal cells]. Med Sci (Paris) 19(8-9): 788-90.

Harper JW, Adami GR, et al. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 75(4): 805-16.

Heldin CH, Miyazono K, et al. (1997). TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature 390(6659): 465-71.

Hermeking H, Lengauer C, et al. (1997). 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. Mol Cell 1(1): 3-11.

Hermesen M, Postma C, et al. (2002). Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. Gastroenterology; 123:1109-19.

Herrera L, Kakati S, et al. (1986). Gardner syndrome in a man with an interstitial deletion of 5q. Am J Med Genet 25(3): 473-6.

I

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

J

Jass JR. (2006). Colorectal cancer: a multipathway disease. Crit Rev Oncog 12:273-87.

Jass JR. (2007). Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 50(1):113-30.

Jass JR, Iino H, et al. (2000). Neoplastic progression occurs through mutator pathways in hyperplastic polyposis of the colorectum. Gut 47(1):43-9.

Jass JR, Baker K, et al. (2006). Advanced colorectal polyps with the molecular and morphological features of serrated polyps and adenomas: concept of a 'fusion' pathway to colorectal cancer. Histopathology 49(2):121-31.

Jin S, Tong T, et al. (2002). GADD45-induced cell cycle G2-M arrest associates with altered subcellular distribution of cyclin B1 and is independent of p38 kinase activity. Oncogene 21(57): 8696-704.

Joerger AC and Fersht AR (2008). Structural biology of the tumor suppressor p53. Annu Rev Biochem 77: 557-82.

Jones PA, Baylin SB (2002). The fundamental role of epigenetic events in cancer. Nature Reviews Genetics 3(6), 415-428.

Jones AM, Douglas EJ, et al (2005). Array-CGH analysis of microsatellite-stable, near-diploid bowel cancers and comparison with other types of colorectal carcinoma. Oncogene 6;24(1):118-29.

Jones S, Chen WD, et al. (2008). Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci U S A 18;105(11):4283-8.

Joshi AD, Corral R, et al. (2009). Red meat and poultry intake, polymorphisms in the nucleotide excision repair and mismatch repair pathways and colorectal cancer risk. Carcinogenesis 30(3):472-9.

Judson H, Stewart A, et al. (2006). Relationship between point gene mutation, chromosomal abnormality, and tumour suppressor gene methylation status in colorectal adenomas. J Pathol 210(3): 344-50.

K

Kaina B, Christmann M, et al. (2007). MGMT: Key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. DNA repair 6:1079-99

Kastan MB, Canman CE, et al. (1995). P53, cell cycle control and apoptosis: implications for cancer. Cancer Metastasis Rev 14(1): 3-15.

Kawasaki T, Ohnishi M, et al. (2008). CpG island methylator phenotype-low (CIMP-low) colorectal cancer shows not only few methylated CIMP-high-specific CpG islands, but also low-level methylation at individual loci. Mod Pathol 21(3):245-255.

Kikuchi R, Takano M, et al. (1995). Management of early invasive colorectal cancer: risk of recurrence and clinical guidelines. Dis Colon Rectum 38:1286-1295.

Kinzler KW, Vogelstein B (1996). Lessons from hereditary colorectal cancer. Cell. 87(2), 159-170.

Komuro K, Tada M, et al. (2005). Right- and left-sided colorectal cancers display distinct expression profiles and the anatomical stratification allows a high accuracy prediction of lymph node metastasis. J Surg Res 124(2):216-24.

Krtolica K, Krajnovic M, et al. (2007). Comethylation of p16 and MGMT genes in colorectal carcinoma: correlation with clinicopathological features and prognostic value. World J Gastroenterol 13:1187-94

L

Laurent-Puig P, Blons H, et al. (1999). Sequence of molecular genetic events in colorectal tumorigenesis. Eur J Cancer Prev 8 Suppl 1: S39-47.

Lengauer C, Kinzler KW, et al. (1998). Genetic instabilities in human cancers. Nature 396(6712): 643-9.

Leslie A, Pratt NR, et al. (2003). Mutations of *APC*, *KRAS*, and p53 are associated with specific chromosomal aberrations in colorectal adenocarcinomas. Cancer Res 63(15): 4656-61.

Levin B, Lieberman DA, et al. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 134(5): 1570-95.

Levine AJ (1997). p53, the cellular gatekeeper for growth and division. Cell 88(3): 323-31.

Levy A, Sauvin LA, et al. (2008). [Everolimus (RAD001) and solid tumours: a 2008 summary]. Bull Cancer 95(12): 1205-11.

Lim W, Hearle N, et al. (2003). Further observations on LKB1/STK11 status and cancer risk in Peutz-Jeghers syndrome. Br J Cancer 21;89(2):308-13.

Liu X, Yue P, et al. (2004). p53 upregulates death receptor 4 expression through an intronic p53 binding site. Cancer Res 64(15): 5078-83.

Longacre TA, Fenoglio-Preiser CM. (1990). Mixed hyperplastic adenomatous polyps/serrated adenomas. A distinct form of colorectal neoplasia. Am J Surg Pathol 14(6):524-37.

Lowe SW, Cepero E, et al. (2004). Intrinsic tumour suppression. Nature 432(7015): 307-15.

Lynch HT, Smyrk T, et al. (1996). Overview of natural history, pathology, molecular genetics and management of HNPCC (Lynch syndrome). Int J Cancer 69 : 38-43.

M

Marsit CJ, Houseman EA, et al. (2006). Examination of a CpG island methylator phenotype and implications of methylation profiles in solid tumors. Cancer Res 66(21): 10621-9.

Martínez ME, Baron JA, et al. (2009). A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. Gastroenterology 136(3):832-41.

McCubrey JA, Steelman LS, et al. (2006). Roles of the RAF/MEK/ERK and PI3K/PTEN/AKT pathways in malignant transformation and drug resistance. Adv Enzyme Regul 46: 249-79.

Michalak E, Villunger A, et al. (2005). Death squads enlisted by the tumour suppressor p53. Biochem Biophys Res Commun 331(3): 786-98.

Modrich P, Lahue R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem 65 : 101-33.

N

Nagasaka T, Goel A, et al. (2008). Methylation pattern of the O6-methylguanine-DNA methyltransferase gene in colon during progressive colorectal tumorigenesis. Int J Cancer 122:2429-36

Nagasaka T, Koi M, et al (2008). Mutations in both KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 134(7):1950–1960.

Nelen MR, Padberg GW, et al. (1996). Localization of the gene for Cowden disease to 10q22-23. Nat Genet 13:114-16.

Nomura M, Shimizu S, et al. (2003). 14-3-3 Interacts directly with and negatively regulates pro-apoptotic Bax. J Biol Chem 278(3): 2058-65.

Nosho K, Shima K, et al. (2009). DNMT3B expression might contribute to CpG island methylator phenotype in colorectal cancer. Clin Cancer Res 15(11):3663-71.

O

O'Brien MJ, Yang S, et al. (2006). Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. Am J Surg Pathol 30:1491–1501.

Ogawara Y, Kishishita S, et al. (2002). Akt enhances Mdm2-mediated ubiquitination and degradation of p53. J Biol Chem 277(24): 21843-50.

Ogino S, Kawasaki T, et al. (2007). TGFBR2 mutation is correlated with CpG island methylator phenotype in microsatellite instability-high colorectal cancer. Hum Pathol 38(4): 614-20.

Ogino S, Nosho K, et al. (2009). Lymphocytic reaction to colorectal cancer is associated with longer survival, independent of lymph node count, microsatellite instability, and CpG island methylator phenotype. Clin Cancer Res 15(20):6412-20.

Olschwang S (2001). [Digestive polyposis: genetic aspects]. Gastroenterol Clin Biol 25(4 Suppl): B26-30.

Olschwang S, Bonaïti C, et al. (2004). Identification et prise en charge du syndrome HNPCC (hereditary non polyposis colon cancer), Prédilection héréditaire aux cancers du côlon, du rectum et de l'utérus. Bull Cancer 91(4):303-15.

Ostwald C, Linnebacher M, et al. (2009). Chromosomally and microsatellite stable colorectal carcinomas without the CpG island methylator phenotype in a molecular classification. Int J Oncol 35(2):321-7.

P

Park DH, Kim HS, et al (2008). Clinicopathologic characteristics and malignant potential of colorectal flat neoplasia compared with that of polypoid neoplasia. Dis Colon Rectum 51(1):43-9.

Peltomaki P (2005). Lynch syndrome genes. Fam Cancer 4(3):227-32.

Peltomaki P and Vasen HF (1997). Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology 113(4): 1146-58.

Phelps RA, Chidester S, et al. (2009). A two-step model for colon adenoma initiation and progression caused by APC loss. Cell 15(137(4):623-34.

Popat S, Hubner R, et al. (2005). Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 23(3): 609-18.

Powell SM, Zilz N, et al. (1992). APC mutations occur early during colorectal tumorigenesis. Nature 359(6392): 235-7.

Pretlow TP, Barrow JB, et al. (1991). Aberrant crypts: putative preneoplastic foci in human colonic mucosa. Cancer Res 51(5): 1564-7.

Pretlow TP and Pretlow TG (2005). Mutant KRAS in aberrant crypt foci (ACF): initiation of colorectal cancer? Biochim Biophys Acta 1756(2): 83-96.

R

Reedijk M, Odorcic S, et al. (2008). Activation of Notch signaling in human colon adenocarcinoma. Int J Oncol 33(6):1223-9.

Ried T, Knutzen R, et al. (1996). Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. Genes Chromosomes Cancer 15(4): 234-45.

Roberts PJ and Der CJ (2007). Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene 26(22): 3291-310.

Roman-Gomez J, Jimenez-Velasco A, et al. (2005). Lack of CpG island methylator phenotype defines a clinical subtype of T-cell acute lymphoblastic leukemia associated with good prognosis. J Clin Oncol 23(28): 7043-9.

Rosenberg DW, Yang S, et al. (2007). Mutations in BRAF and KRAS differentially distinguish serrated versus non-serrated hyperplastic aberrant crypt foci in humans. Cancer Res 67:3551–3554.

Russo A, Bazan V, et al. (2005). The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment. J Clin Oncol 20;23(30):7518-28.

S

Samuels Y and Ericson K (2006). Oncogenic PI3K and its role in cancer. Curr Opin Oncol 18(1): 77-82.

Samowitz WS, Powers MD, et al. (1999). Beta-catenin mutations are more frequent in small colorectal adenomas than in larger adenomas and invasive carcinomas. Cancer Res 59(7): 1442-4.

Sampson JR, Dolwani S, et al. (2003). Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. Lancet 362(9377): 39-41.

Sawyer EJ, A Cerar A et al. (2002). Molecular characteristics of serrated adenomas of the colorectum. Gut 51(2): 200–206.

Schlemper RJ, Riddell RH, et al. (2000). The Vienna classification of gastrointestinal epithelial neoplasia. Gut 47(2):251-5.

Schmierer B and Hill CS (2007). TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. Nat Rev Mol Cell Biol 8(12): 970-82.

Sheikh MS, Burns TF, et al. (1998). p53-dependent and -independent regulation of the death receptor KILLER/DR5 gene expression in response to genotoxic stress and tumor necrosis factor alpha. Cancer Res 58(8): 1593-8.

Shen L, Ahuja N, et al. (2002). DNA methylation and environmental exposures in human hepatocellular carcinoma. J Natl Cancer Inst 94(10): 755-61.

Shen L, Kondo Y, et al. (2005). J Natl Cancer Inst 21;97(18):1330-8.

- Shen L, Toyota M, et al. (2007). Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci USA;104(47):18654–18659.
- Shimizu Y, Ikeda S, et al. (2002). Frequent alterations in the Wnt signaling pathway in colorectal cancer with microsatellite instability. Genes Chromosomes Cancer 33(1):73-81.
- Smith AJ, Stern HS, et al. (1994). Somatic *APC* and *KRAS* codon 12 mutations in aberrant crypt foci from human colons. Cancer Res 54(21): 5527-30.
- Smith G, Carey FA, et al. (2002). Mutations in *APC*, *Kirsten-ras*, and *p53*—Alternative genetic pathways to colorectal cancer. Proceedings of the National Academy of Sciences of the United States of America 99(14), 9433–9438.
- Snover DC, Jass JR, et al. (2005). Serrated polyps of the large intestine: a morphologic and molecular review of an evolving concept. Am J Clin Pathol 124(3):380-91.
- Soreide K, Nedrebo BS, et al. (2009). Endoscopy, morphology, morphometry and molecular markers: predicting cancer risk in colorectal adenoma. Expert Rev Mol Diagn 9(2):125-37.
- Spirio LN, Samowitz W, et al. (1998). Alleles of *APC* modulate the frequency and classes of mutations that lead to colon polyps. Nat Genet 20(4): 385-8.
- Spjut H, Estrada RG (1977). The significance of epithelial polyps of the large bowel. Pathol Annu 12(Pt 1):147–170.
- Spring KJ, Zhao ZZ, et al. (2006). High prevalence of sessile serrated adenomas with *BRAF* mutations: a prospective study of patients undergoing colonoscopy. Gastroenterology 131:1400–1407.
- Stambolic V, MacPherson D, et al. (2001). Regulation of *PTEN* transcription by *p53*. Mol Cell 8(2): 317-25.
- Strathdee G, Appleton K, et al. (2001). Primary ovarian carcinomas display multiple methylator phenotypes involving known tumor suppressor genes. Am J Pathol 158(3): 1121-7.
- Stryker SJ, Wolff BG, et al. (1987). Natural history of untreated colonic polyps. Gastroenterology 93(5): 1009-13.
- Svreck M, Buhard O, et al. (2010). Methylation tolerance due to O6-methylguanine DNA methyltransferase (*MGMT*) field defect in the colonic mucosa: an initiating step in the development on mismatch repair deficient colorectal cancers. In press.
- Sy J, Fung CL, et al. (2010). Tumor budding and survival after potentially curative resection of node-positive colon cancer. Dis Colon Rectum 53(3):301-7.

T

- Teniere P, Songne K, et al. (2002). [Juvenile polyposis coli. The usefulness of a genetic study and the role of surgical treatment]. Gastroenterol Clin Biol 26(11): 1047-50.
- Thorstensen L, Qvist H, et al. (2000). Evaluation of 1p losses in primary carcinomas, local recurrences and peripheral metastases from colorectal cancer patients. Neoplasia 2(6):514-22.
- Thorstensen L, Lind GE, et al. (2005). Genetic and epigenetic changes of components affecting the *WNT* pathway in colorectal carcinomas stratified by microsatellite instability. Neoplasia 7(2):99-108.
- Torlakovic E, Snover DC. (1996). Serrated adenomatous polyposis in humans. Gastroenterology 110(3):748-55.

Torlakovic E, Skovlund E, et al. (2003). Am J Surg Pathol 27(1):65-81.

Toyota M, Ahuja N, et al. (1999). Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. Cancer Res 59(21): 5438-42.

Toyota M, Ahuja N, et al. (1999). CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci USA 96:8681–8686.

Tsafir D, Bacolod M, et al. (2006). Relationship of gene expression and chromosomal abnormalities in colorectal cancer. Cancer Res 15;66(4):2129-37.

U

Ueno H, Mochizuki H, et al. (2004). "Risk factors for an adverse outcome in early invasive colorectal carcinoma." Gastroenterology 127(2): 385-94.

Ueno H, Price AB, et al. (2004). A new prognostic staging system for rectal cancer. Ann Surg 240(5):832-9.

V

Valle L, Serena-Acedo T, et al. (2008). Germline allele-specific expression of TGF β R1 confers an increased risk of colorectal cancer. Science 321:1361-5.

van Krieken JH and van de Velde C (2002). Past, present and future perspectives of colorectal cancer--on the brink of a new era? Eur J Cancer 38(7): 855-7.

van Lier MG, Wagner A, et al. (2010). High cancer risk in Peutz-Jeghers syndrome: a systematic review and surveillance recommendations. Am J Gastroenterol 105(6):1258-64.

Van Rijnsoever M, Elsaleh H, et al. (2003). CpG island methylator phenotype is an independent predictor of survival benefit from 5-fluorouracil in stage III colorectal cancer. Clin Cancer Res 9(8): 2898-903.

Verdecchia A, Francisci S, et al. (2007). "Recent cancer survival in Europe: a 2000-02 period analysis of EURO-CARE-4 data." Lancet Oncol 8(9): 784-96.

Vousden KH and Lu X (2002). Live or let die: the cell's response to p53. Nat Rev Cancer 2(8): 594-604.

W

Walther A, Johnstone E, et al. (2009). Genetic prognostic and predictive markers in colorectal cancer. Nat Rev Cancer 9(7):489-99.

Weber JC, Meyer N, et al. (2007). Allelotyping analyses of synchronous primary and metastasis CIN colon cancers identified different subtypes. Int J Cancer 120(3): 524-32.

Weisenberger DJ, Siegmund KD, et al. (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 38(7): 787-93.

Williams NS, Gaynor RB, et al. (2003). Identification and validation of genes involved in the pathogenesis of colorectal cancer using cDNA microarrays and RNA interference. Clin Cancer Res 9(3):931-46.

Willis SN and Adams JM (2005). Life in the balance: how BH3-only proteins induce apoptosis. Curr Opin Cell Biol 17(6): 617-25.

Winawer SJ, Zauber AG, et al. (2006). Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. Gastroenterology 130:1872-85.

Wiseman A (2006). p53 protein or BID protein select the route to either apoptosis (programmed cell death) or to cell cycle arrest opposing carcinogenesis after DNA damage by ROS. Med Hypotheses 67(2): 296-9.

Wong JJJ, Hawkins NJ, et al. (2007). Colorectal cancer: a model for epigenetic tumorigenesis. Gut 56:140-148.

Wood LD, Parsons DW et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science 318:1108-13.

Wu Y, Berends JM, et al. (2001). Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. Gastroenterology 120(7): 1580-7.

Z

Zavadil J and Bottinger EP (2005). TGF-beta and epithelial-to-mesenchymal transitions. Oncogene 24(37): 5764-74.

Zeng ZS, Weiser MR, et al. (2008). c-Met gene amplification is associated with advanced stage colorectal cancer and liver metastases. Cancer Lett. Jul 8;265(2):258-69.

Zhou BP, Liao Y, et al. (2001). Cytoplasmic localization of p21Cip1/WAF1 by Akt-induced phosphorylation in HER-2/neu-overexpressing cells. Nat Cell Biol 3(3): 245-52.

ANNEXES

Publication 1

Evidence for various 20q status using allelotyping, CGH arrays, and quantitative PCR in distal CIN colon cancers.

Nicolet C, Guérin E, Neuville A, Kerckaert JP, Wicker N, Bergmann E,
Brigand C, Kedinger M, Gaub MP, Guenot D.

Cancer Lett. 2009 Sep 18;282(2):195-204.



Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Evidence for various 20q status using allelotyping, CGH arrays, and quantitative PCR in distal CIN colon cancers

Céline Nicolet^a, Eric Guérin^{a,b}, Agnès Neuville^{a,c}, Jean-Pierre Kerckaert^d, Nicolas Wicker^e, Etienne Bergmann^c, Cécile Brigand^f, Michèle Kedinger^a, Marie-Pierre Gaub^{a,b}, Dominique Guenet^{a,*}

^aInserm, U682, 3 Avenue Motière Strasbourg, F-67200 France; UMR-S682, Univ Louis Pasteur, Faculté de Médecine, Strasbourg, F-67200, France

^bLaboratoire de Biologie Moléculaire et Biochimie, CHRU Strasbourg, Hôpital de Hautepierre, Strasbourg F-67200, France

^cService d'anatomo-pathologie, CHRU Strasbourg, Hôpital de Hautepierre, Strasbourg F-67200, France

^dPlateforme de Génomique Fonctionnelle, Université de Lille II, F-59045 Lille, France

^eLaboratoire de Bioinformatique et Génomique Intégratives, IGBMC, Illkirch F-67000, France

^fService de chirurgie digestive, CHRU Strasbourg, Hôpital de Hautepierre, Strasbourg F-67200, France

ARTICLE INFO

Article history:

Received 3 October 2008

Received in revised form 13 February 2009

Accepted 9 March 2009

Available online xxx

Keywords:

Distal colon cancer

Chromosomal instability

Chromosome 20q

Allelotyping

CGH array

Tetrasomy

Aneusomy

ABSTRACT

The genomic aberration profile of chromosome 20q in distal CIN colon carcinomas was analysed using allelotyping and CGH arrays. Allelotyping revealed carcinomas with allelic imbalance along the full long arm, and carcinomas with fully non-aberrant 20q. Oligonucleotide-based CGH showed that among the carcinomas without allelic imbalance, 47% had in fact a gain. In this subgroup, quantitative PCR for the TOPI gene (20q12) confirmed this gain, and fluorescence in situ hybridization showed that the chromosome 20q gain resulted from tetra/polysomy instead of aneusomy. The 20q gain correlated with a high frequency of aberrations, with allelic imbalance at TP53 locus but not at APC locus, and carcinomas with a disomic 20q showed low frequency of genomic aberrations and were significantly associated to mucinous phenotype. The prognostic value of 20q amplification was not demonstrated in this study. These results indicate that on the basis of aberration frequency, chromosome 20q and TP53/APC locus status, distal CIN carcinomas harbor a high degree of genetic heterogeneity suggesting several pathways for carcinogenesis. This study also indicates that allelotyping needs to be carried out with a complementary technique, such as quantitative PCR.

© 2009 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The transition from adenoma to carcinoma that underlies colon carcinogenesis results from a multi-step process of genomic events that promote activation of proto-oncogenes or inactivation of tumor suppressor genes [1,2]. Although the genetic basis for colon cancer has been increasingly well characterized in recent years, reliable

genomic markers for progression to malignant stages or for resistance to treatment have yet to be described. This may be due to the fact that previous studies were carried out on cohorts whose carcinoma localisation and phenotype varied. Indeed, several studies have suggested that carcinomas from either proximal or distal site, rectal cancers [3], as well as cancers involving microsatellite (MIN) and chromosome unstable (CIN) phenotypes [4], have different properties and consequently should be analysed separately.

The adenoma to carcinoma transition is characterized by the emergence of multiple chromosomal aberrations, including amplification of the chromosome 20q arm [5].

* Corresponding author. Tel.: +33 03 88 27 53 62; fax: +33 03 88 26 35 38.

E-mail address: dominique.guenot@inserm.u-strasbg.fr (D. Guenet).

Overrepresentation of chromosome 20q can contribute to the development of malignancy in the colon [5,6], through the presence of dominant-acting growth regulatory genes assigned to this chromosome [7]. Some authors have reported homogeneous amplification of the 20q arm along its entire length [8,9], and others have also found amplification of specific regions, particularly in subregions 20q13.2 [10] and 20q13.3 [11]. These amplified segments correlate with carcinoma progression and metastatic development in some studies [6,10], but not all of them [12]. This discrepancy may reflect the fact that these studies are sometimes based on cohorts whose carcinoma localisation and phenotype are heterogeneous.

The present report gives a detailed analysis of chromosome 20q status in a collection of carcinomas restricted to distal colon cancer of the CIN phenotype. We used high-resolution allelotyping and CGH arrays (aCGH), as well as real-time quantitative PCR (QPCR) of the TOP1 gene to corroborate the results. These combined approaches demonstrated that 82% of the cases had amplification along the entire length of the chromosome 20q arm but that among these carcinomas, aCGH and QPCR of the TOP1 gene detected a subgroup with amplification but without AI. In these carcinomas, FISH revealed a tetrasomic 20q. The remaining 18% of cases had unaltered disomic chromosome 20q and a mucinous phenotype. These observations lead us to propose that human colon carcinogenesis proceeds via several pathways and among which, pathways that differentiate, among other alterations, the status of chromosome 20q. Moreover, acquisition of tetrasomy rather occurs at the adenoma stage with no further increase while aneusomy is essentially acquired during adenoma to carcinoma transition.

2. Materials and methods

2.1. Tissue collection

A total of 66 human distal adenomas, 223 CIN primary colon carcinomas and matching normal mucosa 5 cm distant from the lesions, were selected on a random basis in a cohort from which proximal, rectal, and MIN carcinomas

have been excluded. Among this cohort, 60 carcinomas were part of the cohort analysed previously [13]. Samples were obtained from patients who underwent surgery at the University Hospital of Strasbourg between 1997 and 2006, and who had been recruited into the study after it had undergone institutional review by the French Ethics Committee of Strasbourg. Genomic DNA was extracted from samples that had been stored in liquid nitrogen using the QIAamp DNA mini kit (Qiagen, Courtaboeuf, France). The anatomo-pathological characteristics are described in Table 1. None of the patients had received preoperative chemotherapy. Clinical staging and percentage of tumor cells were assessed in each sample by an experienced histopathologist. For aCGH, only carcinoma fragments containing more than 40% of cancer cells were analysed, according to the manufacturer recommendation (Agilent Technologies, Santa Clara, CA).

For the event-free survival analysis, only patients who had at least a 36-month follow-up period were considered ($n = 120$; see Table 1).

The different groups were tested with Pearson's Chi² test or Fisher's exact test. The Kaplan–Meier method was performed using the log-rank test. Only p values lower than 0.05 were regarded as statistically significant. Statistical analyses were performed using SPSS software (SPSS for Windows, Version 11.5, Chicago, IL).

2.2. Microsatellite analyses

2.2.1. Overall allelotyping

DNA extracted from the 66 adenomas and 223 carcinomas was amplified by fluorescent PCR as described previously [14]. A panel of 33 microsatellites (MS) targeting 18 chromosomes were chosen, since they correspond to the loci most frequently rearranged in colon cancer [15,16]. These MS allowed, among others, determination of the status of APC (5q, D5S346), TP53 (17p), and TOP1 loci (20q, D20S855). MS and chromosomal locations are described in a previous study [17]. Primers were obtained from the Genome Data Base (<http://www.gdb.org>) or Genemap'99 (<http://www.ncbi.nlm.nih.gov/genemap99/>). Our assay works by detecting a change in the allele ratio

Table 1

Anatomo-pathological characteristics. The 39 carcinomas analysed by aCGH and the 74 carcinomas analysed by allelotyping constitute subgroups of the cohort of 223 distal CIN colon carcinomas. The event-free survival analysis was performed on 120 patients among the cohort of 223 carcinomas with 36 months of clinical follow-up.

Carcinomas	$n = 223$ Allelotyping – QPCR	$n = 120$ Clinical follow-up	$n = 74$ 20q allelotyping	$n = 39$ aCGH	Adenomas	$n = 66$ Allelotyping – QPCR
Mean age (Range)	66 (41–92)	66 (42–89)	67 (46–86)	67 (46–82)	Mean age (Range)	66 (44–88)
Male:Femelle	131:92	73:47	42:32	25:14	Male:Femelle	41:19
<i>IUCC Classification</i>					<i>Histology</i>	
0	6 (2.5%)	1 (0 ^a)	2	2	Tubulovillous/low-grade	39 (59%)
I	25 (11%)	18 (6 ^a)	12	6	Tubulovillous/low-grade	15 (23%)
II	62 (28%)	33 (11 ^a)	22	11	Tubulovillous/high-grade	3 (4.5%)
III	48 (21.5%)	25 (6 ^a)	15	7	Tubulovillous/high-grade	2 (3%)
IV	82 (37%)	43 (40 ^a)	23	13	Villous/low-grade	1 (1.5%)
Differentiated	174 (78%)	98 (82%)	60 (81%)	31 (79.5%)	Hyperplasia	6 (9%)
Mucinous	36 (16%)	18 (15%)	10 (13.5%)	6 (15.5%)		
Poorly differentiated	13 (6%)	4 (3%)	4 (5.5%)	2 (5%)		

^a Patients with clinical events (metastasis – relapse – death related to colon cancer).

in carcinoma relative to the allele ratio in paired normal blood cells and normal mucosa [17]. A modification of the allele ratio in carcinoma relative to the matched control is referred to as AI, and the presence of an additional peak is referred to as MSI. A cut-off value of the intensity of the AI of 20% or greater for a given MS indicates significant AI [14]. In order to exclude carcinomas with an MSI-H phenotype, the panel of five MS recommended by the U.S. National Cancer Institute (BAT25, BAT26, D5S346, D2S123, and D17S250) was included in the analysis.

2.2.2. Allelotyping focused on chromosome 20q

Within the cohort of 223 CIN carcinomas, a subgroup of 74 was selected to be representative of all clinical stages, with or without allelic imbalance at TOP1 locus (D20S855). This subgroup was subjected to focused, high-density allelotyping of chromosome 20q. This specific study was performed by multiplexed PCR amplification of 32 highly polymorphic MS markers located along the long arm of chromosome 20, which has a total length of 33 Mb. This corresponds to an average resolution of 1 Mb (Table 2). Information regarding primer sequences and MS locations were collected by data-mining the GDB Human Genome Database (<http://www.gdb.org>) and Genemap'99 (<http://www.ncbi.nlm.nih.gov/genemap99>).

For both types of allelotyping, multiplex PCR amplification products were separated on an ABI PRISM 3100 genetic analyser (Applied Biosystems). Data analysis involved a two-way unsupervised clustering approach.

Table 2

Polymorphic microsatellite markers used to detect AI over the entire length of chromosome 20q. ID: marker identification name.

Marker	Location	ID	Location (Mb)
Q7	20q11.2	D20S874	30.2
Q8	20q11.21	D20S843	30.7
Q9	20q11.21	D20S195	31.2
Q10	20q11.22	D20S890	31.6
Q11	20q11.22	D20S106	32.8
Q19	20q11.23	D20S909	33.9
Q21	20q11.23	D20S847	34.3
Q1	20q11.23	D20S884	35.5
Q12	20q11.23	D20S859	35.9
F4	20q12	D20S107	38.3
Q5	20q12	D20S855	39.1
W4	20q12	D20S170	39.9
P1	20q12	D20S108	40.2
Q13	20q13.11	D20S899	41.2
Q23	20q13.12	D20S454	42.7
Q28	20q13.12	D20S119	43.1
Q25	20q13.12	D20S838	44.1
Q16	20q13.12	D20S888	44.6
Q27	20q13.12	D20S891	45.4
Q22	20q13.13	D20S178	46
Q29	20q13.13	D20S866	46.8
Q15	20q13.13	D20S887	47.2
Q30	20q13.13	D20S109	48.5
Q31	20q13.13	D20S196	49
Q17	20q13.2	D20S857	49.5
Q18	20q13.2	D20S902	51.2
Q3	20q13.2	D20S120	52.4
Q24	20q13.31	D20S100	53.7
Q6	20q13.32	D20S171	57.2
Q32	20q13.33	–	60.1
Q33	20q13.33	–	61.5
Q34	20q13.33	–	62.5

2.3. aCGH

Pangenomic, high-density aCGH was conducted on 39 samples from within the subgroup of 74 carcinoma samples. This study was performed using Human Genome CGH 44B oligo microarrays (Agilent Technologies, Santa Clara, CA), which contain 44,000, 50-mer oligonucleotide probes with an average spatial resolution of ~35 kb. Hybridization was done using 3 µg of tumor genomic DNA or reference DNA (Promega Pool DNA of American males or females) as already described [18].

The circular binary segmentation algorithm [19] was applied to the raw data of every sample to define regions of equal copy number (referred to as segments). The False Discovery rate FDR method [20] of 5% was used.

2.4. Real time quantitative PCR

QPCR for TOP1 was performed on the collection of 66 adenomas and 223 carcinomas. The DNA copy number of the TOP1 gene was determined relative to the β-globin gene (HBB), a reference gene that aCGH showed not to vary in our carcinoma cohort. The relative TOP1 copy number was determined using the LightCycler PCR system (Roche Molecular Biochemicals, Mannheim, Germany) and the dsDNA-binding dye SYBR Green I (FastStart DNA Master Mix SYBR Green I, Roche Diagnostics). Primers used were as follows: TOP1 forward, 5'-ATGGGTACAGTGTGCT-3'; TOP1 reverse, 5'-AGTTTGGAGGTCCAG-3'; HBB forward, 5'-ACACAACCTGTGTTCACTAGC-3'; HBB reverse, 5'-CAACTTCATCCACGTTCCACC-3'. After verifying the efficiency of both primer pairs (2.02 for TOP1 and 1.97 for HBB), we used the comparative Ct method ($2^{-\Delta\Delta Ct}$) to determine the TOP1 copy number in the carcinoma specimens, which we normalized to a pool of normal mucosa DNA and to the results for the reference gene [21].

2.5. Fluorescent in situ hybridisation

2.5.1. Slide preparation

Imprints of normal and malignant samples were taken from frozen tissues, fixed in PBS containing formaldehyde and 50 mM MgCl₂, and then treated for 13 min in a 10 mg/ml pepsin solution at 37 °C.

2.5.2. Probes and hybridization

A centromeric probe was chosen to identify chromosome 20 (CEP 20, Vysis, Downers Grove, IL). Chromosome 6 (CEP 6, Vysis, Downers Grove, IL) was used as a control, since aCGH indicated minimal changes to this chromosome in the cohort. Probes to chromosomes 6 and 20 were labelled with SpectrumOrange and SpectrumGreen, respectively.

An aliquot (0.5 µl) of each probe was combined with 3.5 µl of hybridization solution (Vysis, Downers Grove, IL) and 1 µl of sterile water; the 5-µl mixture was then applied to the slide. Probe and target DNA were denatured simultaneously at 73 °C for 5 min prior to overnight hybridization at 37 °C in a humidified hybridizer (Dako, Glostrup, Denmark). After 1 post-hybridization wash in 0.4X SSC/0.3% Igepal at 73 °C for 2 min and in 2X SSC/

0.1% Igepal at room temperature for 1 min, nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI) in mounting solution. The number of signals per nucleus was counted using an Olympus BX60 microscope with appropriate filters. At least 100 nuclei were counted in each experiment.

3. Results

In order to characterize chromosome 20 status in a collection of CIN distal colon cancers, we followed a dual approach based on (1) allelotyping, a routine clinical method used to assess colon cancer phenotypes; and (2) aCGH, a high-resolution screening technology to measure DNA copy number over the entire genome, with a particular emphasis on chromosome 20q.

A set of aberrations affecting chromosome arms 8p, 17p, and 18q were observed by aCGH as already described in colon carcinomas [22,23] but a detailed analysis of these aberrations will be described elsewhere. The present paper will focus on chromosome 20.

3.1. Chromosome 20q allelotyping (cohort of 74 carcinomas)

Allelotyping of 32 microsatellites (MS) spread along the full length of chromosome 20q, with an average spacing of 1 Mb, was performed on a set of 74 distal colon carcinomas exclusively of the CIN phenotype. By clusterization, two groups could be clearly identified: a group of 30 samples (40%) with no AI, and a group of 44 samples (60%) with AI along the entire length of the 20q arm (Fig. 1). The clusterization failed to detect any distinct aberrations in any portion along chromosome 20q, indicating that the aberrations, when present, affect the entire 20q arm. In order to extend the analysis to a larger cohort, we performed allelotyping only at the D20S855 MS in the collection of 223 distal CIN carcinomas, and the results indicate that chromosome 20q is altered in 144 cases (*group AI*; 65%) and normal in 79 cases (*group N*; 35%, data not shown). The presence or absence of AI did not correlate with a carcinoma clinical stage ($p = 0.392$).

3.2. aCGH data analysis of chromosome 20 (cohort of 39 carcinomas)

aCGH analysis was conducted on 39 carcinomas from the subset of 74 carcinomas analysed for the 20q allelotyping. The aCGH showed that chromosome 20q was amplified over its entire length in 30 of the 39 samples (77% with genomic gain, *group G*, Fig. 2A). The nine other samples (23% with disomic 20q, *group N*) lacked any aberrations along the entire chromosome (Fig. 2B). Considering that there is an amplification if a log₂ ratio is higher than 0.225 (i.e. for ~2.34 copies [24]), 99% of the oligos indicated amplification in 4–7 carcinomas among the 30 patients of *group G*. Comparing 4–7 among 30 carcinomas with a chi-square proportion test gives a p -value of 0.50. The observed differences are therefore non-significant, so we concluded that no region was differentially amplified. This finding supports the notion that chromosome 20q was amplified uniformly over its entire length.

In *group G*, which featured amplification of the entire chromosome 20q, 16 (53%) showed a gain in 20p, eight (27%) showed a deletion in 20p, and two (7%) showed a gain of the centromeric part together with a deletion in the telomeric region. In the last four samples, chromosome 20p was non-aberrant. In *group N* (no aberration observed on chromosome 20q), chromosome 20p was disomic in eight carcinomas and deleted in the remaining one.

Carcinomas analysed by aCGH were not numerous enough to investigate a correlation between chromosome 20 genomic status and clinical stage. Nonetheless, all clinical stages were equally represented in *group G*, indicating that a correlation, if it exists, is weak.

3.3. Comparison of allelotyping and aCGH data (cohort of 39 carcinomas)

By allelotyping, all carcinomas displaying AI at chromosome 20q (*group AI*) also showed a genomic gain by aCGH (*group AI/G*, $n = 20$, 51%). In contrast, among the 19 carcinomas without AI (*group N*), 10 cases showed a chromosome 20q gain by aCGH (*group N/G*, 26%), and nine carcinomas were therefore disomic for 20q (*group N/N*, 23%, Table 3A).

In order to validate the discrepancy between allelotyping and aCGH data, relative copy number of the TOP1 gene was evaluated by QPCR (Table 3A). The TOP1 gene is located at 20q12, and given that the CGH analysis indicated a uniform amplification of the full arm, assessing TOP1 gene status should be sufficient to reflect 20q amplification. Each case of amplification initially detected by aCGH was successfully validated by an increase in TOP1 copy number (Table 3A). Conversely, when aCGH was normal, no increase in TOP1 copy number was observed. Also, QPCR confirmed that the carcinomas devoid of AI but showing a gain by aCGH (*group N/G*; $n = 10$) had the chromosome 20q amplified. These results indicate the power of the combined approaches used in the present study.

Because of the discrepancy between allelotyping and aCGH analysis, TOP1 relative copy number was assessed for the 79 of 223 carcinomas in which allelotyping did not detect any AI (*group N*). Among this cohort, 37/79 cases showed an increase in TOP1 relative copy number (*group N/G*, 47%) whereas 42/79 carcinomas had no increase (*group N/N*, 53%).

Overall, among the cohort of 223 carcinomas, 144/223 carcinomas belong to the *group AI/G* (65%), 37/223 carcinomas belong to the *group N/G* (17%) and 42/223 to the *group N/N* (18%) suggesting that allelotyping underestimated chromosome 20q gain in 17% of distal CIN colon carcinomas (Table 3B). Interestingly, when considering the histological phenotype of the carcinomas (Table 1), most of the carcinomas were differentiated (78%) but the mucinous phenotype (16%) was significantly associated to a disomic 20q ($p = 0.005$).

Based on the combination of allelotyping and QPCR data, 181 of 223 (82%) of our CIN distal colon cancer samples showed gain of 20q; this did not correlate with clinical stage ($p = 0.37$) as gains were detected in 78% of early stage carcinomas (stages 0/I/II, 74 of 95), 87% of stage III (41 of 47), and 75% of metastatic carcinomas (stage IV, 61 of 81). We tested whether chromosome 20q status could predict a prognosis value by evaluating its influence on the event-free survival rate of patients (Table 1). Using Kaplan–Meier curves and log-rank tests, no correlation was observed in the cohort of 120 patients including stage 0/I to stage IV ($n = 120$, $p = 0.223$, data not shown). We next carried an independent analysis for each stage, but again no correlation was found (stages 0/I, $p = 0.382$; stage II, $p = 0.213$; stage III, $p = 0.71$; stage IV, $p = 0.644$). Similarly, no correlation was observed when early stages were compared to late stages (0/I/II, $p = 0.787$; stages III/IV, $p = 0.566$). Thus, in our homogeneous cohort of distal CIN carcinomas, 20q status had no prognostic value.

3.4. Mechanisms of 20q gain

We hypothesized that a 20q gain that is detected by aCGH but not by allelotyping could result from the duplication of both parental chromosomes (tetra/polysomy), whereas AI results from the duplication of only one parental chromosome (aneusomy). To test this hypothesis, 20q copy number was quantified by FISH for three carcinomas *N/G*, three carcinomas *N/N*, three carcinomas *AI/G*, and one normal colon mucosa; chromosome 6 served as a control. As expected, the *N/N* carcinomas and the normal mucosa were disomic for chromosomes 6 and for chromosome 20 (mean copy number was 2.02 ± 0.21 , Fig. 3A). *AI/G* carcinomas were trisomic for chromosome 20 in 87% of the nuclei (mean copy number was 3.3 ± 0.33 , Fig. 3B), while 88% of nuclei in the three *N/G* carcinomas were tetrasomic for chromosome 20 (mean copy number was 4.3 ± 0.45) and disomic for chromosome 6 (Fig. 3C). In addition, a few nuclei in *N/G* carcinomas were polysomic, with more than four copies of chromosome 20.

In order to determine whether tetrasomy can affect chromosomal regions other than 20q, we used allelotyping to examine the status of MS located on chromosomes shown by aCGH to be fully gained, i.e. chromosome arms 1q (16 carcinomas), 7p+q (20 carcinomas), 8q (21 carcinomas), and 13q (25 carcinomas). The aCGH analysis revealed five carcinomas where chromosome 20q gains were associated with another chromosome gain (either 1q, 7p+q, 8q or 13q), but without AI. Among these five carcinomas, only one showed simultaneous gain at 1q, 13q, and 20q without AI. Thus, tetrasomy preferentially affected chromosome 20q.

3.5. Chromosome 20q and genetic context

In order to achieve better subtyping of the carcinomas at the molecular level, we searched for specific genomic aberrations associated with a normal or amplified chromosome 20q. The statistical FDR 5% method [19]

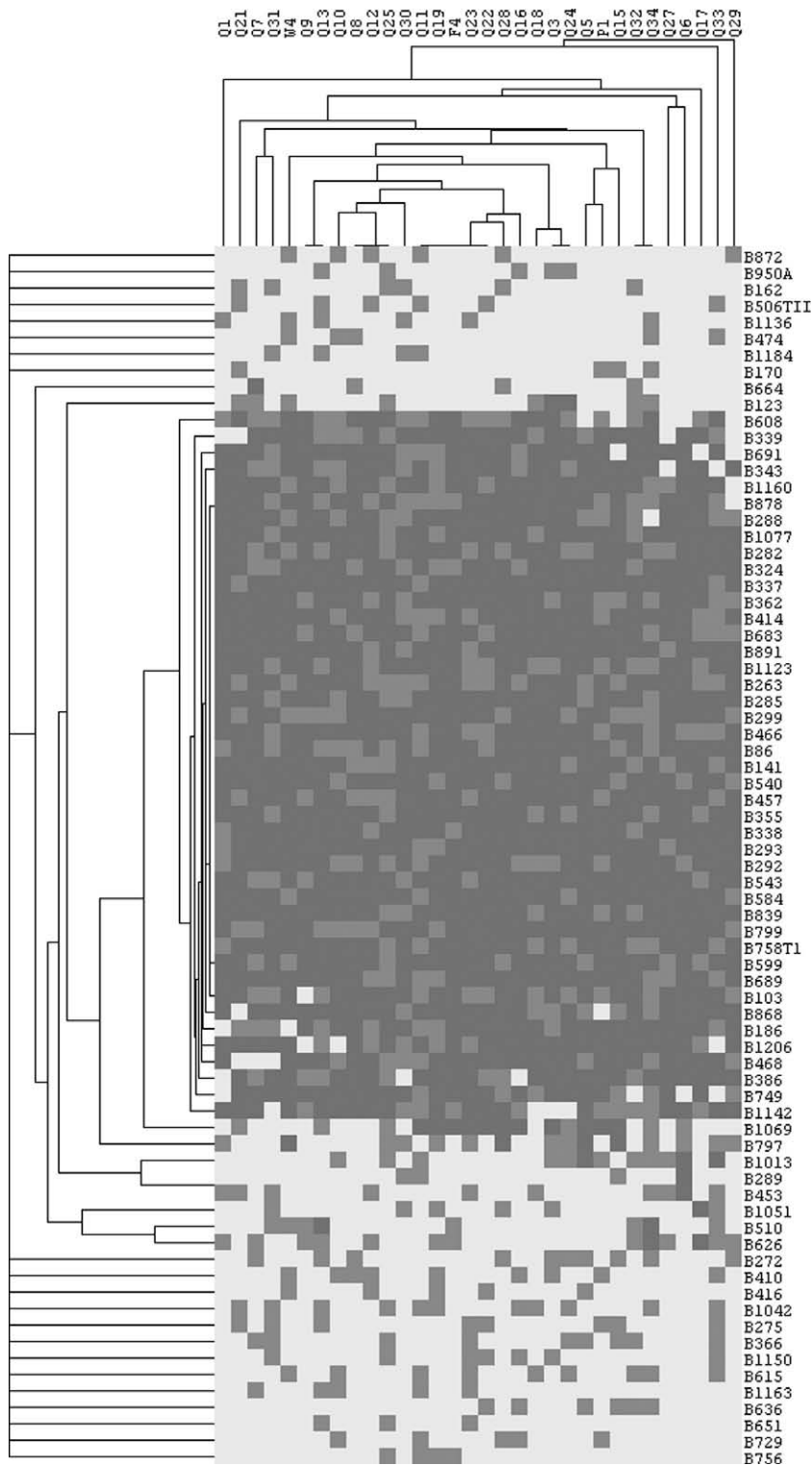


Fig. 1. Two-way unsupervised hierarchical clustering of the allelotyping data using 32 microsatellites covering the 20q on 74 distal CIN colon carcinomas. Normal locus (light gray rectangles); Locus with AI (black rectangles); non-informative locus (dark gray rectangles). Vertical columns correspond to the microsatellites; horizontal lines correspond to the carcinomas.

was performed on the entire genomic aCGH dataset, except for chromosome 20 data and this method did not distinguish between carcinomas

that had gained 20q (*group AI/G* and *group N/G*) and those that were disomic for 20q (*group N/N*). Allelotyping of a microsatellite close to the TP53

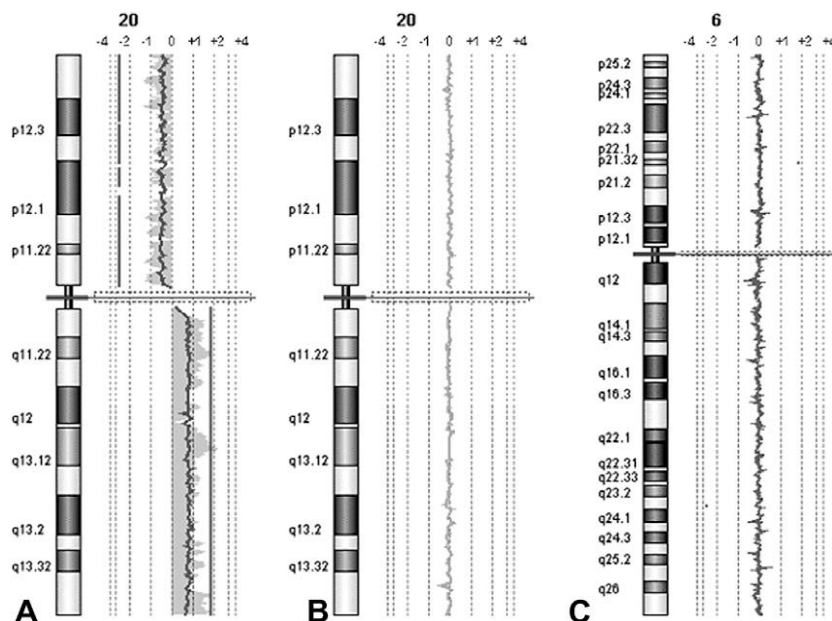


Fig. 2. CGH array profiles for chromosome 20 and chromosome 6 as a control. Genomic profiles for two stage IV carcinomas: (A) AI by allelotyping, 20q gain as shown by aCGH, and QPCR (*group AI/G*), (B) no aberration by allelotyping, aCGH or QPCR (*group N/N*). Chromosome 6 was used as control in FISH experiment (C).

gene in the cohort of 223 carcinomas showed the TP53 locus to be aberrant more frequently in the group of carcinomas with a 20q gain (129 of 149 informative for TP53 locus, 87%) than in the group disomic for 20q (20 of 33 informative for TP53, 61%, $p = 0.001$). In contrast, no significant association with 20q status was observed for the APC locus (196 informative carcinomas, $p = 0.308$). The APC, TP53, and 20q status of carcinomas is summarized in Table 4, and this indicates that 45% of the cases possessed the three aberrations simultaneously.

We showed in previous work [13] that human CIN colon carcinomas can be divided based on global microsatellite AI frequency into “high” ($65.2\% \pm 16.7\%$ of loci altered), “low” ($38.9\% \pm 9.4\%$ of loci altered), and “very low” ($14\% \pm 8.1\%$ of loci altered) groups. Interestingly, Fisher’s exact test carried out on the 223 carcinomas in this study shows that most of the carcinomas disomic for 20q (*group N/N*) belong to the low group (35% AI frequency, $n = 42$), whereas most carcinomas that gained 20q (either *groups AI/G* or *N/G*) belong to the high group (54% AI frequency, $p = 0.001$).

3.6. Chromosome 20q gain: an early event

In order to determine if chromosome 20q gain was also observed in precancerous lesions and then possibly involved in initiation or adenoma-carcinoma transition, we analysed the D20S855 MS status by allelotyping in a cohort of 66 distal adenomas (Table 3C). The results indicated that chromosome 20q was altered in 8/66 adenomas (*group AI*, 12%) and normal in 58/66 cases (*group N*, 88%). To estimate the presence of tetraploid 20q, TOP1 relative copy number was performed on the 58 adenomas without AI. QPCR indicated increased TOP1 relative copy number in 11/58 adenomas corresponding to 17% in the cohort of 66 adenomas (*group NR/G*). Thus overall, the 20q is gained in 19/66 adenomas (*group G*, 29%; Table 3C). TP53 and APC loci were altered in 9/52 (17%, 14 non-informative cases) and 18/51 (35%, 15 non-informative cases) adenomas respectively. Using Chi2 test, we observed an association between 20q and TP53 genomic status but not with APC in adenomas ($p = 0.029$ and $p = 0.193$ respectively). The Fisher exact test performed on the allelotyping data of the 66 distal CIN adenomas showed the disomic 20q (*group N/N*) to be associated to the “very-low” frequency group ($p = 0.007$). On the other hand, no correlation was found between the grade or histology of the adenomas and the 20q genomic status ($p = 0.088$ and $p = 0.205$ respectively).

Therefore, using two independent genomic approaches, we have been able to define two groups of carcinomas based on chromosome 20q status: a group in which chromosome 20q is gained, and this gain is associ-

ated with a high level of genomic aberrations but does not necessarily show AI at the APC locus; and a group of carcinomas disomic for 20q, and this disomy is associated with a low level of aberrations, none of which have been specifically identified so far. In addition, our results point to a weakness of the allelotyping approach in detecting genomic aberrations resulting from simultaneous amplification of parental alleles.

4. Discussion

The genomic status of chromosome 20q in colon cancer was analysed using allelotyping and aCGH approaches in a homogeneous cohort restricted to distal colon cancer of CIN phenotype. Several reports have emphasized the importance of separating the analyses of carcinomas of proximal and distal origins, as well as of CIN and MIN phenotypes [3,4]. In a carcinoma cohort that was homogeneous for origin and phenotype, allelotyping indicated that the 20q arm is either aberrant or unaltered, and bioinformatics analyses of aCGH data showed that the 20q aberration is due to a homogeneous gain. Complementary studies of allelotyping and aCGH also showed that 20q gain results from both aneuploidy, with only one of the two parental chromosomes gained, and tetrasomy, with both parental chromosomes gained together. Comparison of 20q status in adenomas and carcinomas suggest that 20q gain resulting from tetraploidy would be an initiating event in adenoma occurrence whereas the gain resulting from aneuploidy would rather be implicated in adenoma to carcinoma transition.

4.1. Chromosome 20q status

Allelotyping is a fast, cheap, and sensitive approach used in routine clinical diagnosis to detect AI and to type CIN and MIN carcinomas. In our study, allelotyping of the

Table 3

Comparison of allelotyping/aCGH/QPCR data to determine the status of chromosome 20q in carcinomas and adenomas. Values of $2^{-\Delta\Delta Ct} > 1.31$ indicate a genomic gain, while $X < 2^{-\Delta\Delta Ct} > 1.31$ indicates a disomic chromosome. (A) Chromosome 20q status: Concordance between 20q allelotyping, aCGH, and QPCR data for 39 distal carcinomas of the aCGH cohort. (B) Chromosome 20q status: Concordance between allelotyping and QPCR data for 223 distal carcinomas. Allelotyping was performed at the D20S855 locus and QPCR quantified *TOP1* copy number. (C) Chromosome 20q status: Concordance between allelotyping and QPCR data for 66 distal adenomas. Allelotyping was performed at the D20S855 locus and QPCR quantified *TOP1* copy number. (D) Schematic representation of chromosome 20q status in distal adenomas and carcinomas.

	20q Allelotyping	aCGH and QPCR <i>TOP1</i>	Final 20q status
(A)	20/39 (51%) Group AI	20/39 (51%) Group AI/G	30/39 (77%) Group G
	19/39 (49%) Group N	10/39 (26%) Group N/G 9/39 (23%) Group N/N	9/39 (23%) Group N
	D20S855 Allelotyping	QPCR <i>TOP1</i>	20q status
(B)	144/223 (65%) Group AI	144/223 (65%) Group AI/G	181/223 (82%) Group G
	79/223 (35%) Group N	37/223 (17%) Group N/G 42/223 (18%) Group N/N	42/223 (18%) Group N
	D20S855 Allelotyping	D20S855 Allelotyping and QPCR <i>TOP1</i>	20q status
(C)	8/66 (12%) (Group AI)	8/66 (12%) Group AI/G	19/66 (29%) Group G
	58/66 (88%) (Group N)	11/66 (17%) Group N/G 47/66 (71%) Group N/N	47/66 (71%) Group N
(D)			

Panel A: AI, allelic imbalance at D20S855; G, chromosome 20q gain as shown by aCGH and QPCR for *TOP1*; N, no allelic imbalance at D20S855; N/N, no allelic imbalance at D20S855 and no chromosome 20q gain as shown by aCGH and QPCR for *TOP1*.

Panel B: AI, allelic imbalance at D20S855; G, chromosome 20q gain as shown by aCGH and QPCR; N, no allelic imbalance at D20S855; N/N, no allelic imbalance at D20S855 and no chromosome 20q gain as shown by aCGH and QPCR.

Panel C: AI, allelic imbalance at D20S855; G, chromosome 20q gain as shown by QPCR; N, no allelic imbalance at D20S855; N/N, no allelic imbalance at D20S855 and no chromosome 20q gain as shown by QPCR.

Panel D: AI, allelic imbalance at D20S855; G, chromosome 20q gain as shown by QPCR; N, no allelic imbalance at D20S855; N/N, no allelic imbalance at D20S855 and no chromosome 20q gain as shown by QPCR.

chromosome 20q arm allowed us to show that colon carcinogenesis may arise not only from a 20q gain, but also from a disomic chromosome 20q. To gain information about the gene copy number level, allelotyping has been supplemented with aCGH. Since the average resolution of a 44 K array is ~ 35 kb, copy number changes can be directly related to genes involved in a pathological process [25]. Using this approach, we found that AI of 20q in our cohort was due to a full gain of the entire arm, with all loci homogeneously amplified. Chromosome gains have been described in many other cancers, including breast (1q, 8q [26]), melanoma (1q, 7, 8q, 17q, 20q [27]), ovary (3q, 8q, [28]), and esophageal cancers (3q, 8q, 9q, 12q, 16p, 17, 19, 20q, 22 [29]). In fact, several authors have reported that

copy number at the 20q13.2 region is higher in carcinomas with liver metastases than in carcinomas without metastasis [5,30] and this higher copy number correlates with faster carcinoma progression and lower rates of patient survival [10]. On the other hand, Diep et al. [31] showed that carcinomas with 20q amplification have a longer survival rate, but only above a certain level of amplification. This difficulty in determining whether chromosome 20q amplification affects clinical outcome in cancer patients probably reflects the fact that most studies have included all carcinoma locations and phenotypes in constructing their cohorts. By working with a cohort homogeneous for carcinoma location and phenotype, the present study found no correlation between chromosome 20q status

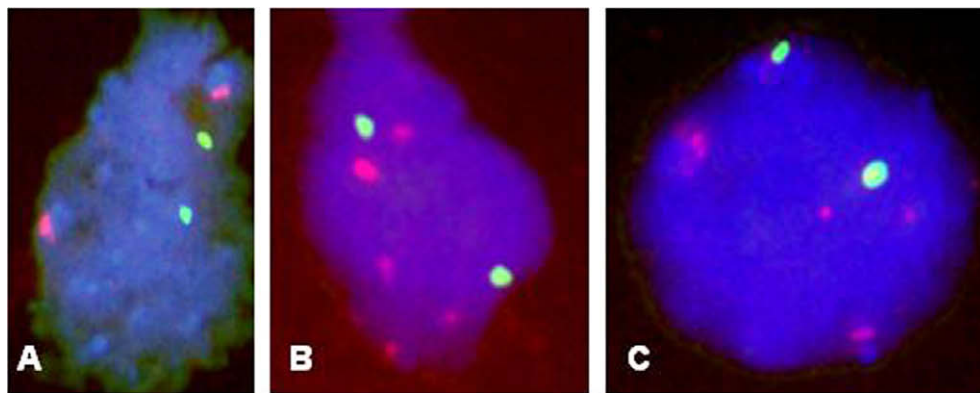


Fig. 3. FISH experiment performed using centromeric probes for examination of chromosome 20 status (red spots) and chromosome 6 (green spots) as a control. Magnification 1000X. (A) A diploid carcinoma nucleus showing two chromosome 20 signals (group N/N, stage III). (B) An aneuploid carcinoma nucleus showing five chromosome 20 signals (group AI/G, stage II). (C) A tetraploid carcinoma nucleus showing four chromosome 20 signals (group N/G, stage II).

Table 4

Summary of genomic aberrations found through allelotyping 48 informative adenomas and 163 informative tumors for APC and TP53 loci and through allelotyping and QPCR for chromosome 20q. (AI) with allelic imbalance; (N) without allelic imbalance.

20q QPCR	TP53	APC	Number		%	
			Carcinomas	Adenomas	Carcinomas	Adenomas
AI	AI	AI	73	5	45	10.5
AI	AI	N	44	2	27	4
N	AI	AI	13	/	8	/
AI	N	N	10	5	6	10.5
N	AI	N	7	1	4	2
N	N	N	8	22	5	46
N	N	AI	4	10	2.5	21
AI	N	AI	4	3	2.5	6

and clinical stage or the event-free survival rate. Concerning the prevalence of mucinous phenotype in the normal 20q subgroup, some reports observed that the mucinous phenotype could confer a lower frequency of metastasis to an axillary lymph node in breast cancer [32] and a greater survival benefit from chemotherapy in colon cancer [33]. However, the status of chromosome 20q confers no particular benefit to patients studied in the cohort of distal tumors.

4.2. Mechanism of 20q gain

In our study, aneusomy was observed in carcinomas of all clinical stages, which displayed allelic imbalance and amplification detected by both aCGH and by FISH, which detected an uneven number of chromosomes 20. In contrast, chromosome 20q tetrasomy was present in a subgroup of carcinomas of all clinical stages that presented no AI but amplification detectable by aCGH. FISH analysis confirmed that most nuclei in these carcinomas had four copies of chromosome 20.

Tetra- or polyploid cells arising from cellular stress and ageing are genetically unstable and can act as an intermediate state towards aneuploidy [34,35]. This idea is appealing because it may explain why many cancers contain large-scale changes in chromosome numbers but rarely

show simple gains or losses of individual chromosomes. However, in the present study, carcinomas at all clinical stages were found to be affected by a 20q tetra/polysomy, not only early stage carcinomas. This suggests that tetra/polyploidy, in addition to being an intermediate state towards aneuploidy, is a state that does not evolve further during cancer progression. Mechanisms leading to tetra/polyploidy are diverse [36] and can involve either non-disjunction of mitotic chromosomes [37] or a defective tetraploid checkpoint that fails in its normal function of arresting and eliminating tetra/polyploid cells in G1. This checkpoint can become defective as a result of mutations in the TP53 gene or in related genes [38,39]. In either scenario, TP53 is a leading actor, and it may or may not be associated with disruption of the APC gene, which has been proposed to increase the incidence of tetra/polyploidy in cancer [40]. Interestingly, the present study found that 92% of the N/G carcinomas had AI at the TP53 locus, but only 64% had AI at the APC locus.

At present, there is no simple explanation for the frequent centromeric breakage observed in chromosome 20. This phenomenon has been hypothesized to result from hypomethylation of DNA in the centromeric region, with subsequent decondensation of centromeres [41], promoting breakage and frequent recombination. Moreover, changes in components such as kinetochore-associated

checkpoint proteins or in centromeric proteins such as CENP may prevent the proper attachment of kinetochores, leading to disrupted chromosome segregation, delayed mitosis, and chromosome breakage malignancies [42]. A better understanding of the mechanisms of chromosome 20q gain may require further analysis of the proteins implicated in the segregation of chromosomes and in the spindle checkpoint.

4.3. Chromosome 20q gain: early event

Chromosome 20q aberration may be associated with the adenoma-carcinoma transition rather than with carcinoma progression, as previously proposed [43,44]. Indeed, the 20q gain in the present study was observed at the same frequency in early and late stage carcinomas, suggesting that this gain occurs early in colon carcinogenesis. In fact, the 20q gain was already present in precancerous lesions, as 29% of adenomas had increased TOP1 gene copy number, compared with 78% of early stage carcinomas. Looking more specifically at 20q gained by tetraploidy, we found that the incidence is similar in adenomas and carcinomas (17%), suggesting that the tetraploidy would be an event of adenomas initiation while aneuploidy would rather be critical for the progression from premalignant adenoma to invasive carcinoma.

In conclusion, we show in the present work that in a homogeneous cohort of human distal CIN colon cancers, chromosome 20q was either normal or amplified, that this gain can arise from aneuploid or tetrasomic chromosome 20q. The fact that colon cancer can proceed through all stages with a normal chromosome 20q, in the absence of APC or TP53 aberrations and indeed with few genomic aberrations in general, reinforces the concept that several pathways are involved in colon carcinogenesis. Moreover, the diversity of combinations of aberrations further increases the difficulty in identifying the key events in carcinoma progression. Finally, if allelotyping continues to be used widely for fast diagnosis, clinicians will fail to detect abnormalities resulting from duplication of both parental chromosomes. To avoid this problem, allelotyping should be used in conjunction with a complementary method suitable for routine clinical testing, such as QPCR.

Conflicts of interest

None.

Acknowledgements

We thank Sabine Quief and Céline Villenet for performing the CGH arrays (IRC, Inserm, Lille, France), Béatrice Lannes and Martine Muckensturm (CHRU, Strasbourg, France) for helpful assistance in FISH experiments, Florence Ackermann (CHRU, Strasbourg, France), Mathilde Arrive, and Sylvie Delacour (CHRU, Strasbourg, France) for their excellent technical assistances, Nicolas Meyer (Département de Santé Publique, CHRU Strasbourg) for the achievement of the two-way unsupervised clustering, Marc Guennegues (CGE, Strasbourg, France) for his critical reading of the manuscript.

This work was supported by Inserm, the Institut National du Cancer (INCA, ACI 2004) and the Ligue Régionale contre le Cancer. C. Nicolet is a recipient of a Region Alsace/Inserm grant program.

References

- [1] K.W. Kinzler, B. Vogelstein, Lessons from hereditary colorectal cancer, *Cell* 87 (1996) 159–170.
- [2] B. Vogelstein, K.W. Kinzler, Cancer genes and the pathways they control, *Nat. Med.* 10 (2004) 789–799.
- [3] G. Matanoski, X.G. Tao, L. Almon, A.A. Adade, J.O. Davies-Cole, Demographics and tumor characteristics of colorectal cancers in the United States, 1998–2001, *Cancer* 107 (2006) 1112–1120.
- [4] F.A. Sinicrope, R.L. Rego, N. Foster, D.J. Sargent, H.E. Windschitl, L.J. Burgart, T.E. Witzig, S.N. Thibodeau, Microsatellite instability accounts for tumor site-related differences in clinicopathologic variables and prognosis in human colon cancers, *Am. J. Gastroenterol.* 101 (2006) 2818–2825.
- [5] A. Leslie, A. Stewart, D.U. Baty, D. Mehan, L. MCGreavey, G. Smith, C.R. Wolf, M. Sales, N.R. Pratt, R.J. Steele, F.A. Carey, Chromosomal changes in colorectal adenomas: relationship to gene mutations and potential for clinical utility, *Genes Chromosomes Cancer* 45 (2006) 126–135.
- [6] S. Hidaka, T. Yasutake, H. Takeshita, M. Kondo, T. Tsuji, A. Nanashima, T. Sawai, H. Yamaguchi, T. Nakagoe, H. Ayabe, Y. Tagawa, Differences in 20q13.2 copy number between colorectal cancers with and without liver metastasis, *Clin. Cancer Res.* 6 (2000) 2712–2717.
- [7] F. Mertens, B. Johansson, M. Hoglund, F. Mitelman, Chromosomal imbalance maps of malignant solid tumors: a cytogenetic survey of 3185 neoplasms, *Cancer Res.* 57 (1997) 2765–2780.
- [8] D.E. Aust, R.F. Willenbacher, J.P. Terdiman, L.D. Ferrell, C.G. Chang, D.H. Moore 2nd, A. Molinaro-Clark, G.B. Baretton, U. Loehrs, F.M. Waldman, Chromosomal alterations in ulcerative colitis-related and sporadic colorectal cancers by comparative genomic hybridization, *Hum. Pathol.* 31 (2000) 109–114.
- [9] D. Tsafrir, M. Bocolod, Z. Selvanayagam, I. Tsafrir, J. Shia, Z. Zeng, H. Liu, C. Krier, R.F. Stengel, F. Barany, W.L. Gerald, P.B. Paty, E. Domany, D.A. Nottelman, Relationship of gene expression and chromosomal abnormalities in colorectal cancer, *Cancer Res.* 66 (2006) 2129–2137.
- [10] D.E. Aust, M. Muders, A. Kohler, M. Schmidt, J. Diebold, C. Muller, U. Lohrs, F.M. Waldman, G.B. Baretton, Prognostic relevance of 20q13 gains in sporadic colorectal cancers: a fish analysis, *Scand. J. Gastroenterol.* 39 (2004) 766–772.
- [11] E.J. Douglas, H. Fiegler, A. Rowan, S. Halford, D.C. Bicknell, W. Bodmer, I.P. Tomlinson, N.P. Carter, Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas, *Cancer Res.* 64 (2004) 4817–4825.
- [12] W.M. Korn, T. Yasutake, W.L. Kuo, R.S. Warren, C. Collins, M. Tomita, J. Gray, F.M. Waldman, Chromosome arm 20q gains and other genomic alterations in colorectal cancer metastatic to liver, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization, *Genes Chromosomes Cancer* 25 (1999) 82–90.
- [13] J.C. Weber, N. Meyer, E. Pencreach, A. Schneider, E. Guerin, A. Neuville, C. Stemmer, C. Brigand, P. Bachellier, S. Rohr, M. Keding, C. Meyer, D. Guenot, P. Oudet, D. Jaeck, M.P. Gaub, Allelotyping analyses of synchronous primary and metastasis cin colon cancers identified different subtypes, *Int. J. Cancer* 120 (2007) 524–532.
- [14] A. Schneider, S. Borgnat, H. Lang, O. Regine, V. Lindner, M. Kassem, C. Saussine, P. Oudet, D. Jacqmin, M.P. Gaub, Evaluation of microsatellite analysis in urine sediment for diagnosis of bladder cancer, *Cancer Res.* 60 (2000) 4617–4622.
- [15] O. Delattre, S. Olschwang, D.J. Law, T. Melot, Y. Remvikos, R.J. Salmon, X. Sastre, P. Validire, A.P. Feinberg, G. Thomas, Multiple genetic alterations in distal and proximal colorectal cancer, *Lancet* 2 (1989) 353–356.
- [16] S. Olschwang, R. Hamelin, P. Laurent-Puig, B. Thuille, Y. De Rycke, Y.J. Li, F. Muzeau, J. Girodet, R.J. Salmon, G. Thomas, Alternative genetic pathways in colorectal carcinogenesis, *Proc. Natl. Acad. Sci. USA* 94 (1997) 12122–12127.
- [17] D. Guenot, E. Guerin, S. Aguillon-Romain, E. Pencreach, A. Schneider, A. Neuville, M.P. Chenard, I. Duluc, S. Du Manoir, C. Brigand, P. Oudet, M. Keding, M.P. Gaub, Primary tumour genetic alterations and intra-tumoral heterogeneity are maintained in xenografts of human colon cancers showing chromosome instability, *J. Pathol.* 208 (2006) 643–652.

- [18] F. Jardin, P. Ruminy, J.P. Kerckaert, F. Parmentier, J.M. Picquenot, S. Quief, C. Villenet, G. Buchonnet, M. Tosi, T. Frebourg, C. Bastard, H. Tilly, Detection of somatic quantitative genetic alterations by multiplex polymerase chain reaction for the prediction of outcome in diffuse large B-cell lymphomas, *Haematologica* 93 (2008) 543–550.
- [19] W.R. Lai, M.D. Johnson, R. Kucherlapati, P.J. Park, Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data, *Bioinformatics* 21 (2005) 3763–3770.
- [20] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, *JRSS Ser. B (Methodological)* 57 (1995) 289–300.
- [21] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative pcr and the 2(-delta delta c(T)) method, *Methods* 25 (2001) 402–408.
- [22] G.A. Meijer, M.A. Hermsen, J.P. Baak, P.J. Van Diest, S.G. Meuwissen, J.A. Belien, J.M. Hoovers, H. Joenje, P.J. Snijders, J.M. Walboomers, Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation, *J. Clin. Pathol.* 51 (1998) 901–909.
- [23] M. Hoglund, D. Gisselsson, N. Mandahl, B. Johansson, F. Mertens, F. Mitelman, T. Sall, Multivariate analyses of genomic imbalances in solid tumors reveal distinct and converging pathways of karyotypic evolution, *Genes Chromosomes Cancer* 31 (2001) 156–171.
- [24] K. Nakao, K.R. Mehta, J. Fridlyand, D.H. Moore, A.N. Jain, A. Lafuente, J.W. Wiencke, J.P. Terdiman, F.M. Waldman, High-resolution analysis of DNA copy number alterations in colorectal cancer by array-based comparative genomic hybridization, *Carcinogenesis* 25 (2004) 1345–1357.
- [25] D. Pinkel, R. Seagraves, D. Sudar, S. Clark, I. Poole, D. Kowbel, C. Collins, W.L. Kuo, C. Chen, Y. Zhai, S.H. Dairkee, B.M. Ljung, J.W. Gray, D.G. Albertson, High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays, *Nat. Genet.* 20 (1998) 207–211.
- [26] A. Kallioniemi, O.P. Kallioniemi, J. Piper, M. Tanner, T. Stokke, L. Chen, H.S. Smith, D. Pinkel, J.W. Gray, F.M. Waldman, Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization, *Proc. Natl. Acad. Sci. USA* 91 (1994) 2156–2160.
- [27] G. Jonsson, C. Dahl, J. Staaf, T. Sandberg, P.O. Bendahl, M. Ringner, P. Guldberg, A. Borg, Genomic profiling of malignant melanoma using tiling-resolution array CGH, *Oncogene* 26 (2007) 4738–4748.
- [28] H. Iwabuchi, M. Sakamoto, H. Sakunaga, Y.Y. Ma, M.L. Carcangiu, D. Pinkel, T.L. Yang-Feng, J.W. Gray, Genetic analysis of benign, low-grade, and high-grade ovarian tumors, *Cancer Res.* 55 (1995) 6172–6180.
- [29] S.D. Pack, J.D. Karkera, Z. Zhuang, E.D. Pak, K.V. Balan, P. Hwu, W.S. Park, T. Pham, D.O. Ault, M. Glaser, L. Liotta, S.D. Detera-Wadleigh, R.G. Wadleigh, Molecular cytogenetic fingerprinting of esophageal squamous cell carcinoma by comparative genomic hybridization reveals a consistent pattern of chromosomal alterations, *Genes Chromosomes Cancer* 25 (1999) 160–168.
- [30] H. Aragane, C. Sakakura, M. Nakanishi, R. Yasuoka, Y. Fujita, H. Taniguchi, A. Hagiwara, T. Yamaguchi, T. Abe, J. Inazawa, H. Yamagishi, Chromosomal aberrations in colorectal cancers and liver metastases analyzed by comparative genomic hybridization, *Int. J. Cancer* 94 (2001) 623–629.
- [31] C.B. Diep, L. Thorstensen, G.I. Meling, E. Skovlund, T.O. Rognum, R.A. Lothe, Genetic tumor markers with prognostic impact in dukes' stages b and c colorectal cancer patients, *J. Clin. Oncol.* 21 (2003) 820–829.
- [32] L. Scopsi, S. Andreola, S. Pilotti, R. Bufalino, M.T. Baldini, A. Testori, F. Rilke, Mucinous carcinoma of the breast. A clinicopathologic, histochemical, and immunocytochemical study with special reference to neuroendocrine differentiation, *Am. J. Surg. Pathol.* 18 (1994) 702–711.
- [33] M. Morris, C. Platell, B. Iacopetta, Tumor-infiltrating lymphocytes and perforation in colon cancer predict positive response to 5-fluorouracil chemotherapy, *Clin. Cancer Res.* 14 (2008) 1413–1417.
- [34] L. Comai, The advantages and disadvantages of being polyploid, *Nat. Rev. Genet.* 6 (2005) 836–846.
- [35] Z. Storchova, D. Pellman, From polyploidy to aneuploidy, genome instability and cancer, *Nat. Rev. Mol. Cell Biol.* 5 (2004) 45–54.
- [36] N.J. Ganem, Z. Storchova, D. Pellman, Tetraploidy, aneuploidy and cancer, *Curr. Opin. Genet. Dev.* 17 (2007) 157–162.
- [37] Q. Shi, R.W. King, Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines, *Nature* 437 (2005) 1038–1042.
- [38] M. Castedo, A. Coquelle, S. Vivet, I. Vitale, A. Kauffmann, P. Dessen, M.O. Pequignot, N. Casares, A. Valent, S. Mouhamad, E. Schmitt, N. Modjtahedi, W. Vainchenker, L. Zitvogel, V. Lazar, C. Garrido, G. Kroemer, Apoptosis regulation in tetraploid cancer cells, *Embo. J.* 25 (2006) 2584–2595.
- [39] R.L. Margolis, O.D. Lohez, P.R. Andreassen, G1 tetraploidy checkpoint and the suppression of tumorigenesis, *J. Cell Biochem.* 88 (2003) 673–683.
- [40] D. Dikovskaya, D. Schiffmann, I.P. Newton, A. Oakley, K. Kroboth, O. Sansom, T.J. Jamieson, V. Meniel, A. Clarke, I.S. Nathke, Loss of Apc induces polyploidy as a result of a combination of defects in mitosis and apoptosis, *J. Cell Biol.* 176 (2007) 183–195.
- [41] J.R. Sawyer, G. Tricot, S. Mattox, S. Jagannath, B. Barlogie, Jumping translocations of chromosome 1q in multiple myeloma: evidence for a mechanism involving decondensation of pericentromeric heterochromatin, *Blood* 91 (1998) 1732–1741.
- [42] R.D. Everett, W.C. Earnshaw, J. Findlay, P. Lomonte, Specific destruction of kinetochore protein cenp-c and disruption of cell division by herpes simplex virus immediate-early protein vnm110, *Embo. J.* 18 (1999) 1526–1538.
- [43] T. Ried, R. Knutzen, R. Steinbeck, H. Blegen, E. Schrock, K. Heselmeyer, S. Du Manoir, G. Auer, Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors, *Genes Chromosomes Cancer* 15 (1996) 234–245.
- [44] M. Hermsen, C. Postma, J. Baak, M. Weiss, A. Rapallo, A. Sciuotto, G. Roemen, J.W. Arends, R. Williams, W. Giaretti, A. De Goeij, G. Meijer, Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability, *Gastroenterology* 123 (2002) 1109–1119.

Publication 2

IR spectral imaging for histopathological characterization of xenografted human colon carcinomas.

Wolthuis R, Travo A, Nicolet C, Neuville A, Gaub MP, Guenot D, Ly E, Manfait M, Jeannesson P, Piot O.

Anal Chem. 2008 Nov 15;80(22):8461-9.

IR Spectral Imaging for Histopathological Characterization of Xenografted Human Colon Carcinomas

Rolf Wolthuis, Adrian Travo, Céline Nicolet, Agnès Neuville, Marie-Pierre Gaub, Dominique Guenot, Elodie Ly, Michel Manfait, Pierre Jeannesson, and Olivier Piot

Anal. Chem., **2008**, 80 (22), 8461-8469 • DOI: 10.1021/ac801191x • Publication Date (Web): 11 October 2008

Downloaded from <http://pubs.acs.org> on December 2, 2008

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)

IR Spectral Imaging for Histopathological Characterization of Xenografted Human Colon Carcinomas

Rolf Wolthuis,[†] Adrian Travo,[†] Céline Nicolet,[‡] Agnès Neuville,^{‡,§} Marie-Pierre Gaub,^{‡,||} Dominique Guenot,[‡] Elodie Ly,[†] Michel Manfait,[†] Pierre Jeannesson,^{*,†} and Olivier Piot[†]

MéDIAN UMR CNRS 6237 MEDyC, Université de Reims Champagne-Ardenne, UFR Pharmacie, 51 rue Cognacq-Jay, 51096 Reims Cedex, France, UMR-S682 Inserm, Université Louis Pasteur, Faculté de Médecine, 67200 Strasbourg, France, Laboratoire de Biologie Moléculaire et Biochimie, CHRU Strasbourg, Hôpital de Hautepierre, 67200 Strasbourg, France, and Service d'Anatomo-Pathologie, CHRU Strasbourg, Hôpital de Hautepierre, 67200 Strasbourg, France

This study aims to develop IR imaging of tumor tissues for generating an automated IR-based histology. Formalin-fixed paraffin-embedded xenografts of human colon carcinomas were analyzed. Chemometric and statistical multivariate treatments of spectral data permitted to probe the intrinsic chemical composition of tissues, directly from paraffinized sections without previous dewaxing. Reconstructed color-coded spectral images revealed a marked tumor heterogeneity. We identified three spectral clusters associated to tumoral tissues, whereas HE staining revealed only a single structure. Nine other clusters were assigned to either necrotic or host tissues. This spectral histology proved to be consistent over multiple passages of the same xenografted tumor confirming that intratumoral heterogeneity was maintained over time. In addition, developing an innovative image analysis, based on the quantification of neighboring pixels, permitted the identification of two main sequences of spectral clusters related to the tissue spatial organization. Molecular attribution of the spectral differences between the tumor clusters revealed differences of transcriptional activity within these tumor tissue subtypes. In conclusion, IR spectral imaging proves to be highly effective both for reproducible tissue subtype recognition and for tumor heterogeneity characterization. This may represent an attractive tool for routine high throughput diagnostic challenges, independent from visual morphology.

Histopathological characterization of tissues is a paramount step for the diagnosis of tumoral lesions and as well as for a better understanding of mechanisms associated with the tumor progression. Among the different types of cancer, colon cancer not only represents a high incidence pathology but also one of the most

investigated lesions at the tissular level, from mucin analysis^{1,2} to genetic studies.^{3–5} Unlike classical histology mainly morphology based, or immunohistochemistry specifically targeting certain proteins, vibrational spectroscopy (Raman scattering and Fourier transform infrared (FT-IR) absorption spectroscopy) probes the chemical composition of the tissue and can detect subtle molecular structural changes. It also avoids inherent operator variability. For many different tumors, spectroscopy based histology has proved to allow discrimination between malignant or benign tissues and normal tissues^{6–12} as well as detection of drug resistance in cancer cell lines.¹³ Rigas et al.¹⁴ studied malignant versus normal colon tissues using IR spectroscopy. Focusing on colonic tissue, they found changes correlated with malignancy in specific vibrational frequencies attributed to the nucleic acids and spectral peak shifts related to loss of hydrogen bonds in cell proteins. Other authors demonstrated the diagnostic potential of FT-IR microspectroscopy applied to formalin-fixed chemically deparaffined tissues: carcinomas, polyps and normal tissues were distinguished with a

- (1) Forgue-Lafitte, M. E.; Fabiani, B.; Levy, P.; Maurin, N.; Flejou, J. F.; Bara, J. *Int. J. Cancer* **2007**, *121*, 1543–49.
- (2) Byrd, J.; Bresalier, R. *Cancer Metastasis Rev.* **2004**, *23*, 77–99.
- (3) Guenot, D.; Guérin, E.; Aguillon-Romain, S.; Pencreach, E.; Schneider, A.; Neuville, A.; Chenard, M. P.; Duluc, I.; Du Manoir, S.; Brigand, C.; Oudet, P.; Keding, M.; Gaub, M. P. *J. Pathol.* **2006**, *208*, 643–52.
- (4) Weber, J. C.; Meyer, N.; Pencreach, E.; Schneider, A.; Guérin, E.; Neuville, A.; Stemmer, C.; Brigand, C.; Bachellier, P.; Rohr, S.; Keding, M.; Meyer, C.; Guenot, D.; Oudet, P.; Jaeck, D.; Gaub, M. P. *Int. J. Cancer* **2006**, *120*, 524–32.
- (5) Kinzler, K. W.; Vogelstein, B. *Cell* **1996**, *87*, 159–70.
- (6) De Jong, W. D.; Bakker Schut, T. C.; Maquelin, K.; van der Kwast, T.; Bangma, C. H.; Kok, D. J.; Puppels, G. J. *Anal. Chem.* **2006**, *78*, 7761–69.
- (7) Lasch, P.; Haensch, W.; Naumann, D.; Diem, M. *Biochim. Biophys. Acta* **2004**, *1688*, 176–86.
- (8) Krishna, C. M.; Sockalingum, G. D.; Bhat, R. A.; Venteo, L.; Kushtagi, P.; Pluot, M.; Manfait, M. *Anal. Bioanal. Chem.* **2007**, *387*, 1649–56.
- (9) Steller, W.; Eienkel, J.; Horn, L. C.; Braumann, U. D.; Binder, H.; Salzer, R.; Krafft, C. *Anal. Bioanal. Chem.* **2006**, *384*, 145–54.
- (10) Fernandez, D. C.; Bhargava, R.; Hewitt, S. M.; Levin, I. W. *Nat. Biotechnol.* **2005**, *23*, 469–74.
- (11) Krishna, C. M.; Kegelaer, G.; Adt, I.; Rubin, S.; Kartha, V. B.; Manfait, M.; Sockalingum, G. D. *Biopolymers* **2006**, *82*, 462–70.
- (12) Lasch, P.; Haensch, W.; Lewis, E. N.; Kidder, L. H.; Naumann, D. *Appl. Spectrosc.* **2002**, *56*, 1–9.
- (13) Lasch, P.; Chiriboga, L.; Yee, H.; Diem, M. *Technol. Cancer Res. Treat.* **2002**, *1*, 1–7.
- (14) Rigas, B.; Morgello, S.; Goldman, I. S.; Wong, P. T. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8140–44.

* To whom correspondence should be addressed. Phone: 33 3 26 91 35 67. Fax: 33 3 26 91 35 50. E-mail: pierre.jeannesson@univ-reims.fr.

[†] MéDIAN UMR CNRS 6237, Université de Reims Champagne-Ardenne.

[‡] UMR-S682 Inserm, Université Louis Pasteur.

[§] Laboratoire de Biologie Moléculaire et Biochimie, CHRU Strasbourg, Hôpital de Hautepierre.

^{||} Service d'Anatomo-Pathologie, CHRU Strasbourg, Hôpital de Hautepierre.

significant over 80% accuracy.^{15–17} Lasch et al.^{7,12,18} demonstrated FT-IR microspectroscopy to be a tool for histopathological analysis by combining spectral analysis to tissue section mapping and by exploiting spectral data using statistical multivariate analysis methods. This was carried out on cryosections of colon biopsies, which permitted distinction of non-neoplastic crypt cells from adenocarcinoma cells. These studies demonstrated the strength of vibrational spectroscopy, i.e. the ability to detect, in a fully automated way, the presence of malignancy. But they also illustrated some limitations: the spectral differences, such as found by Rigas et al.,¹⁴ are hard to relate to specific biological processes and events. By choosing to reproduce classical histology with spectral imaging and multivariate data analysis, as in the study of Lasch et al.,⁷ the additional information that might be present in the spectral data, i.e. chemical composition, is not fully exploited.

New tools are required to further analyze the level of intratumor heterogeneity and its prognostic or predictive value. Thus, using a collection of xenografted colon carcinomas, our study aimed at a further exploration of the possibilities of the FT-IR approach for histological recognition, with a particular emphasis on tumor heterogeneity. The tumors analyzed are from xenografted human colon carcinomas from different patients partly described before by Guenot et al.³ Xenografting of human colon tumors in nude mice bridges the gap between cell line based models and patient material, offering the complexity of the original tumor with the experimental freedom of the cell lines. Xenografted colon tumors have been shown to maintain their three-dimensional morphology and their genetic characteristics over multiple passages.³ In addition, as in xenografting only tumor cells survive, all other cells implicitly originate from the host.

In our approach, we develop image analysis computational methods and demonstrate how general tumor growth patterns can be extracted from the IR spectral images, collected directly on paraffin-embedded tissue sections. Our findings allow access to quantitative spatially resolved information on intratumor heterogeneity that cannot be obtained by conventional histology.

EXPERIMENTAL SECTION

Xenografts. The xenografts and the xenografting protocol were described elsewhere.³ The xenograft experiments were approved by the CCPRB d'Alsace No. 1 Strasbourg, France ethics committee. Twenty-two xenografts of colon carcinomas and liver metastasis with a staging ranging from T2N0M0 to T4N1M1, from 12 male and 7 female patients, were included in this study. The frequency of chromosomal alterations based on the allelotyping on 33 microsatellite markers ranged from 6 to 97%.³ From 4 patients, 2 different passages of the xenografts were included. Clinical and histological characteristics are summarized in Table S1 in the Supporting Information. Xenografts from 9 patients (11 samples) were used to generate the predictive models for histology (so-called training set), and the other samples (9 patients,

11 samples) were used to validate the generated models (so-called validation set). The sets were chosen so that both groups are as equivalent as possible based on sex, tumor location, tumor stage and allelotyping.

IR Spectroscopy. Formalin-fixed paraffin-embedded xenografted tumors were cut to 7 μm thick slices and put on 2 mm thick ZnSe substrates (Crystal GmbH, Berlin, Germany). No attempt was made to remove the paraffin from the tissues. The consecutive slice was put on standard glass and hematoxylin eosin (HE) staining performed for conventional histology. IR spectral images were measured on a Spotlight 300 system (Perkin-Elmer Life Sciences, Courtaboeuf, France) with a spatial resolution of approximately 10 μm (6.25 μm step size). The spectral range of the spectra was 648–2000 cm^{-1} , at 8 cm^{-1} resolution with 2 measurements per spectrum. For each slice, an area containing only paraffin was also measured at the same conditions as the tissue spectra. Water vapor spectra were measured on the Perkin-Elmer Spectrum One detector which is part of the Spotlight 300 system. For that, the sample compartment was first flushed with nitrogen to remove all traces of water and carbon dioxide. A background spectrum was measured under this condition. The flushing was then stopped and air from the surrounding was allowed to diffuse into the sample compartment. During that rise in humidity, measurements were taken continuously until to reach a stable end point; no change was observed any longer. The experiment was repeated several times, leading to a total number of 58 spectra collected. The paraffin and water vapor spectra were used to correct the tissue spectra for the presence of these interferences. In addition, all the experiments were performed at constant room temperature (22 °C).

The total number of spectral images collected was 68, with, in average, an image size of 135,000 spectra, covering a substantial part of each slice. This whole data set comprises more than 9 million spectra.

Data Processing. Extended Multiplicative Signal Correction (EMSC)-Based Preprocessing. Several sources and factors contribute to the measured IR spectra. Not all contributions are equally relevant; some of them even mask the information one is looking for. The most important ones for this study are listed below:

Atmospheric water vapor gives narrow absorption lines in the spectral ranges 1300–1900 cm^{-1} . The intensity depends mainly on the humidity.

The paraffin used to embed the tissue gives bands at 888, 1377 and 1464 cm^{-1} . Each band may be composed of several overlapping bands and the relative intensity of these bands may be influenced by the paraffin embedding protocol or by the presence of other material such as tissue.

Scatter effects, such as Mie scatter, and interference effects give rise to wavelength dependent backgrounds in the spectra.

The sample thickness is directly related to the infrared absorption according to the Lambert–Beer law. But the sample thickness is not necessarily related to the sample composition.

EMSC is an extension of multiplicative scatter correction (MSC) introduced in 1983 for near-infrared spectra processing.^{19,20} MSC, further developed for IR data, offers an integrated approach

(15) Argov, S.; Salman, J. R. A.; Goldstein, I. S. J.; Guterman, H.; Mordechai, S. *J. Biomed. Opt.* **2002**, *7*, 248–54.

(16) Sahu, R. K.; Argov, S.; Bernshtain, E.; Salman, A.; Walfisch, S.; Goldstein, J.; Mordechai, S. *Scand. J. Gastroenterol.* **2004**, *6*, 557–66.

(17) Salman, A.; Argov, S.; Ramesh, J.; Goldstein, J.; Sinelnikov, I.; Guterman, H.; Mordechai, S. *Cell Mol. Biol.* **2001**, *47*, 159–64.

(18) Lasch, P.; Wäsche, W.; McCarthy, W. J.; Müller, G.; Naumann, D. In *Infrared spectroscopy: new tool in medicine*; Mantsch H. H., Jackson, M., Ed.; Proc SPIE: Bellingham, WA, 1998; 3257, pp 187–98.

(19) Martens, H.; Jensen, S. A.; Geladi, P. Multivariate linearity transformation for near-infrared reflectance spectrometry; Christie, O. H. J., Ed.; Proc. Nordic. Symp. Applied Statistics, Stokkand Forlag Publishers: Stavanger, 1983; pp 205–34.

to minimize the influence of all these factors.^{21,22} In our study, we have adapted EMSC to the processing of paraffined tissue IR data. Briefly, each spectrum \mathbf{s} , considered as a vector, in the data set can be modeled as a linear combination of known interference spectra, estimates of backgrounds and an estimate of the data set:

$$\mathbf{s} = a\hat{\mathbf{s}} + \mathbf{bI} + \mathbf{cP} + \mathbf{e} \quad (1)$$

with a (scalar), \mathbf{b} and \mathbf{c} (both vectors) regression coefficients, $\hat{\mathbf{s}}$ an estimate of the data set, \mathbf{I} the matrix of known interference spectra, \mathbf{cP} a polynomial modeling background and offsets due to scatter effects and interferences and \mathbf{e} the model error. If there is not a good estimate of the spectrum $\hat{\mathbf{s}}$ available, one could choose the average of the whole data set for $\hat{\mathbf{s}}$.

The corrected spectrum is now calculated as

$$\mathbf{s}_{\text{EMSC}} = \mathbf{e}/a + \hat{\mathbf{s}} \quad (2)$$

Thus, the variance due to interference and scatter effects is removed and the spectra are normalized to have comparable intensity as $\hat{\mathbf{s}}$. In the case of a large and complicated set of interference spectra, this set can be modeled using a principal component analysis (PCA) on non-mean-centered data. This simplifies the model of interferences and increases the numerical robustness of the calculation.

Histology by Spectroscopy: k-Means and Linear Discriminant Analysis (LDA). A k -means algorithm²³ was applied on a subset, so-called training set, of the measured spectral images to find spectral clusters. This clustering was compared with the HE stained sections to make a histological assignment for each cluster. The number of clusters was chosen so that each cluster has a unique histological assignment; multiple clusters with the same histological assignment were allowed. This k -means classification was subsequently used to generate a predictive LDA model enabling the classification of samples not included in the training set. Prior to the LDA, the dimensionality of the data was reduced by PCA. The LDA model was validated using the samples in the validation set.

Software Platform. Data were processed using Matlab 2006a (Mathworks, Natick, MA) based software. Routines were partly written in house and partly written at the Centre for Optical Diagnostics and Therapy of the Erasmus Medical Centre (Rotterdam, NL). The used LDA routines were from the Discriminant Analysis Toolbox 0.3, an implementation of the algorithms as published by Ripley.²⁴ The PCA routine was from the PLS Toolbox version 1.5.3 of Eigenvector (Wenatchee, WA).

RESULTS

Preprocessing of Spectra of Paraffin-Embedded Sections of Human Colon Xenografts. Spectra were cut to a spectral range of 724–1720 cm^{-1} which is assumed to be the most

informative spectral region for biological materials.²⁵ As explained in the materials and methods, paraffin and water vapor are significant undesirable contributions to the spectra measured. As an estimate of the measured spectra, necessary for the EMSC correction, the mean of the spectra in the training set was taken. Paraffin proved to give spectra with a highly variable background, probably due to interferences. Instead of using a single spectrum for paraffin only, it was judged that the first 8 principal components (PC) of a set of spectral images of paraffin would give a better description of the variations encountered. More elaborate discussion about the model construction was given by Ly et al.²⁶ For this experiment, a number of 8 different slices, measured during different measurement sessions, provided a total of 255,000 spectra. This set of spectra was reduced to 8 PC after scaling each spectrum to zero mean and unit variance. This set of 8 PC, obtained from non-mean-centered data, was included in the paraffin model to capture the variation of experimental conditions. Paraffin spectra with high backgrounds (above 0.35 absorbance) and spectra with a maximum absorbance less than 0.3 (low signal) of the main paraffin band or exceeding 1 (saturation of the detector) were considered aberrant and excluded from the calculation. In addition to the paraffin PC, the first 2 PC of a set of 58 water vapor spectra were added as well to take into account the day to day variation of atmospheric water vapor. This set of PC was used as interference spectra in the EMSC.

First results (not shown) indicated that, if paraffin and tissue were found in conjunction, the relative intensity of spectral bands of paraffin may change in a way not covered by the paraffin principal components. These spectral features could be due to molecular interactions between paraffin and tissue.²⁷ To compensate for that, a set of Gaussian curves, calculated from a band decomposition of the spectral bands of paraffin, was added to the set of interference spectra. The full set of interferences taken into account is shown in Figure 1A and B. Finally, to model offsets and Mie scatter effects resulting in slowly curving backgrounds, a fourth order polynomial was included as well. This approach was used, instead of that described by Mohlenhoff et al.²⁸ or Kholer et al.²⁹ because it permits to speed up the calculation, and was proved to be effective on the spectral region of interest (724–1720 cm^{-1}).

Tissue spectra to be included in the data analysis were selected, based on the three following rules: 1. The sum of squares of \mathbf{e} , a measure for the fit error, must be smaller than an empirical cut off. Effectively noisy spectra and spectra with very deviant shapes are excluded based on this rule. 2. The first fit coefficient must be larger than a certain value larger than zero. It is the requirement to ensure that a spectrum contains a significant amount of tissue signal. 3. The other fit coefficients must be inside a range ± 3 times the standard deviation around the mean for

(20) Geladi, P.; McDougall, D.; Martens, H. *Appl. Spectrosc.* **1985**, *39*, 491–500.

(21) Martens, H.; Stark, E. J. *J. Pharm. Biomed. Anal.* **1991**, *9*, 625–35.

(22) Kohler, A.; Kirschner, C.; Oust, A.; Martens, H. *Appl. Spectrosc.* **2005**, *59*, 707–716.

(23) Jain, A. K.; Dubes, R. C. Clustering methods and algorithms. In *Algorithms for clustering data*; Prentice Hall: Englewood Cliffs, 1988; pp 55–142.

(24) Ripley, B. D. Linear discriminant analysis. In *Pattern recognition and neural networks*; Cambridge University Press: Cambridge, 1996; pp 91–120.

(25) Parker, F. S. *Applications of infrared, Raman, and resonance Raman spectroscopy in biochemistry*; Plenum Press: New York, 1983.

(26) Ly, E.; Piot, O.; Wolhuis, R.; Durlach, A.; Bernard, P.; Manfait, M. *Analyst* **2008**, *133*, 197–205.

(27) Vrabie, V.; Gobinet, C.; Piot, O.; Tfyli, A.; Bernard, P.; Huez, R.; Manfait, M. *Biomed. Signal Process. Control* **2007**, *2*, 40–50.

(28) Mohlenhoff, B.; Romeo, M.; Diem, M.; Wood, B. R. *Biophys. J.* **2005**, *88*, 3635–40.

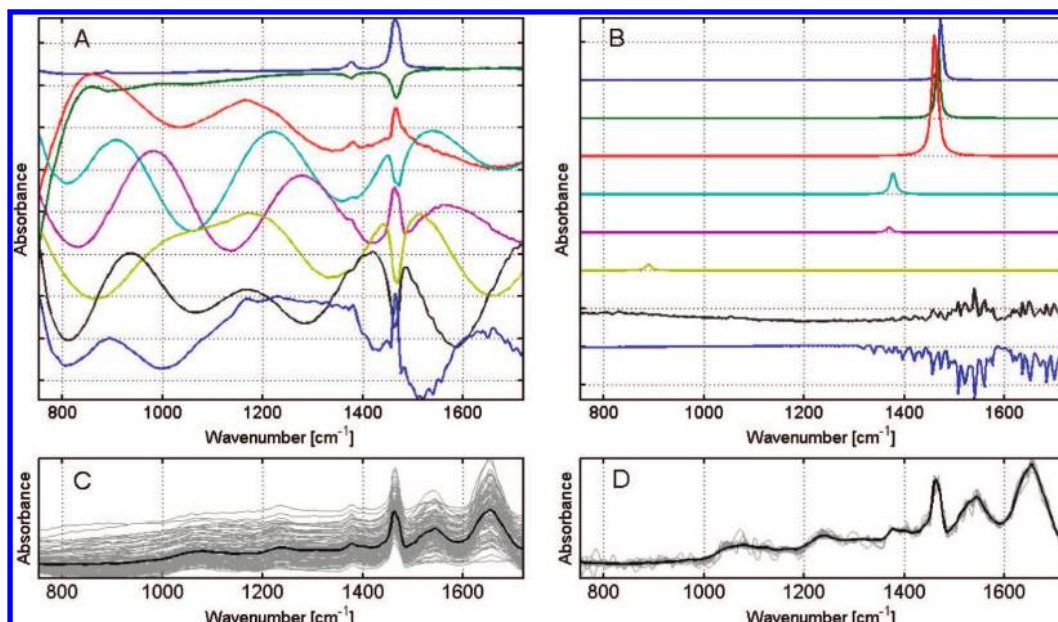


Figure 1. Preprocessing of IR spectra of formalin-fixed paraffin-embedded tissue sections of xenografted human colon carcinomas. The preprocessing was performed using extended multiplicative signal correction (EMSC) as described in the Experimental Section. A. From top to bottom, the first 8 principal components of paraffin signal used as interference spectra in the EMSC model. B. From top to bottom, 6 additional Gaussian curves modeling paraffin bands and the first 2 principal components of water vapor spectra. C. Untreated raw spectra collected on a tissue section (average spectrum drawn in bold line). D. Spectra after EMSC: backgrounds are removed, variabilities due to paraffin and water vapor are reduced; and outlier spectra without tissue features are removed.

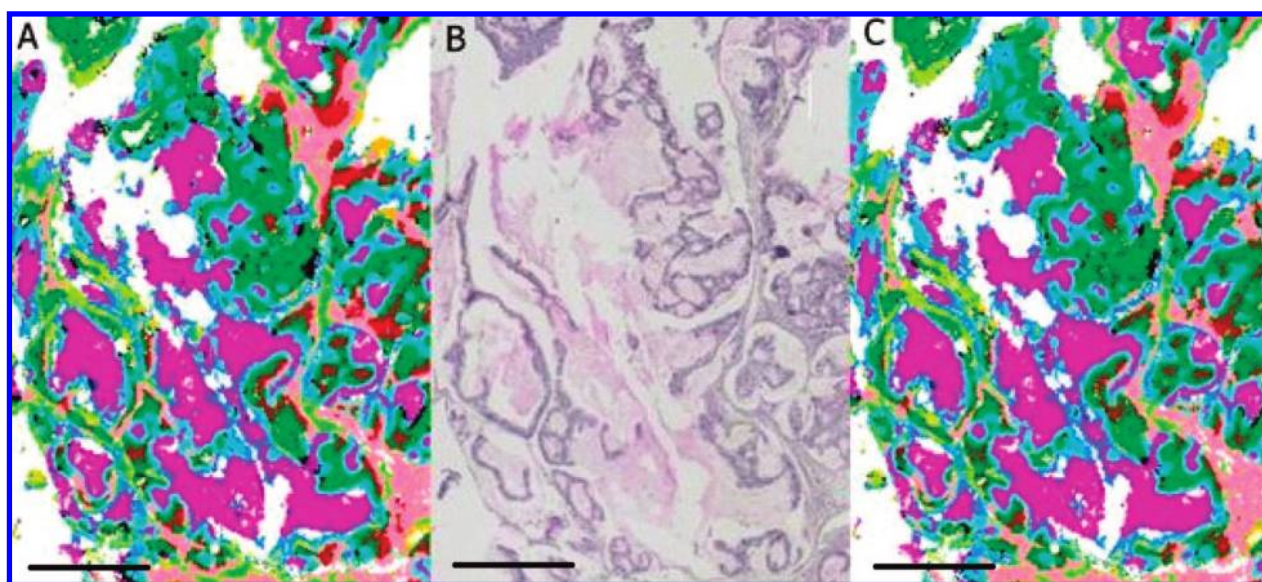


Figure 2. IR color coded spectral images of a tissue section of xenografted colon carcinoma. A. IR image from a tissue sample (sample #212; cf. Table S1 in the Supporting Information) constructed using *k*-means clustering of the spectral data after EMSC preprocessing as shown in Figure 1. B. Hematoxylin eosin (HE) stained section corresponding to the measured IR image. C. IR image of the same tissue section constructed by predictive LDA (linear discriminant analysis) model. Note slight differences between *k*-means clustering (A) and LDA (C). Scale bar, 400 μm .

each coefficient. This rule also excludes spectra with very deviant shapes.

The effect of the preprocessing is shown in Figure 1C and D: Spectra are normalized, backgrounds are corrected for, interferences are minimized and outliers (spectra that fail toward the three rules) are removed.

(29) Kohler, A.; Sulé-Suso, J.; Sockalingum, G. D.; Tobin, M.; Bahrami, F.; Yang, Y.; Pijanka, J.; Dumas, P.; Cotte, M.; Martens, H. *Appl. Spectrosc.* **2008**, *62*, 259–266.

Construction of Spectral Images and Correlation with HE Staining. Spectra of the training set were clustered into 12 clusters using a *k*-means algorithm. The *k*-means algorithm partitions the data in the given number of clusters, so that the sum of the distances of the data points to their respective cluster centers are minimized. The method is, apart from the number of clusters, unsupervised. The algorithm is especially suited to cluster large data sets such as the spectral images in this study.

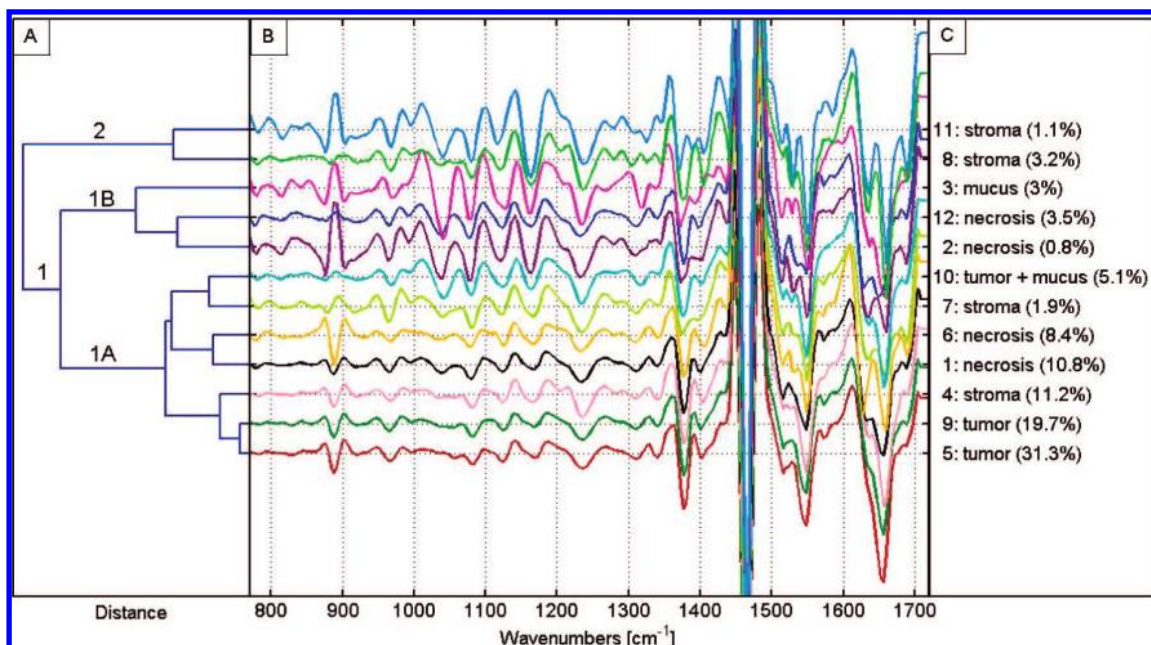


Figure 3. IR signatures of spectral clusters determined by *k*-means classification. A. Dendrogram obtained by hierarchical classification analysis of the spectral clusters. The dendrogram allows assessment of the interspectral distances (calculated using Ward's algorithm). B. Average spectra of the clusters. Second derivative spectra are shown to enhance the spectral features; the colors correspond to the colors of the cluster images (Figure 2A). C. Histogramical assignment of the 11 spectral clusters. The percentages indicate the proportions of pixels belonging to each cluster. These percentages were calculated from the set of IR images constructed by *k*-means classification on the samples of the training set.

Based on the *k*-means clustering, pseudocolor images were constructed by assigning a different color to each cluster (Figure 2A). White areas correspond to spectra (outliers) omitted in the analysis based on the rules listed above. The pseudocolor images allow a comparison with the HE stained slices (Figure 2B). A specific histological structure could be assigned to each cluster using all images in the training set. A number of 12 clusters allows for a unequivocal reproduction of HE based histology.

In Figure 3, the assignment to histology as identified from HE stained slices is given together with the average spectrum of each cluster, as well as a dendrogram visualizing the distance between the *k*-means cluster averages. Spectra are plotted as second order derivatives to enhance spectral differences. The dendrogram was calculated on the Euclidean distance between the cluster averages using Ward's algorithm for hierarchical cluster analysis.

Based on the generated images, the following histological assignment was made: 3 clusters for tumor tissue, 4 clusters for stromal tissue, 4 clusters for necrotic tissue and 1 cluster for excreted mucus. At this point, it has to be reminded that all stromal tissues originate from the host. Interestingly, it was found that the tumor clusters 5, 9 and 10 may appear simultaneously in a single sample, thus revealing an intratumor heterogeneity. Branch 1A of the dendrogram indicates that the two tumor clusters 5 and 9 are the closest and linked to cluster 4 assigned to stroma.

These three clusters represent 62% of the measured spectra in the training set of the samples. In addition, inside branch 1A, tumor cluster 10 is separated from the other tumor clusters 5 and 9. Its spectral signature in the range 1000–1100 cm⁻¹ reveals the presence of mucin³⁰ confirming the assignment of this cluster to tumor areas containing mucus producing cells. Branch 1B is composed of cluster 3 (accumulated mucus without cells) and of a subcluster formed by clusters 2 and 12 (both necrotic tissue).

Spectra corresponding to the necrotic tissue belong to separate branches (clusters 1 and 6 in branch 1A, clusters 2 and 12 in branch 1B) reflecting different degrees of necrosis. Besides branch 1, the dendrogram is composed of branch 2, containing clusters 8 and 11 both assigned to stroma.

In conclusion, these data demonstrate the presence of several clusters with the same histological assignment. This reveals a marked heterogeneity in the biochemical composition of each tissue subtype, supporting additional levels of tissue organization, that are undetectable by conventional HE staining.

Training and Validation of IR-Imaging-Based Histology.

In order to automatically classify spectra from new samples according to the established cluster membership, LDA,²⁴ a supervised classification method, was used to transform the *k*-means results into a predictive model. Besides classification, LDA also gives the probability that a spectrum belongs to a particular class. A spectrum is then assigned to the cluster with the highest probability. In addition, this probability information can be useful to assess the relevancy of the classification obtained. Together with the established histological assignment of the *k*-means clusters, the LDA model allows for an IR-imaging-based histology for new samples. The LDA model has to be validated using samples from the validation set. This validation on independent data is necessary to prove that the calculated model performs well. For computational reasons, the LDA model was built using 10% of the available spectra in the training set; every tenth spectrum was included. The EMSC preprocessing described before was applied to the spectra. In addition, the dimensionality of the chosen data set was reduced by applying PCA and taking only the first 25 PC into account.

(30) Khajehpour, M.; Dashnau, J. L.; Vanderkooi, J. M. *Anal. Biochem.* 2005, 348, 40–48.

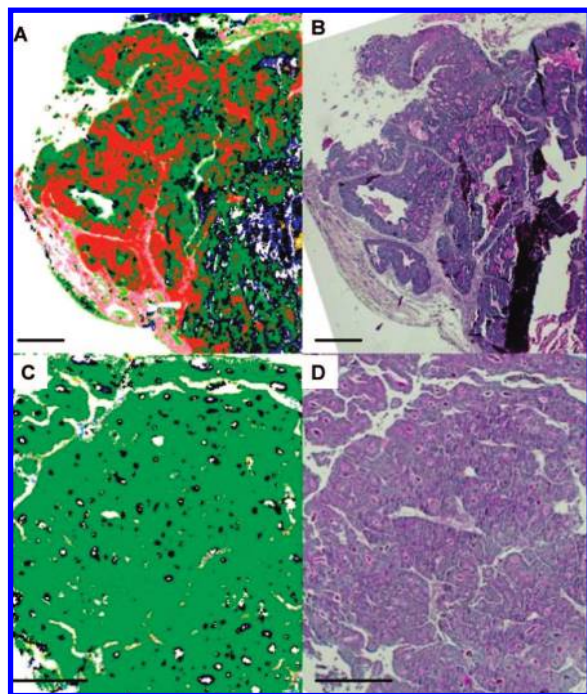


Figure 4. IR color-coded images of two samples of the validation set obtained by the LDA model. A. IR image of the sample #250. B. HE stained section corresponding to the calculated image in A. C. IR image of the sample #314. D. HE stained section corresponding to the calculated image in B. (See Table S1 in the Supporting Information for description of samples #250 and #314). Scale bar, 400 μm .

First, the validation of the LDA model consisted in comparing the outcomes of the LDA with the *k*-means results. As shown in Figure 2A and C, the LDA predictive model leads to reconstruct a color coded image highly similar to the *k*-means image. Only slight discrepancies appear mainly on the margins of the clusters, corresponding to pixels with a probability less than 0.75 of belonging to the assigned cluster indicating that these margin pixels have intermediate spectral signatures. Usually, the other considered pixels have a probability higher than 0.9.

Second, the predictive value of the LDA model was assessed by comparing the LDA results of the validation set with the HE-based histology. Data of Figure 4 and of Figures S1–S4 in the Supporting Information indicate that the LDA predicted histology for the validation set was totally consistent with the conventional staining procedure. Repeated measurements on the same sample area with a time interval of a few weeks proved the repeatability of the IR measurements and the data processing (data not shown). Figure S5 in the Supporting Information depicts the relevancy of the spectral clustering by showing the class membership probabilities for the two closest clusters (#5 and #9). This significance of clusters 5 and 9 was demonstrated for three characteristic samples: #241 and #314 presenting mainly the clusters 5 and 9 respectively, and the mixed sample #250 containing both clusters.

In addition, since the validation set included different passages from the same xenografted tumor, the IR analysis of these samples indicate stability over sequential passages (Figure 4). This stability at the level of the spectral-based histology confirms previous results showing that the xenografted tumors are stable, not only in their morphology but also in the chromosomal alterations.³

Indeed, the spectral images show that the relative contribution of the tumor clusters 5, 9 and 10 is similar between the different passages.

Spectral Analysis of the Tissue Organization and Identification of Tissue Patterns. As illustrated by the assignment of two clusters (5 and 9) to tumor tissue, tissue organization not revealed by HE stained images, was observed on IR images. This organization was analyzed quantitatively by applying image analysis methods at the pixel level and by analyzing the spectral differences. The data of the validation set and the training set were pooled for this analysis and the cluster assignment was calculated using the LDA model. White areas in the images were omitted in the calculation. For this approach, histograms (Figure 5A) presenting the chance for a given cluster to find another cluster in its neighborhood were calculated as follows: for each pixel in the images, the cluster membership of the direct neighbors (left, right, above, under) was determined. If a neighboring pixel belonged to the same cluster, the event was omitted. The result per cluster was normalized and plotted in a set of histograms. Each histogram gives the probability distribution for the cluster membership of the neighboring cluster. For example: cluster 5 borders, on average, for 32% with cluster 4 and 63% with cluster 9 but hardly with any other cluster. From these histogram columns, sequences or morphological patterns were characterized by using the following rules: start from the column of a particular cluster (i.e., cluster 4), determine the cluster with the highest probability of bordering (i.e., cluster 5), go to the column of this last cluster and determine in this column the most probable bordering cluster (i.e., cluster 9) as previously (if this cluster is the same as the preceding one, take the second highest) and continue until an end point is reached. For example, starting from this cluster 4, the most abundant stroma cluster, this gives the following sequence: 4 \rightarrow 5 \rightarrow 9 \rightarrow 1 \rightarrow 12.

Using all clusters as a starting point, 12 different sequences can be found. Comparing all the sequences, it is observed that they all have the sequence 4 \rightarrow 5 \rightarrow 9 in common. Considering that stroma cluster 8 borders almost exclusively with cluster 4, the full sequence may start with cluster 8: 8 \rightarrow 4 \rightarrow 5 \rightarrow 9 \rightarrow 1 \rightarrow 12. This pattern suggests that stroma clusters 4 and 8 and tumor cluster 5 are not in direct contact with necrotic tissues 1 and 12, but always buffered by tumor cluster 9. Another sequence of interest, however less probable, is the one involving cluster 10 as it is also a tumor tissue and mucus cluster 3 as it represents the mucus producing tumors: 8 \rightarrow 4 \rightarrow 5 \rightarrow 9 \rightarrow 10 \rightarrow 3. These schematic patterns can be visualized in simple diagrams (Figure 5B).

The spectral origins of the sequence involving the predominant tumor clusters (5 and 9) and neighboring clusters (1 and 12 assigned by HE staining as necrotic tissue) were also investigated (Figure 6). For this, difference spectra between the successive clusters 5 \rightarrow 9 \rightarrow 1 \rightarrow 12 were calculated in such a way that no negative spectral features appear. In addition, to better appreciate the differences, second derivatives of these difference spectra are shown as well. The difference spectrum between clusters 5 and 9 (blue spectrum) puts in evidence bands at 996, 1116, 1240 cm^{-1} (red vertical lines), associated with RNA.²⁶ This observation indicates that cluster 5 is richer in RNA, which is consistent with a higher transcriptional activity in cluster 5 than in cluster 9.

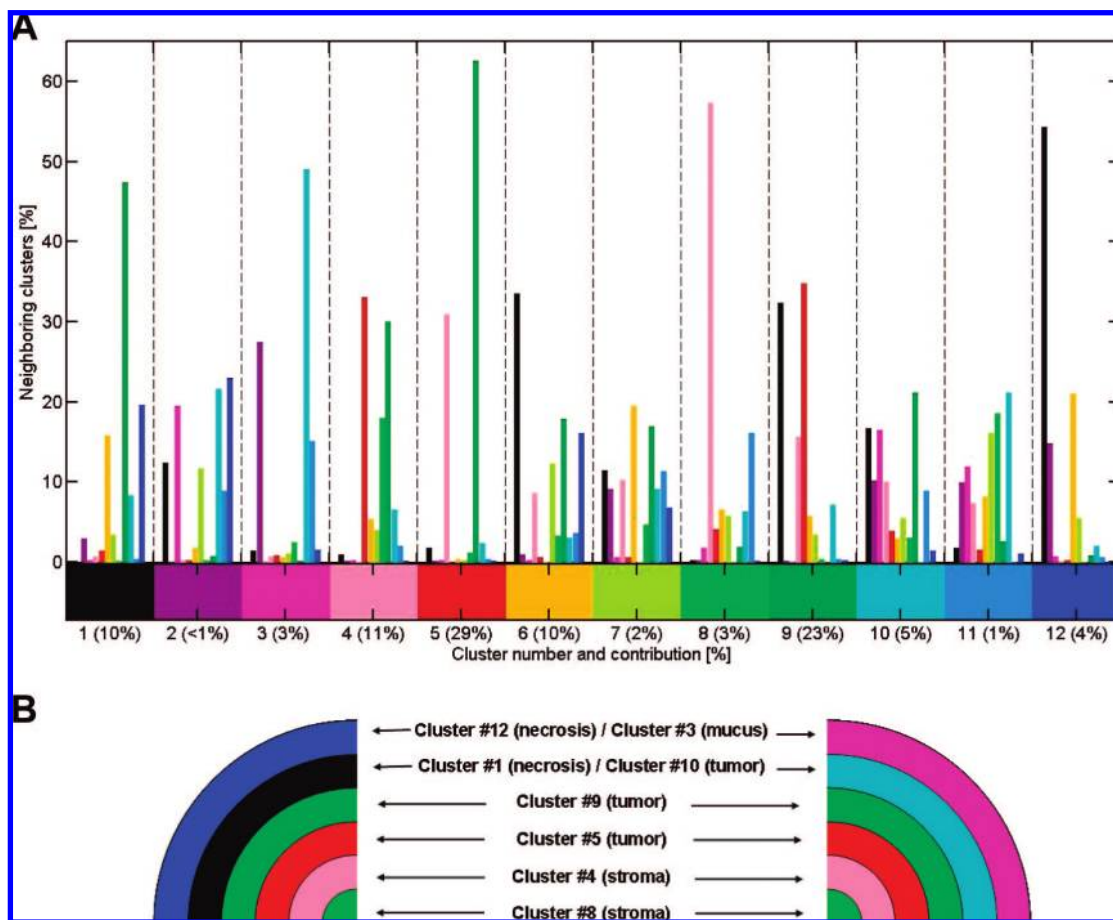


Figure 5. Spatial ordering of the spectral clusters. A. Histograms of the cluster bordering. The percentages of neighboring pixels were calculated from IR images constructed by the LDA model and by considering the totality of the analyzed samples (training and validation sets). The histogram colors correspond to the cluster colors, recalled on the color scale. The contribution of each cluster as a percentage of the total number of spectra is also indicated. B. Schematic representation of two main tumor tissue patterns including tumor clusters #5, 9 and 10. The rules used to characterize the sequences of these morphological patterns are indicated in the Results section, paragraph “Spectral Analysis of the Tissue Organization and Identification of Tissue Patterns”. The first pattern (left) includes necrotic tissues as observed in Figure S4A in the Supporting Information, the second pattern (right) includes mucinous tissue as observed in Figure 2C.

Concerning DNA, specific vibrations do not appear on this difference spectrum indicating a constant content in DNA between the two tumor clusters 5 and 9. The difference spectra between clusters 9 and 1 (green spectra) and clusters 1 and 12 (red spectra) exhibit vibration bands specific of DNA (bands at 1055, 1088 and 1718 cm^{-1}).³¹ The decrease of DNA bands indicates a progressive loss of DNA between the tumor cluster 9 and the necrotic tissue clusters 1 and 12.

In conclusion, both the analyses of cluster sequences and of spectral differences via IR imaging allow measures of intrinsic molecular changes and reveal a marked heterogeneity in the tumor tissue organization, otherwise not identifiable by conventional histology.

DISCUSSION

In our study, an IR-imaging-based histology of human xenografted colon tumors was established using *k*-means clustering. Based on this *k*-means clustering, a predictive LDA model was generated allowing the prediction of the histology of new samples. Using this approach, paraffin-embedded tissue sections can be

segmented into identifiable regions of differing pseudocolors. This segmentation was obtained by classification of spectral data into 12 clusters, each of them representing a set of spectra with similar vibrational features and assigned to a single histological structure. On this basis, the proposed IR-based histology reflects, at the micrometric scale, the intrinsic chemical composition of the tissue and consequently brings additional information; it does not require dyes or molecular probes unlike the conventional morphology-based histology.

To our knowledge, this spectroscopic approach is the first to carry out direct spectral analysis of formalin-fixed paraffin-embedded specimens, without requiring chemical dewaxing. This method of numerical dewaxing avoids the tedious and time-consuming experimental steps of deparaffining that necessitates highly toxic organic solvents. In addition, compared to spectroscopic investigations led on cryosections,⁷ our analysis allows long-term retrospective studies that involve mostly paraffin-embedded archival samples. This opens the way to detection of spectral markers specific of the early stages of disease and with a prognosis value. The correction of paraffin spectral interferences takes into account the high variability of the IR paraffin signal by modeling it by means of several components (principal compo-

(31) Arakawa, H.; Neault, J. F.; Tajmir-Riahi, H. A. *Biophys. J.* 2001, 81, 1580–87.

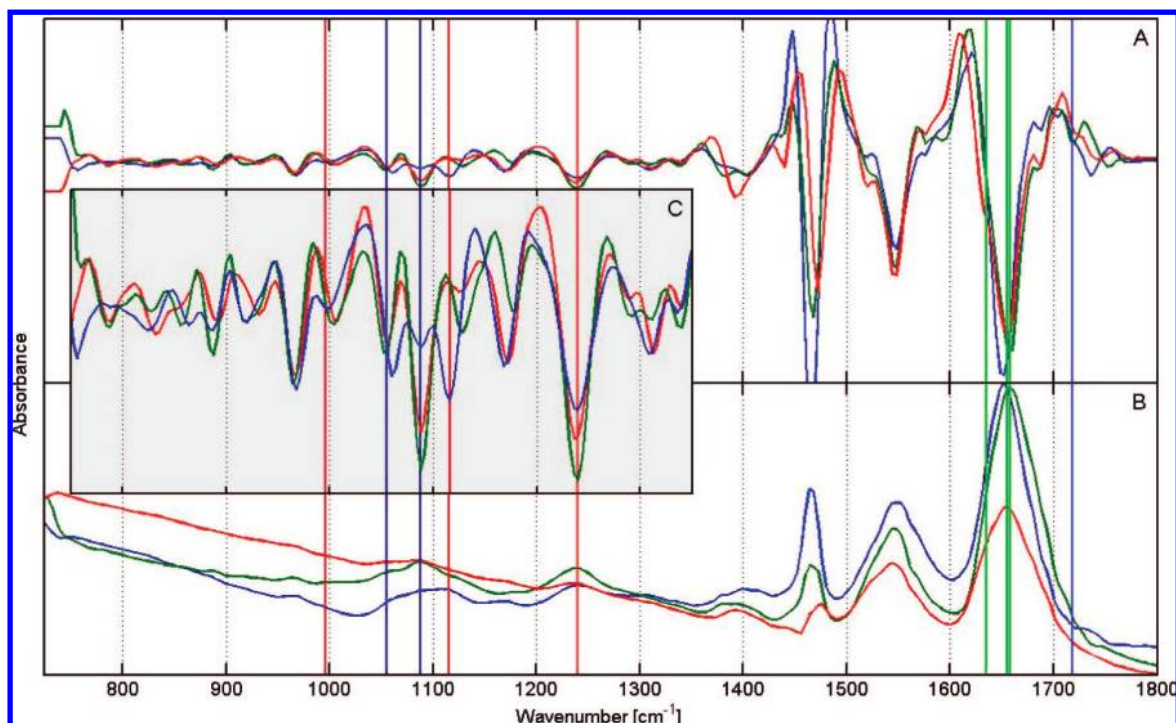


Figure 6. Spectral differences within the first cluster pattern (Figure 5B). A. Difference spectra between clusters 1 minus 12 (red), 9 minus 1 (green) and 5 minus 9 (blue). The subtractions were scaled such that no negative features appeared. B. Second order derivatives of the difference spectra. Vibration bands appear negative on these derivative spectra. C. Inset showing second derivative spectra magnified in the range 750–1350 cm^{-1} . Vertical lines indicate band assignments of RNA (red), DNA (blue) and proteins (green).

nents analysis and Gaussian curves). For specimens coming from different pathology laboratories, it is necessary to reconstruct an appropriate model of the paraffin signal as recently shown.²⁶

Here, IR spectral imaging demonstrates a marked intratumor heterogeneity, undetectable by conventional histological staining. This spectral heterogeneity within a tumor tissue has never been previously reported. Spectroscopic imaging is usually limited to reproducing classical histology allowing the identification of spectral markers associated only to each tissue subtype (i.e tumor, stroma, necrosis...). Here, three spectral clusters are found to be assigned with tumor tissue, clusters 5, 9 and 10, this last one being associated with mucin producing cells. These clusters are present simultaneously in one single sample but at variable degrees. The development of an innovative treatment of the cluster images leads to additional information on the encountered tumor heterogeneity. This image analysis is based on the concept of pixel connectivity, to quantify the chances that two clusters are in direct contact with each other. The probability distributions reflect the general tissue organization in the analyzed images. From these distributions, two main general patterns of tissue organization (Figure 4B) were identified: one revealing a gradient from stromal tissue, to tumor and then to progressive grades of necrosis, and one from stroma to tumor, to mucin containing tumor and finally to excreted mucus. In particular, tumor cluster 5 is most likely found next to stromal tissue and, contrary to tumor cluster 9, it is almost never found in direct contact with necrotic tissue. This organization would indicate that cluster 5 is associated to the more vital tumor tissue. This hypothesis can be reinforced by calculating spectral differences between clusters 5 and 9. Indeed, the difference spectrum reveals that cluster 5 is richer in RNA, pointing at a higher transcriptional activity in the associated tissue. The decrease of

DNA bands indicates a progressive loss of DNA in the cluster sequence $9 \rightarrow 1 \rightarrow 12$. This suggests the presence of different stages of progressive necrosis and could explain the identification of two distinctive spectral clusters (1 and 12) assigned to the necrotic tissue. In addition, different passages of the same tumor give IR images with a similar cluster distribution indicating that intratumoral heterogeneity was maintained over successive passages. Such results are in agreement with those of Guenot et al.³ reporting that the xenografted tumors maintain their genetic fingerprint and their tissue morphology over time. The pathophysiological significance of IR intratumoral heterogeneity needs to be more thoroughly investigated. Indeed, as recently suggested by Wang et al.,³² subtle IR features may reveal early molecular changes well before morphological changes occur. Further studies combining FT-IR microimaging for precise localization of tumor areas and laser dissection will permit complementary proteomic and genetic analyses on these regions of interest.

CONCLUSION

In conclusion, IR imaging in combination with spectral image analysis represents a potential avenue for a new type of histology, based on optical biopsies and more informative than conventional staining procedures. This methodology, independent of the intraobserver variability, shows great promises for applications in routine diagnostic medicine.

ACKNOWLEDGMENT

This study was supported by a grant of Institut National du Cancer (INCa), Canceropôle Grand Est. R.W. is a recipient of a

(32) Wang, T. D.; Triadafilopoulos, G.; Crawford, J. M.; Dixon, L. R.; Bhandari, T.; Sahbaie, P.; Friedland, S.; Soetikno, R.; Contag, C. H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15864–69.

postdoctoral fellowship from INCa. We thank Sylvie Record for linguistic assistance. A.T. is a recipient of a doctoral fellowship from Ministère de l'Enseignement Supérieur et de la Recherche.

SUPPORTING INFORMATION AVAILABLE

Clinical and histological characteristics of the studied human colon tumors in Table S1. In Figures S1–S4, IR color-coded images, constructed with the predictive LDA model, and corresponding HE staining sections of various samples, including two

xenografts from the same tumor after different passages (Figure S1). In Figure S5, class membership probabilities for clusters #5 and #9. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review June 12, 2008. Accepted September 11, 2008.

AC801191X

Publication 3

**Metabolic characterization of primary human colorectal cancers
using high resolution magic angle spinning ¹H magnetic
resonance spectroscopy**

Imperiale A, ElBayed K, Piotto M, Neuville A, Crille Blondet C, Brigand C,
Bellocq JP, Namer IJ.

Metabolomics 2009 5;292-301

Metabolic characterization of primary human colorectal cancers using high resolution magic angle spinning ^1H magnetic resonance spectroscopy

M. Piotto · F.-M. Moussallieh · B. Dillmann ·
A. Imperiale · A. Neuville · C. Brigand · J.-P. Bellocq ·
K. Elbayed · I. J. Namer

Received: 21 October 2008 / Accepted: 10 December 2008 / Published online: 24 December 2008
© Springer Science+Business Media, LLC 2008

Abstract Colorectal cancer is one of the most frequent and most lethal forms of cancer in the western world. The aim of this study is to characterize by ^1H high resolution magic angle spinning NMR spectroscopy (HRMAS) the metabolic fingerprint of both tumoral and healthy tissue samples obtained from a cohort of patients affected by primary colorectal adenocarcinoma. By analyzing HRMAS data using multivariate statistical analysis (PLS-DA), the two types of tissues could be discriminated with a high level of confidence. The identification of the metabolites at the origin of this discrimination revealed that adenocarcinomas are richer in taurine, glutamate, aspartate, and lactate whereas healthy tissues contain a higher amount of myo-inositol and β -glucose. The statistical model resulting from the PLS-DA analysis was subsequently used to perform a blind test on tumoral and healthy colon

biopsies. The results of the classification showed that the HRMAS analysis has very high sensitivity and specificity.

Keywords Colorectal cancer · Biopsies · HRMAS · Statistical analysis · Taurine · Myo-inositol

1 Introduction

Colorectal cancer is one of the most frequent cancers in the western world (Jemal et al. (2007); Libutti et al. (2008)). The most common colon cancer is adenocarcinoma which is a malignant epithelial tumor that originates from the glandular epithelium of the colorectal mucosa. Patient prognosis is based on several parameters like the clinical stage of the disease, the histological type and grade of the tumor (Cusack et al. (1996)) and the analysis of oncogenes and molecular markers (Fearon et al. (1990); Leichman et al. (1997)). In order to develop more efficient diagnostic and prognostic tools, current research focuses on both the genetic and the molecular basis of the disease.

^1H high resolution magic angle spinning (HRMAS) is a nuclear magnetic resonance (NMR) technique (Andrew et al. (1958); Lippens et al. (1999); Lowe (1959)) that allows the characterization of the metabolic phenotypes of intact cells, tissues and organs, under both normal and pathological conditions (Cheng et al. (1997); Cheng et al. (1998); Cheng et al. (2000); Martínez-Bisbal et al. (2004); Martínez-Granados et al. (2006); Sitter et al. (2002); Sjøbakk et al. (2008); Tzika et al. (2002)). HRMAS allows the assessment of cellular metabolic networks that are directly related to specific genetic information and to the regulation of specific gene transcripts. Several human gynecological, neurological, and urological malignancies have been studied by HRMAS with promising results.

M. Piotto (✉)
Bruker BioSpin, 34 rue de l'industrie, 67166 Wissembourg,
France
e-mail: martial.piotto@bruker.fr

M. Piotto · F.-M. Moussallieh · B. Dillmann · K. Elbayed
Institut de Chimie, Université Louis Pasteur, 4 rue Blaise Pascal,
67000 Strasbourg, France

F.-M. Moussallieh · B. Dillmann · A. Imperiale ·
I. J. Namer (✉)
Department of Biophysics and Nuclear Medicine, University
Hospitals of Strasbourg, 67098 Strasbourg, France
e-mail: Izzie.Jacques.NAMER@chru-strasbourg.fr

A. Neuville · J.-P. Bellocq
Department of Pathology, University Hospitals of Strasbourg,
67098 Strasbourg, France

C. Brigand
Department of Digestive Surgery, University Hospitals
of Strasbourg, 67098 Strasbourg, France

The aim of the present study is to characterize by ^1H HRMAS the metabolic fingerprint of both tumoral and healthy tissue biopsy samples obtained from a cohort of patients affected by primary colorectal adenocarcinoma. One of the objectives of the project is to develop and validate in a clinical set-up a statistical metabolic model capable of achieving a reliable classification of cancerous versus healthy biopsies. In order to obtain reliable statistical results, the study was performed on a relatively large cohort of patients ($n = 44$).

The study of the metabolic content of colon biopsies by magnetic resonance has been the topic of only a few papers in the literature. In fact, the study presented in this manuscript is, to our knowledge, the first one involving the direct comparison of the metabolic content of healthy and cancerous human colon biopsies using intact biopsy samples.

2 Materials and methods

2.1 Patient population

This study involves the use of human colon biopsies obtained from the Tumor bank of the University Hospitals of Strasbourg, France. Between March 2007 and December 2007, 44 patients (26 men, 18 women; mean age: 68 ± 12 years, age range: 45–90 years) with histologically proven colorectal adenocarcinoma were prospectively selected. All patients underwent a surgical resection of the primary lesion and a radical lymphadenectomy. Neither neo-adjuvant chemotherapy nor radiotherapy were performed prior to the operation. For the HRMAS analysis, two specimens containing respectively tumoral and healthy tissue were obtained from each patient. For four patients, only the healthy or the tumoral biopsy was available. The total number of biopsies was therefore equal to 84. Tissue specimens were collected with minimum ischemic delays after resection (average time 50 ± 25 min) and snap-frozen in liquid nitrogen before being stored at -80°C . All tissue samples exhibited a viable tumor/necrosis ratio and were quantitatively and qualitatively adequate to perform a correct HRMAS analysis.

2.2 HRMAS analysis

2.2.1 Rotor preparation

Each biopsy sample was prepared at -10°C by introducing 15 to 20 mg of biopsy into a disposable 30 μl Kelf insert. To provide a lock frequency for the NMR spectrometer, 10 μl of D_2O containing 1% w/w TSP were also added to the insert. The insert was then sealed tightly with a conical plug and stored at -80°C until the HRMAS analysis. The

insert insures that the entire biopsy sample is detected by the radio-frequency coil of the probe and that no leaks occur during the HRMAS analysis. Shortly before the HRMAS analysis, the insert was placed into a standard 4 mm ZrO_2 rotor and closed with a cap. The ensemble was then inserted into a HRMAS probe pre-cooled at 3°C . All HRMAS experiments were performed at 3°C and were started immediately after the temperature inside the probe had reached the equilibrium condition (5 min). Upon completion of the HRMAS analysis (20 min), the insert was taken out of the rotor and stored back at -80°C . This methodology prevents sample cross-contamination and allows storing all the biopsies in a reliable matter. A considerable amount of time is also saved by eliminating the cleaning process of the rotor. Furthermore, complementary NMR analysis can be performed at a later stage on any of the stored biopsy inserts.

2.2.2 Data acquisition

HRMAS spectra were recorded on a Bruker Avance III 500 spectrometer operating at a proton frequency of 500.13 MHz. This instrument is installed at the Hautepierre university hospital in Strasbourg and is dedicated to the analysis of biopsies by HRMAS. It is operated by qualified scientific and medical personnel in the context of the CARMen project which aims at the creation of an extensive metabolic database covering most current human tumors. The spectrometer is equipped with a 4 mm double resonance (^1H , ^{13}C) gradient HRMAS probe. A Bruker Cooling Unit (BCU) is used to regulate the temperature at 3°C by cooling down the bearing air flowing into the probe. All NMR experiments were conducted on samples spinning at 3,502 Hz in order to keep the rotation sidebands out of the spectral region of interest. For each biopsy sample, a one-dimensional proton spectrum using a Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence was acquired (Bruker *cpmgpr1d* pulse sequence). This sequence allows a better observation of the metabolites by reducing the intensity of lipid signals in the spectrum. The inter-pulse delay between the 180° pulses of the CPMG pulse train was synchronized with the sample rotation and set to 285 μs ($1/\omega_r = 1/3502 = 285 \mu\text{s}$) in order to eliminate signal losses to B_1 field inhomogeneities (Elbayed et al. (2005); Piotto et al. (2001)). The number of loops was set to 328 giving the CPMG pulse train a total length of 93 ms. In order to quantify, in an absolute manner, the amount of metabolites present in each sample, a calibrated Eretic signal (Akoka et al. (1999)) was included in all the 1D experiments. The Lorentzian Eretic signal was added digitally in the spectrum after data acquisition and processing. This procedure allows obtaining an Eretic signal which is perfectly in phase with the NMR signals and

which does not compromise the quality of the baseline. Using a standard solution of known concentration of lactate, the amplitude of the Eretic signal was calibrated to correspond to 0.62 μmol of protons. In order to accurately calculate the metabolite concentration in each biopsy, a correction factor taking into account the difference in pulse widths between the standard lactate sample and the different biopsies was also applied (Wider and Dreier (2006)). The CPMG experiment was acquired with the following parameters: sweep width 14.2 ppm, number of points 32 k, relaxation delay 2 s and acquisition time 2.3 s. A total of 128 FID were acquired resulting in an acquisition time of 10 min. All spectra were recorded in such a manner that only a zero phase order correction was necessary to properly phase the spectrum. The FID was multiplied by an exponential weighing function corresponding to a line broadening of 0.3 Hz prior to Fourier transformation. All spectra were processed using automatic base line correction routines. ^1H spectra were referenced by setting the lactate doublet chemical shift to 1.33 ppm.

In order to confirm resonance assignments, two-dimensional homonuclear and heteronuclear experiments were also recorded on eight samples. 2D DIPS12 spectra (Shaka et al. (1988)) were acquired with a 170 ms acquisition time, a 60 ms mixing time, a 14.2 ppm spectral width and a 1.5 s relaxation delay. Thirty-two transients were averaged for each of the 512 increments during t1, corresponding to a total acquisition time of 8 h. Data were zero filled to a 2 k * 1 k matrix and weighted with a shifted square sine bell function prior to Fourier transformation. 2D ^1H - ^{13}C HSQC experiments using echo-antiecho gradient selection for phase-sensitive detection (Davis et al. (1992)) were acquired using a 73 ms acquisition time with GARP ^{13}C decoupling and a 1.5 s relaxation delay. A total of 116 transients were averaged for each of 256 t1 increments, corresponding to a total acquisition time of 15 h. Two 1 ms sine-shaped gradient pulses of strength 40 and 10.05 G/cm were used in the experiment. Data were zero-filled to a 2 k * 1 k matrix and weighted with a shifted square sine bell function before Fourier transformation.

2.2.3 Multivariate statistical analysis

The spectral region between 4.7 and 0.5 ppm of each 1D CPMG NMR spectrum was automatically binned into regions of 0.01 ppm using the AMIX 3.8 software (Bruker GmbH, Germany). This procedure minimizes the effect of peak shifts due to pH variations. The peak integral within each 0.01 ppm region was computed and normalized with respect to the total integral of the spectrum in the 4.7–0.5 ppm region. This process generated an X data matrix containing 421 columns (chemical shifts) and 84 rows (corresponding to healthy and tumoral biopsies of 44

patients). Data sets were then imported into the SIMCA P 11.0 software (Umetrics AB, Umeå, Sweden) and pre-processed using unit variance scaling of the X columns by weighing each integral region by $1/SD_k$, where SD_k represents the standard deviation of the kth column in the X matrix. This procedure gives equal weight to metabolites present in high and low concentrations while minimizing the contribution from noise.

The X matrix was analyzed using principal component analysis (PCA) within the SIMCA P 11.0 software package. This procedure allows to quickly evaluate the quality of the data and to identify possible outliers. After the PCA analysis, Partial Least Square Discriminant Analysis (PLS-DA) (Tenenhaus (1998)) was conducted in order to build a statistical model that optimizes the separation between the two classes of patients. The number of components of the PLS-DA model was determined by cross-validation. The class membership of each sample was iteratively predicted, using the results to generate a goodness of fit measure ($Q^2 = 1 - \text{PRESS}/\text{SS}$) for the overall model. PRESS is the predicted squared sum of error and represents the squared differences between observed and predicted Y values when each sample is kept out of model development and SS is the residual sum of squares of the previous dimension. Since PLS-DA models are usually built using a relatively small number of samples compared to the number of variables, an extensive cross-validation of the model is mandatory to avoid overfitting of the data (Westerhuis et al. (2008)). The maximum theoretical value for Q^2 is equal to 1 for a perfect prediction. However, a Q^2 value superior to 0.5 is generally considered to be a decent predictor. In order to validate the model, a circular validation procedure was performed. The class labels of the biopsies (cancerous versus healthy) in the Y column matrix were randomly permuted and for each permutation, a new statistical model was computed along with its Q^2 value. This procedure was repeated for a large number of permutations and the Q^2 value were plotted as a function of the number of permutations in the Y matrix. A model was considered as valid if the regression line obtained by circular permutation showed a negative slope and if all the estimated Q^2 were smaller than the original one. This feature proves that none of the models computed on the permuted Y matrix are better than the model computed using the original Y matrix.

3 Results and discussions

3.1 Spectral assignment

Representative HRMAS CPMG magnetic resonance spectra of adenocarcinomas and healthy tissues originating from the

same patient are presented in Fig. 1. Metabolites were assigned using standard metabolite chemical shift tables available in the literature (Martínez-Bisbal et al. (2004)). 2D DIPSI2 and HSQC spectra were also used to confirm some assignments. A total of about 30 different metabolites were identified from these spectra (Table 1). The respective metabolic content of healthy and cancerous biopsies can be directly inferred from Fig. 1 since the intensity of each

spectrum was normalized with respect to the amplitude of the Eretic signal and the weight of biopsy present in each sample. Figure 1 tends to show that adenocarcinomas are characterized by a higher level of taurine while healthy tissues contain a larger amount of myo-inositol. It is important to emphasize that since both biopsies were obtained simultaneously from the same patient, the direct comparison of their metabolic content is extremely reliable since the time of

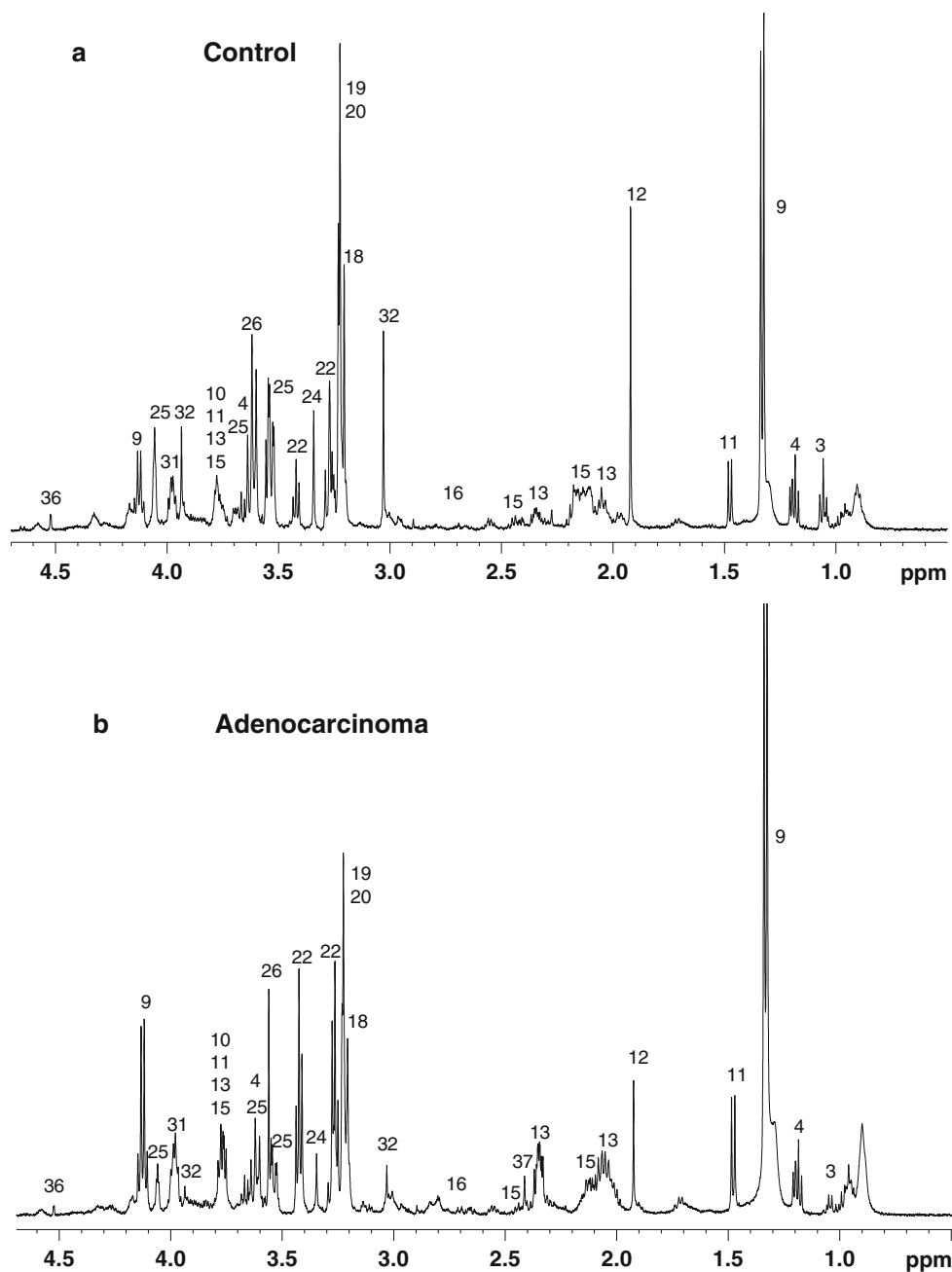


Fig. 1 Representative 1D ¹H CPMG HRMAS spectra of healthy (a) and adenocarcinoma (b) colorectal tissues originating from the same patient. For display purposes, the amplitude of the lactate peak at 1.33 ppm has been cut. Partial metabolite assignment in the

4.7–0.5 ppm region is indicated. The numbers refer to the metabolites listed in Table 1. The Eretic peak at 10.3 ppm, which is not shown in this plot, is a synthetic Eretic peak that is used for the absolute quantification of the metabolites

Table 1 ^1H resonance assignments of the metabolites present in cancerous and healthy human colons. The nomenclature used for fatty acids is defined in the publication of Martínez-Bisbal (Martínez-Bisbal et al. (2004))

Metabolites	Group	^1H chemical shift (ppm)	^{13}C chemical shift
1 Isoleucine	δCH_3	0.93	13.77
	γCH_3	1.00	17.37
	γCH_2	1.48	27.13
	αCH	3.67	62.25
2 Leucine	δCH_3	0.95	23.36
	$\delta'\text{CH}_3$	0.96	24.61
	γCH	1.71	26.69
	βCH_2	1.71	42.40
3 Valine	αCH	3.73	56.06
	γCH_3	0.98	19.14
	$\gamma'\text{CH}_3$	1.04	20.56
4 Ethanol	βCH	2.27	31.77
	CH_3	1.17	19.49
	CH_2OH	3.65	60.15
5 Fatty acids (a)	(2) CH_2	1.28	34.46
	(1) CH_2	1.29	25.35
6 Fatty acids (b)	(2) CH_2	2.04	27.39
	CH_2	2.80	28.05
7 Fatty acids (a) (b)	(n) CH_2	1.28	32.24
8 Fatty acids (c)	(2) CH_2	1.58	27.31
	(1) CH_2	2.25	36.06
9 Lactate	CH_3	1.33	22.65
	CH	4.12	71.05
10 Lysine	γCH_2	1.46	24.19
	δCH_2	1.71	29.14
	βCH_2	1.90	32.50
	εCH_2	3.01	41.84
11 Alanine	βCH_3	1.47	18.80
	αCH	3.78	53.18
12 Acetate	CH_3	1.92	25.97
13 Glutamate	βCH_2	2.06	29.69
	γCH_2	2.34	35.93
	αCH	3.76	57.19
14 Methionine	εCH_2	2.12	16.50
15 Glutamine	βCH_2	2.14	NA
	γCH_2	2.44	33.37
	αCH_2	3.77	NA
16 Aspartic acid	$\beta\text{CH}_2(\text{u})$	2.70	39.15
	$\beta\text{CH}_2(\text{d})$	2.80	39.15
	αCH	3.90	54.81
17 Phosphoethanolamine	$-\text{CH}_2-\text{NH}_3^+$	3.22	43.21
	$-\text{CH}_2-\text{O}-$	3.98	62.94
18 Choline	$-\text{N}^+(\text{CH}_3)_3$	3.21	NA
	βCH_2	3.52	69.84

Table 1 continued

Metabolites	Group	^1H chemical shift (ppm)	^{13}C chemical shift
19 Phosphorylcholine	$-\text{N}^+(\text{CH}_3)_3$	3.22	56.49
	βCH_2	3.61	69.03
	αCH	4.18	60.59
20 Glycerophosphocholine	$-\text{CH}_2-\text{NH}_3^+$	3.23	NA
	αCH_2	4.29	56.36
21 Arginine	βCH_2	1.91	30.17
	δCH_2	3.22	43.23
22 Taurine	$-\text{CH}_2-\text{NH}_3^+$	3.26	50.07
23 Proline	$-\text{CH}_2-\text{SO}_3^-$	3.42	38.07
	γCH_2	2.00	26.45
24 Scyllo-Inositol	$\beta\text{CH}_2(\text{u})$	2.07	31.70
	$\beta\text{CH}_2(\text{d})$	2.35	31.70
	$\delta\text{CH}_2(\text{u})$	3.33	48.75
	$\delta\text{CH}_2(\text{d})$	3.41	48.75
25 Myo-Inositol	αCH	4.13	63.81
	All Hs	3.34	76.25
26 Glycine	C5H	3.27	76.96
	C1H, C3H	3.54	73.80
	C4H, C6H	3.62	75.05
	C2H	4.05	74.90
27 Threonine	αCH	3.56	44.10
	αCH	3.59	63.14
28 Glycerol	βCH	4.26	68.55
	1,3 $\text{CH}_2\text{OH}(\text{u})$	3.56	65.10
	1,3 $\text{CH}_2\text{OH}(\text{d})$	3.65	65.10
29 β -Glucose	$-\text{CH}(\text{OH})-$	3.78	74.82
	C6H(u)	3.73	63.53
	C6H(d)	3.89	63.53
30 α -Glucose	C1H	4.65	NA
	C1H	5.23	NA
31 Serine	αCH	3.84	59.08
	βCH	3.97	62.90
32 Creatine	CH_3	3.03	39.67
	CH_2	3.93	56.32
33 Asparagine	αCH	4.00	53.94
34 Tyrosine	CH 3,5	6.88	118.34
	CH 2,6	7.18	133.32
35 Phenylalanine	CH 2,6	7.32	131.82
	C4	7.40	131.69
36 Ascorbic acid	C4H	4.53	NA
37 Succinic acid	($\alpha,\beta\text{CH}_2$)	2.40	NA

ischemia is identical for both samples. The nature and the concentration of the metabolites present in each biopsy can therefore be reliably compared.

3.2 Statistical analysis

PCA analysis of the CPMG data applied to the 4.7–0.5 ppm chemical shift range and using the entire patient population was not able to discriminate adenocarcinomas from healthy tissues in a statistically significant manner. A few outliers were however identified by this analysis. In this type of studies, outliers can occur because of erroneous spectrum referencing, sample pollution due to the fixing medium (o.c.t. compound, Tissue-Tek, Sakura) employed by the pathologists or an unusual amount of lipids. In our study, 12 biopsies out of a total of 84 were removed, leaving a cohort of 72 biopsies. PLS-DA analysis of the CPMG data was initially performed on the same population of 72 biopsies. As for the PCA analysis, buckets within the 4.7–0.5 ppm range were initially used. Biopsies were classified in a Y column matrix as being either adenocarcinomas (value of 0) or healthy tissues (value of 1). The analysis generated a two component PLS-DA model characterized by a faithful representation of the Y data ($R^2Y = 0.80$) and, more importantly, by a very good cumulative confidence criterion of prediction ($Q^2 = 0.68$). The score plot of the PLS-DA model (Fig. 2a) showed a very clear separation of the two sets of biopsies. The loading plot (Fig. 2b) of the PLS-DA model revealed a statistically significant elevation of taurine, glutamate, aspartate, and lactate in adenocarcinomas compared to healthy tissues. Conversely, the metabolic signature of normal colon tissue consisted of a higher concentration of myo-inositol and β -glucose. The assessment of the metabolites responsible for this classification can lead to the detection of pathological biomarkers of tumor metabolism. It is necessary to stress that the discriminant power of each metabolite is linked to its relative difference in concentration between each groups (tumor versus healthy tissue) for each clinical variable considered (i.e., histological classification). Therefore, a metabolite present in subtle concentration in one group but completely absent in the second group, will play a major role in discriminant analysis. Conversely, a metabolite detected in high concentration in both groups will be less significant in the clustering process (inconsistent relative difference).

Since the initial PLS-DA model using the entire set of biopsies gave extremely promising results, a second PLS-DA model was built using only a subset of the available biopsies. More precisely, only the first 50 biopsies of our sample set, corresponding to 27 tumoral cases and 23 controls, were used to build a new model. Moreover, since the most discriminant metabolites appeared consistently to be, in a decreasing order of importance, taurine, myo-inositol, glutamate, and aspartate only these metabolites were used for model development. This choice resulted in an X matrix containing only 30 columns, thereby decreasing considerably the risk of overfitting the data. The resulting

PLS-DA model was characterized by a R^2Y of 0.83 and a Q^2 of 0.80. The score and the loading plot of the new PLS-DA model are shown in Fig. 3a, b respectively. The detailed analysis of these plots shows that the average position of the different patients is similar to the model using the whole set of patients (Fig. 2a). The PLS-DA model was validated by circular validation (Fig. 3c). The remaining 22 colon biopsies (14 tumoral cases and 8 controls) were then used as a test set and subjected to a blind classification process using the previously determined statistical model. The outcome of the PLS-DA classification process is presented in Fig. 4. A visual inspection of the score plot shows that the 14 cancerous patients and the 8 healthy patients are classified in the correct region. This result is very promising and shows that the HRMAS analysis has very good sensitivity and specificity. Following this classification process, a new PCA analysis was performed on the whole biopsy cohort ($n = 72$) using only data points corresponding to taurine, myo-inositol, glutamate, and aspartate. The PCA results showed this time a clear differentiation between the two groups of biopsies (data not shown). This result is extremely important since PCA alone is rarely capable of separating biological samples in an accurate way. This feature proves the high discriminating power of taurine, myo-inositol, glutamate, and aspartate in healthy and tumoral colon tissues. Since the weight of biopsy present in each sample is accurately known, the Eretic peak can be used to quantify the resolved taurine peak at 3.42 ppm by simple integration. The average taurine concentration of cancerous and healthy biopsy samples was found to be 4.5 ± 1.7 mmol/kg and 2.0 ± 0.6 mmol/kg, respectively. This result is clearly in agreement with the conclusions of the statistical analysis.

The results presented in this manuscript were finally compared to related studies published in the literature. In 1996, Moreno and Arús (Moreno and Arús (1996)) compared the metabolic content of tumoral colon biopsies ($n = 16$) with normal mucosa biopsies ($n = 10$) using extraction techniques and high resolution liquid state ^1H NMR. They found that colon tumors were characterized by a higher level of taurine, glutamate, aspartate, lactate, spermine, putrescine, glutathione, and glycerophosphoethanolamine. Colon tumors were also associated with a significant decrease in myo- and scyllo-inositol. The results presented in our manuscript agree to a large extent with the results obtained by Moreno and Arús. In our spectra of intact biopsies, taurine, glutamate, aspartate, and lactate are clearly present in higher concentrations in adenocarcinomas while myo-inositol and β -glucose are detected in higher concentrations in healthy tissues. Spermine and putrescine are however not detected in our HSQC spectra. Like Moreno and Arús, we find that, among these metabolites, taurine and myo-inositol are the most promising biomarkers. More recently, Seierstad (Seierstad et al. (2008)) reported a HRMAS study of human

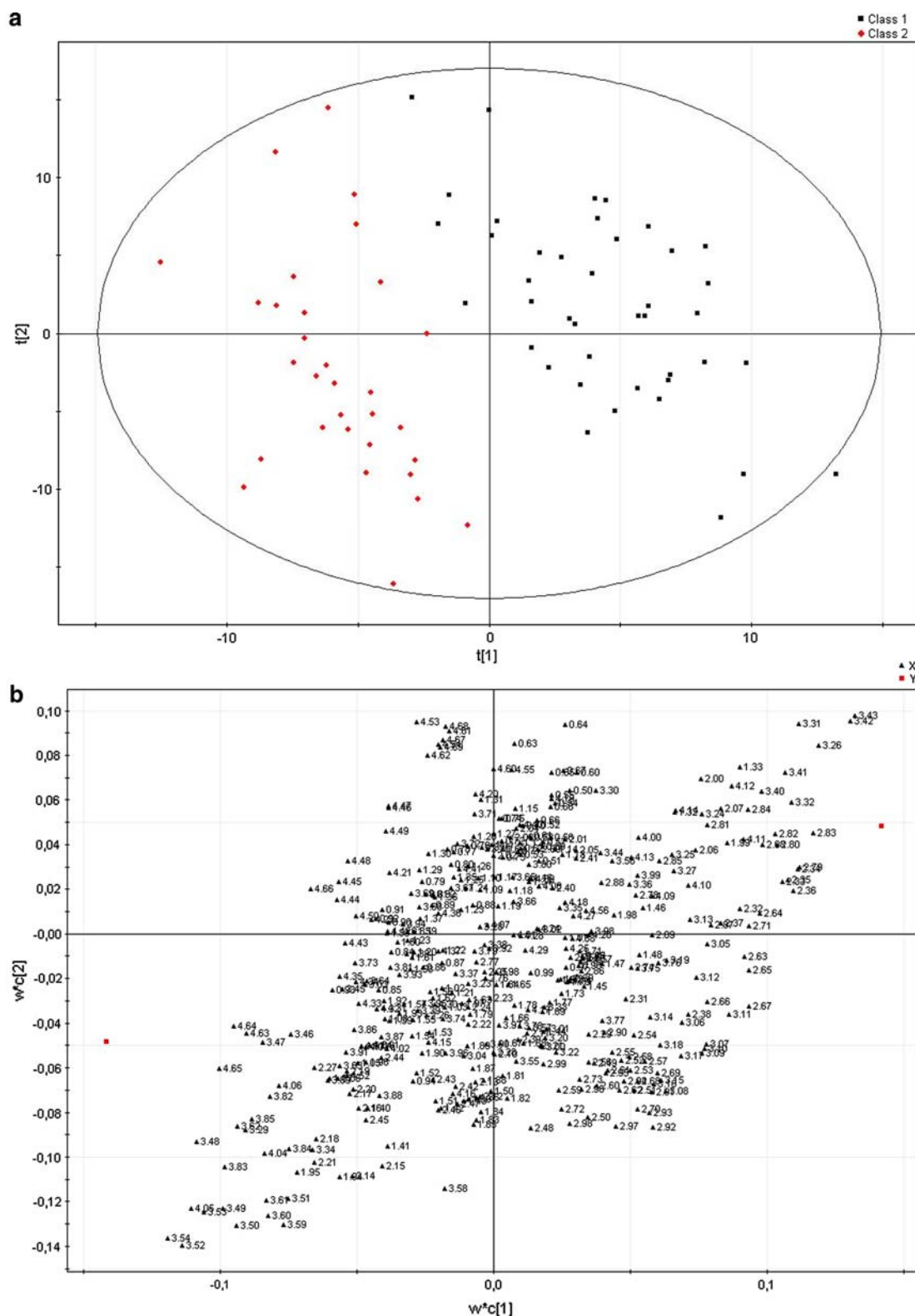
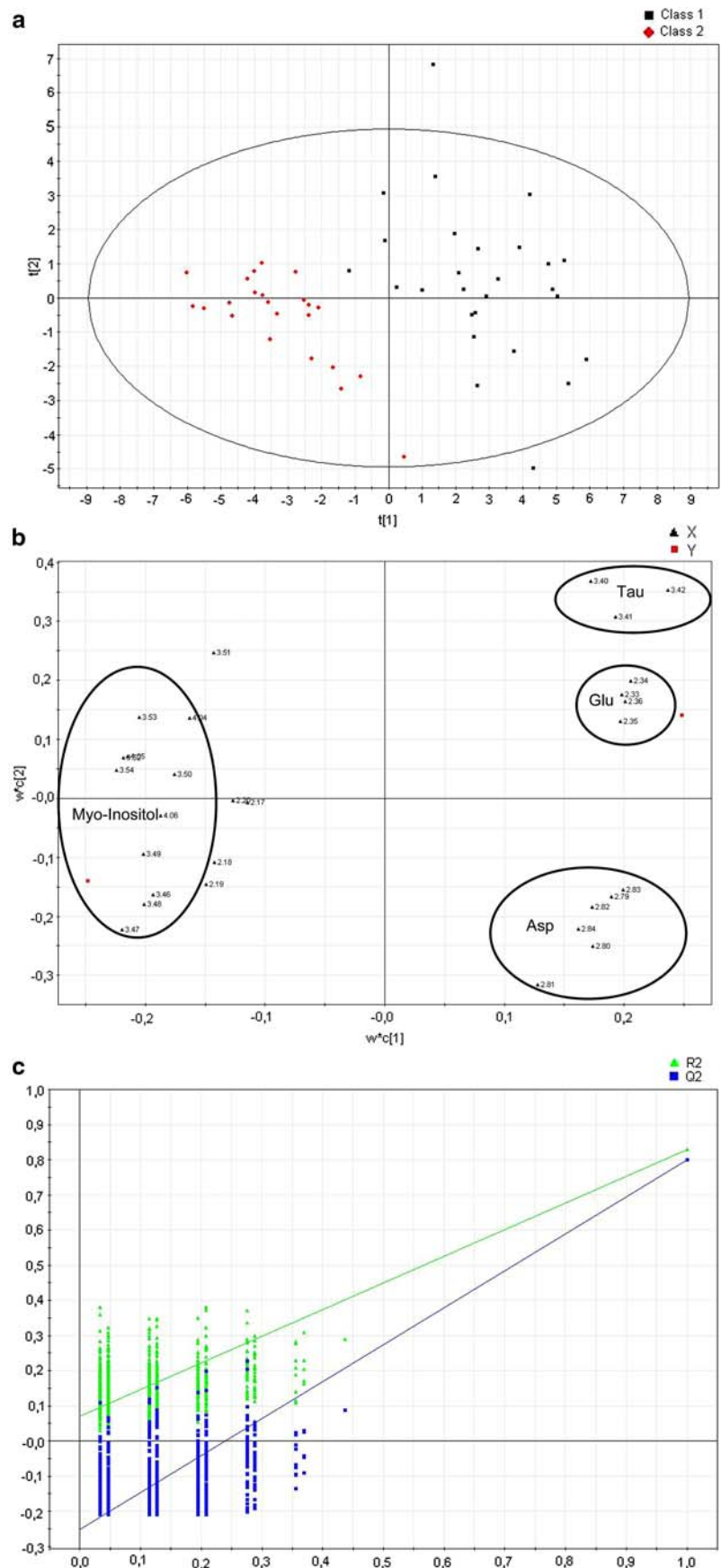


Fig. 2 Two component PLS-DA model obtained when classifying the entire cohort of colorectal biopsies ($n = 72$) according to histopathological criteria and using the 4.7–0.5 ppm spectral region. The PLS-DA model is characterized by the following parameters: $R^2Y = 0.80$ and $Q^2 = 0.68$. The high Q^2 value indicates good predictive capabilities for the model. **a** PLS-DA score plot. Cancerous

and healthy colons are represented by black squares and red rhombs, respectively. **b** PLS-DA loading plot. Each point (chemical shift in ppm) represents the center of a single bucket corresponding to an interval of 0.01 ppm in the HRMAS spectrum. These buckets are associated with one or more metabolite resonances

Fig. 3 Two component PLS-DA model obtained when classifying a subset of the cohort of colorectal biopsies (27 tumoral cases and 23 controls) according to histopathological criteria and using only the most important discriminant variables determined in the previous model (taurine, glutamate, aspartate, and myo-inositol regions). The PLS-DA model is characterized by the following parameters: $R^2Y = 0.83$ and $Q^2 = 0.80$. **a** PLS-DA score plot. Cancerous and healthy colons are represented by black squares and red rhombs, respectively. **b** PLS-DA loading plot. Each point (chemical shift in ppm) represents the center of a single bucket corresponding to an interval of 0.01 ppm in the HRMAS spectrum. **c** Model validation using 999 permutations. The regression line represents the correlation coefficient between the original Y (histopathological classification) and the permuted Y versus the cumulative Q^2 . A negative slope of regression line suggests that there is no model overfit. Q^2 and R^2Y are represented by blue squares and green squares, respectively



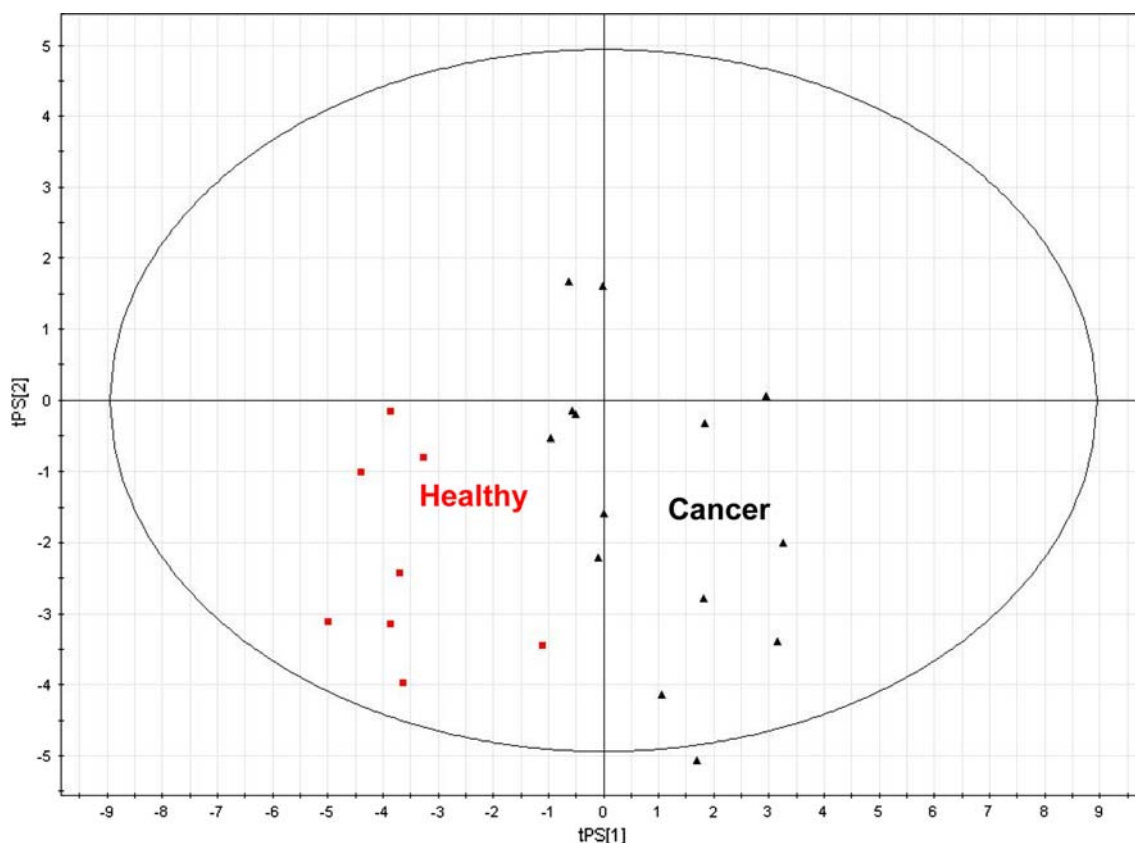


Fig. 4 Blind test classification of colon biopsies (14 tumoral cases and 8 controls) using the metabolic model presented in Fig. 3. Cancerous and healthy colons are represented by black triangles and red squares, respectively

rectal cancer biopsies and colorectal xenografts. The aim of their study was to identify the most appropriate human rectal xenograft model for preclinical magnetic resonance spectroscopy studies. They made no attempts at differentiating tumor from healthy human colon samples. However, the results they published are clearly in agreement with those presented here. The metabolites they identified are similar to the ones reported in Table 1 and colon tumors are characterized by a higher level of taurine.

Our results are therefore in agreement with previous studies already published in the literature. The strength of our approach is clearly to work on intact colon biopsies with an analytical technique (HRMAS) capable of delivering some information as to the nature of the tissue in a reasonable amount of time (ca. 20 min).

4 Concluding remarks

In conclusion, our results show that adenocarcinomas can be separated with a good specificity and sensibility from healthy colorectal biopsies solely on the basis of their metabolic content. An elevated amount of taurine, glutamate aspartate, and lactate in a colorectal biopsy is highly

symptomatic of an adenocarcinoma. Conversely, a high amount of myo-inositol and β -glucose is a signature of a healthy tissue. Our HRMAS study is one of the first ones to actually use a PLS-DA metabolic model to classify tumoral and healthy colon biopsies. The next crucial step will be to investigate if a metabolic analysis can discriminate between different patient populations within the cancer group—the purpose being the identification of possible markers of unfavorable patient evolution (i.e., appearance of precocious metastasis or local recurrence). Such monitoring requires a careful longitudinal follow-up of patients and this will be the object of our next study.

Acknowledgments This work is part of the CARMEN project and was supported by grants from Région Alsace, Oséo, Communauté Urbaine de Strasbourg, Conseil Départemental du Bas-Rhin, Bruker BioSpin, University Louis Pasteur and University Hospitals of Strasbourg. The technical assistance of Dr. J. Raya, Mrs. H. Kada, and Mrs. T. Tong is gratefully acknowledged.

References

- Akoka, S., Barantin, L., & Trierweiler, M. (1999). Concentration measurement by proton NMR using the ERETIC method. *Analytical Chemistry*, 71(13), 2554–2557. doi:10.1021/ac981422i.

- Andrew, E. R., Bradbury, A., & Eades, R. G. (1958). Removal of dipolar broadening of nuclear magnetic resonance spectra of solids by specimen rotation. *Nature*, *183*, 1802–1803. doi:10.1038/1831802a0.
- Cheng, L. L., Anthony, D. C., Comite, A. R., et al. (2000). Quantification of microheterogeneity in glioblastoma multiforme with ex vivo high-resolution magic-angle spinning (HRMAS) proton magnetic resonance spectroscopy. *Neuro-Oncology*, *2*(2), 87–95. doi:10.1215/15228517-2-2-87.
- Cheng, L. L., Chang, I. W., Louis, D. N., & Gonzalez, R. G. (1998). Correlation of high-resolution magic angle spinning proton magnetic resonance spectroscopy with histopathology of intact human brain tumor specimens. *Cancer Research*, *58*(9), 1825–1832.
- Cheng, L. L., Ma, M. J., Becerra, L., et al. (1997). Quantitative neuropathology by high resolution magic angle spinning proton magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(12), 6408–6413. doi:10.1073/pnas.94.12.6408.
- Cusack, J. R., Giacco, G. G., Cleary, K., et al. (1996). Survival factors in 186 patients younger than 40 years old with colorectal adenocarcinoma. *Journal of the American College of Surgeons*, *183*, 105–112.
- Davis, A. L., Keeler, J., Laue, E. D., & Moskau, D. (1992). Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *Journal of Magnetic Resonance (San Diego Calif)*, *98*(1), 207–216.
- Elbayed, K., Dillmann, B., Raya, J., Piotta, M., & Engelke, F. (2005). Field modulation effects induced by sample spinning: application to high-resolution magic angle spinning NMR. *Advances in Magnetic Resonance*, *174*, 2–26. doi:10.1016/j.jmr.2004.11.017.
- Fearon, E., Cho, K. R., Nigro, J., et al. (1990). Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science*, *247*, 49–56. doi:10.1126/science.2294591.
- Jemal, A., Siegel, R., Ward, E., et al. (2007). Cancer Statistics. *CA: A Cancer Journal for Clinicians*, *57*, 43–66.
- Leichman, C. G., Lentz, H. J., Danenberg, K., et al. (1997). Quantitation of intratumoral thymidylate synthase expression predicted for disseminated colorectal cancer response and resistance to protracted infusion of 5-fluorouracil and weekly leucovorin. *Journal of Clinical Oncology*, *15*, 3223–3229.
- Libutti, S. K., Saltz, L. B., & Tepper, J. E. (2008). Principles and practice in oncology. In V. T. De Vita, T. S. Lawrence, & S. A. Rosenberg (Eds.), *Colon cancer in cancer* (pp. 1232–1285). Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins.
- Lippens, G., Bourdonneau, M., Dhalluin, C., et al. (1999). Study of compounds attached to solid supports using high resolution magic angle spinning NMR. *Current Organic Chemistry*, *3*, 147–169.
- Lowe, I. J. (1959). Free induction decays of rotating solids. *Physical Review Letters*, *2*, 285–287. doi:10.1103/PhysRevLett.2.285.
- Martínez-Bisbal, M. C., Martí-Bonmatí, L., Piquer, J., et al. (2004). ¹H and ¹³C HR-MAS spectroscopy of intact biopsy samples ex vivo and in vivo ¹H MRS study of human high grade gliomas. *NMR in Biomedicine*, *17*(4), 191–205. doi:10.1002/nbm.888.
- Martínez-Granados, B., Monleon, D., Martínez-Bisbal, M. C., et al. (2006). Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning ¹H NMR spectroscopy. *NMR in Biomedicine*, *19*(1), 90–100. doi:10.1002/nbm.1005.
- Moreno, A., & Arús, C. (1996). Quantitative and qualitative characterization of ¹H NMR spectra of colon tumors, normal mucosa and their perchloric acid extracts: Decreased levels of myo-inositol in tumours can be detected in intact biopsies. *NMR in Biomedicine*, *9*(1), 33–45. doi:10.1002/(SICI)1099-1492(199602)9:1<33::AID-NBM391>3.0.CO;2-G.
- Piotta, M., Bourdonneau, M., Furrer, J., et al. (2001). Destruction of magnetization during TOCSY experiments performed under magic angle spinning: Effects of Radial B1 inhomogeneities. *Journal of Magnetic Resonance (San Diego, Calif.)*, *149*, 114–118. doi:10.1006/jmre.2001.2287.
- Seierstad, T., Røe, K., Sitter, B., et al. (2008). Principal component analysis for the comparison of metabolic profiles from human rectal cancer biopsies and colorectal xenografts using high-resolution magic angle spinning ¹H magnetic resonance spectroscopy. *Molecular Cancer*, *7*, 1–13. doi:10.1186/1476-4598-7-33.
- Shaka, A. J., Lee, C. J., & Pines, A. (1988). Iterative schemes for bilinear operators: Application to spin decoupling. *Journal of Magnetic Resonance (San Diego, Calif.)*, *77*, 274–293.
- Sitter, B., Sonnewald, U., Spraul, M., Fjosne, H. E., & Gribbestad, I. S. (2002). High-resolution magic angle spinning MRS of breast cancer tissue. *NMR in Biomedicine*, *15*(5), 327–337. doi:10.1002/nbm.775.
- Sjøbakk, T. E., Johansen, R., Bathen, T. F., et al. (2008). Characterization of brain metastases using high-resolution magic angle spinning MRS. *NMR in Biomedicine*, *21*(2), 175–185. doi:10.1002/nbm.1180.
- Tenenhaus, M. (1998). *La régression PLS: Théorie et Pratique*. Paris.
- Tzika, A. A., Cheng, L. L., Goumnerova, L., et al. (2002). Biochemical characterization of pediatric brain tumors by using in vivo and ex vivo magnetic resonance spectroscopy. *Journal of Neurosurgery*, *96*(6), 1023–1031.
- Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., et al. (2008). Assessment of PLS-DA cross validation. *Metabolomics*, *4*(1), 81–89. doi:10.1007/s11306-007-0099-6.
- Wider, G., & Dreier, L. (2006). Measuring protein concentrations by NMR spectroscopy. *Journal of the American Chemical Society*, *128*(8), 2571–2576. doi:10.1021/ja055336t.