

Hierarchical clustering identifies a subgroup of colonic adenocarcinomas expressing crypt-like differentiation markers, associated with MSS status and better prognosis

Laure Droy-Dupré · Céline Bossard · Christelle Volteau · Stéphane Bezieau · Christian L. Laboisse · Jean-François Mosnier

Received: 15 March 2014 / Revised: 21 December 2014 / Accepted: 20 January 2015 / Published online: 11 February 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract The aim of this study was to identify in the group of colonic adenocarcinomas, not otherwise specified (NOS), subgroups of oncogenetic and prognostic significance based on the expression of immunohistochemical markers of epithelial cell differentiation of the gastrointestinal tract. Hierarchical clustering analysis of 122 adenocarcinomas (NOS) identified four clusters based on how closely their profile of immunohistochemical expression of differentiation markers was related: (i) a major cluster of 83 adenocarcinomas (68 %) called crypt-like carcinoma (CLA) with a immunohistochemically expressing colonic crypt differentiation markers (cytokeratin 20+, CDX2+, MUC2+ or MUC2–) and (ii) three minor clusters, characterized by the loss of colonic crypt differentiation markers and/or the acquisition of expression of markers of metaplastic foveolar gastric differentiation (MUC5AC+) and/or aberrant cytokeratin 7 expression. CLAs were invariably MSS (χ^2 test: $p < 0.0001$). The sole parameters associated with worse overall survival of the 122 patients with adenocarcinoma (NOS) were pT stage, pN+ stage, and advanced clinical stage. Interestingly, CLA lineage

of differentiation was an independent prognostic parameter for better overall survival among the 40 patients with an adenocarcinoma (NOS) stage III. In conclusion, hierarchical clustering led to the identification of a main cluster of adenocarcinoma (NOS) with crypt-like differentiation, associated with MSS status and better prognosis. Its value as a biomarker of response to conventional chemotherapeutic agents deserves to be examined in randomized therapy trials.

Keywords Colonic adenocarcinomas · Lineage of differentiation · Hierarchical clustering analysis

Introduction

Recent progress regarding genetics and molecular oncogenesis of colonic adenocarcinoma has led to the concept of heterogeneity, resulting in several clinico-pathological entities characterized by distinct molecular genetic pathways modulated by epigenetic events [1]. The now classical view of colorectal oncogenesis, the so-called canonical pathway derived from the Familial Adenomatous Polyposis model, involves the accumulation of mutations of *APC*, *KRAS*, and *TP53* [2–4]. The so-called mutator phenotype [5–7], characterized by genetic or epigenetic alterations of mismatch repair (MMR) genes *MLH1*, *MSH2*, *MSH6*, and *PMS2* resulting in microsatellite instability (MSI-H), represents an alternative pathway of colorectal oncogenesis. While in Lynch syndrome MMR gene abnormalities result from germline mutations, in sporadic cases they result from hypermethylation of CpG Islands of the *MLH1* promoter [8].

Some subtypes of sporadic colonic adenocarcinomas, i.e., mucinous or medullary carcinomas, are frequently

L. Droy-Dupré · C. Bossard · C. L. Laboisse · J.-F. Mosnier
Department of Pathology, Centre Hospitalier Universitaire,
Nantes, France

C. Volteau
Biostatistics Department, Centre Hospitalier Universitaire,
Nantes, France

S. Bezieau
Department of Genetics, Centre Hospitalier Universitaire,
Nantes, France

J.-F. Mosnier (✉)
Service d'Anatomie Pathologique, Hôtel Dieu, 1 place Ricordeau,
44093 Nantes, France
e-mail: jfmosnier@live.fr

MSI-H [9–11]. For adenocarcinoma not otherwise specified (NOS), simply called adenocarcinoma in the last edition of the WHO classification [12], no histopathological features of tumor cell differentiation are associated with molecular characteristics. In this context, the aim of our study was to examine whether some immunophenotypical characteristics of differentiation of colonic adenocarcinoma (NOS) might be relevant in terms of molecular oncogenesis and/or clinical behavior. To address this issue, hierarchical clustering analysis of a cohort of patients from a single institution was performed, based on a panel of immunohistochemical markers of differentiation of gastrointestinal epithelia (CDX2, cytokeratin 20 and cytokeratin 7, MUC2, and MUC5AC). Clusters of differentiation resulting from the hierarchical clustering analysis were correlated with immunophenotypic and genotypic markers of colonic oncogenesis (MSI/MSS status, MLH1, MSH2, MSH6) as well as to evolution of disease in terms of overall survival.

Material and methods

Patients

The case series consisted of 122 patients with colonic adenocarcinoma (NOS), surgically treated, consecutively selected when tissue was available, at Nantes University Hospital between 1998 and 2008. Patients with other histological subtypes of colonic carcinomas as defined by WHO classification [12] were excluded from this study. The adenocarcinomas (NOS) are defined as gland-forming with variability in the size and configuration of the glandular structures. The epithelial tumor cells are usually large and tall. Adenocarcinomas are graded according to the WHO recommendations as the percentage of the tumor showing formation of gland-like structures [12]. Tumors were staged according to UICC-TNM staging 7th ed. [13]. Our case series consisted of 5 pT1, 8 pT2, 84 pT3, and 25 pT4 colon carcinomas including 69 colon carcinomas with lymph nodes metastases. Synchronous visceral metastases were detected in 33 patients, i.e., in the liver in 21 patients, lungs in 5 patients, peritoneal cavity in 4 patients, bone in 2 patients, and ovary in 1 patient. The case series consisted of 8 stage I, 41 stage II, 40 stage III, and 33 stage IV colon carcinomas. All patients with adenocarcinoma (NOS) stage III received chemotherapy. None of the patients underwent neoadjuvant therapy before surgery, patients with rectal carcinomas being excluded from the study. The following information was scored from the clinical charts: age, gender, tumor location, treatment, follow-up information including metachronous metastases (location and date of diagnosis), date of death.

Tissue micro-array construction

A tissue micro-array including the 122 colonic adenocarcinomas (NOS) was constructed. Tissue cores with 1.0 mm diameter were taken from donor paraffin blocks of chosen tumor areas ($n=3$) as well as from the distant normal mucosa ($n=1$). These tissue cores were inserted into recipient paraffin blocks using a tissue arrayer (Beecher Instrument, Silver Springs, MD). Cores of placental tissue were added to each array and served both as control and location landmarks.

Immunohistochemistry

Sections of tissue microarray blocks were immunostained using peroxidase-labeled polymer method (Envision Flex Plus kit, Dako, Glostrup, Denmark). Primary antibodies against the following antigens were used: cytokeratin 7 (clone OV-TL 12-30, Dako, dilution 1:100), cytokeratin 20 (clone KS20.8, Dako, dilution 1:100), MUC2 (clone Ccp58, Novocastra, Newcastle upon Tyne, UK, dilution 1:100), MUC5AC (clone 45 M1, gift of Pr Bara, dilution 1:50), CDX2 (clone CDX2-88, Biogenex, Fremont, CA, dilution 1:100), MLH1 (clone G168-15, BD Biosciences, Erembodegem, Belgium, dilution 1:100), MSH2 (clone D06571-8, Oncogene, Cambridge, MA, dilution 1:100), and MSH6 (clone 44, BD Transduction Laboratories, Lexington, KY, dilution 1:40). For visualization of bound antibody, the sections were immersed in 3-3' diaminobenzidine (Envision Flex Plus kit, Dako) and counterstained with hematoxylin.

Evaluation of immunohistochemistry

All cases were scored by two pathologists. Immunoreactivity for cytoplasmic cytokeratin 20, cytokeratin 7, MUC2, MUC5AC, and nuclear CDX2 was scored as the percentage of positive cells. The final immunohistochemical score for a given marker corresponded to the mean of the scores given for the three tumor cores from each patient.

Hierarchical clustering analysis of immunohistochemical differentiation of adenocarcinomas, NOS

In a first step, the immunohistochemical data on expression of the differentiation markers cytokeratin 20, cytokeratin 7, MUC2, MUC5AC, and CDX2 were specifically reformatted for hierarchical clustering analysis [14]. A four-point scale was used with 0=negative (<5 % of cells staining), 1=uninterpretable (e.g., because of loss of the tissue or excessive background staining), 2=weak staining (i.e., 5–50 % of cells showing weak to intermediate intense staining), and 3=strong staining (>50 % of cells showing weak to moderately intense staining or >5 % of cells showing strong staining). Staining was considered relevant when it reproduced the cytoplasmic

MUC2 or CK20 or nuclear CDX2 expression in normal colonic epithelial cells. In addition, MUC5AC staining was considered relevant when reproducing the cytoplasmic expression normally present in foveolar gastric cells. Scores were entered into a Microsoft Excel spreadsheet following the model described by Liu et al. [14].

For univariate and multivariate survival analysis, both percentage and reformatted values of immunopositive cells were tested.

The four-point scale was initially designed for hierarchical clustering analysis. Hierarchical clustering analysis was performed using Cluster and TreeView software, freely available at <http://rana.lbl.gov/EisenSoftware.htm/> [14, 15]. An Excel macro TMA-Deconvoluter was also downloaded from Eisen Software for processing of raw tissue microarrays staining data into a format compatible with Cluster software [14, 15]. The clustered data were finally viewed using TreeView [15].

Microsatellite instability status and mismatch repair proteins

To determine MSI/MSS status, DNA was extracted from two whole 10 μm thick sections of donor blocks from 122 colon carcinomas. MSI/MSS status was determined by PCR, through amplification of five mononucleotide markers BAT25, BAT26, NR21, NR22, and NR24, designed and validated previously [16].

The 3- μm sections of tissue microarrays were immunostained for three mismatch repair proteins MLH1, MSH2, and MSH6 using the peroxidase-labeled polymer method as described before. Tumor mismatch repair protein expression was assessed as retained (normal) or lost.

Statistics

Hierarchical clustering analysis was performed to assess associations between the five evaluated markers of differentiation (see above) [14, 15]. Associations between categorical variables (cluster of differentiation, gender, MSI/MSS status, tumor location) were calculated using two-tailed χ^2 test.

Overall survival (OS) was measured from the date of surgery to the date of death related to colon carcinoma (i.e., death unrelated to colon carcinoma was censored at the time of death). Univariate survival analysis was performed using the log-rank test and illustrated using Kaplan–Meier curves. Cox hazard proportional model was used for multivariate survival analysis. The hazard risk ratio (HR) and 95 % confidence intervals (95 % CI) were estimated for each association. Based on the whole cohort, UICC stage, MSI/MSS status, grade, cluster of differentiation, or each single marker of differentiation were entered in a first set of multivariate analyses. In a second set of multivariate analyses, the prognostic value of the same parameters was evaluated in each well defined UICC stage. *p* values <0.05 were considered

statistically significant. Statistical analyses were performed using SAS software version 9.3 for Windows (SAS Institute, Cary, NC, USA).

Results

Clinico-pathological data

The case series comprised of 122 colon adenocarcinomas NOS, in 44 women and 78 men (age range 19–91 years, mean 68 years, median 68 years). Of the carcinomas, 61 were located in the right colon and 61 in the left colon.

Immunohistochemical differentiation profile of the normal colon

In the normal colon, CDX2 expression, restricted to the nucleus of epithelial cells, was more intense in crypt bottom than in surface epithelial cells. Both goblet cells and eosinophilic columnar cells expressed CDX2. Cytokeratin 20 expression was restricted to surface and upper crypt epithelial cells, which corresponds to those most differentiated. MUC2 was expressed by goblet cells of the entire crypt without a gradient of expression. MUC2 expression appeared as cytoplasmic granular staining peripheral to mature secretory vesicles. MUC5AC and cytokeratin 7 were not expressed in normal colonic mucosa sampled at a distance from the tumor (Fig. 1).

Immunohistochemical expression of markers of differentiation in colonic adenocarcinomas (NOS)

CDX2 expression was high in 84 adenocarcinomas (69 %) and low (less than 50 % immunopositive-cells) in 38 cases (31 %). In 16 of these 38 cases, CDX2 expression was lost. Cytoplasmic cytokeratin 20 expression was found in 105 adenocarcinomas; it was lost in 16 adenocarcinomas (13 %). MUC2 expression was found in both carcinomatous goblet cells (granular cytoplasmic staining peripheral to mature secretory vesicles) and eosinophilic columnar cells (granular homogeneous cytoplasmic staining) in 61 adenocarcinomas (50 %). Cytoplasmic cytokeratin 7 expression was present in eight adenocarcinomas (7 %). This staining was intense, displayed by nearly 100 % of tumor cells in two cases including an adenocarcinoma in a patient with ulcerative colitis. Ectopic MUC5AC staining was observed in 20 adenocarcinomas (16 %); MUC5AC staining was present in the cytoplasm of carcinomatous cells.

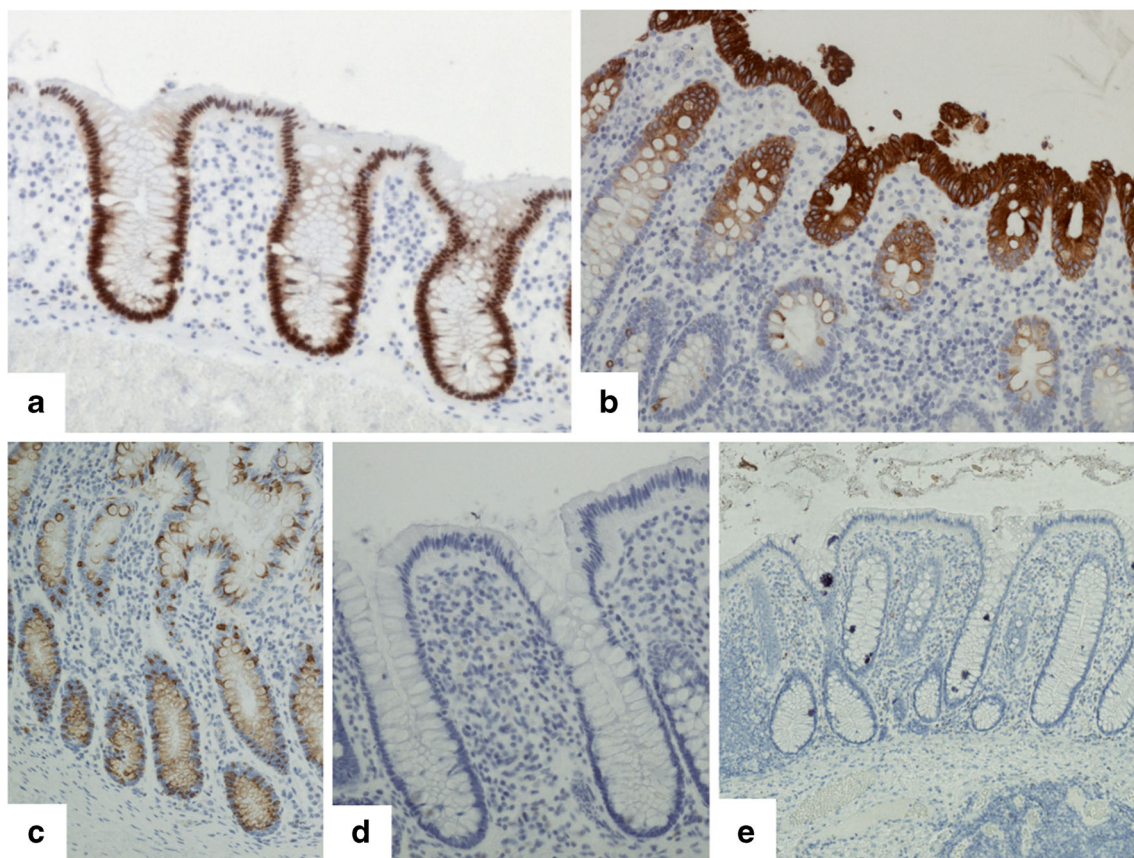


Fig. 1 Expression of CK20, CDX2, MUC, CK7, and MUC5AC in the normal colonic crypt. **a** CDX2 expression along the entire colonic crypt with an attenuation of expression in the intercryptic surface epithelium; **b**

strong expression of cytokeratin 20 in the surface epithelium as well as in the upper third of the crypt; **c** MUC2 expression restricted to goblet cells; **d** the normal crypts do not express either MUC5AC or cytokeratin 7

Hierarchical clustering analysis of 122 adenocarcinomas (NOS) based on immunohistochemical markers of differentiation

The hierarchical clustering analysis divided the 122 colonic adenocarcinomas (NOS) into four groups based on the relatedness of the markers of differentiation immunoprofile (Fig. 2). These four groups were designated cluster of differentiation (CD) I, II, III, and IV. As can be seen from the TreeView output (Fig. 2), CD I consisted of 83 adenocarcinomas (68 %) showing the colonic crypt differentiation immunoprofile (cytokeratin 20-positive, CDX2-positive, MUC2-positive or MUC2-negative) (Fig. 3a). In this group, there was a good level of correlation between the cases ($r=0.611$). CD II consisted of 11 adenocarcinomas (9 %) exhibiting a mixed immunoprofile of both normal colonic crypt and metaplastic foveolar gastric differentiation (cytokeratin 20-positive, CDX2-positive, MUC2-positive or MUC2-negative, MUC5AC-positive) (Fig. 3b). In this group, there was a nearly perfect level of correlation between the cases ($r=0.926$). CD III consisted of six adenocarcinomas (5 %) showing in part loss of colonic crypt differentiation immunoprofile inconstantly combined with gain of

metaplastic foveolar gastric differentiation immunoprofile (cytokeratin 20-negative, CDX2-positive, MUC2-positive or MUC2-negative, MUC5AC-positive or MUC5AC-negative). In this group, there was a high level of correlation between the cases ($r=0.541$). CD IV was the most heterogeneous. It consisted of 22 adenocarcinomas (18 %) showing a loss of colonic crypt differentiation immunoprofile and a metaplastic foveolar gastric differentiation immunoprofile (cytokeratin 20-positive or cytokeratin 20-negative, CDX2-negative, MUC2-positive or MUC2-negative, MUC5AC-positive or MUC5AC-negative, cytokeratin 7-positive or cytokeratin 7-negative). The level of correlation between the cases was moderate ($r=0$). The preferential location of the adenocarcinomas belonging to CD I cluster was the left colon (53/83, 64 %) while the preferential location of the CD II, CD III, and CD IV clusters was the right colon (31/39, 79 %) (χ^2 test: $p<0.0001$).

Adenocarcinomas belonging to CD I are associated with MSS status

Genetic analysis based on five markers of microsatellite instability showed that 22 colonic adenocarcinomas

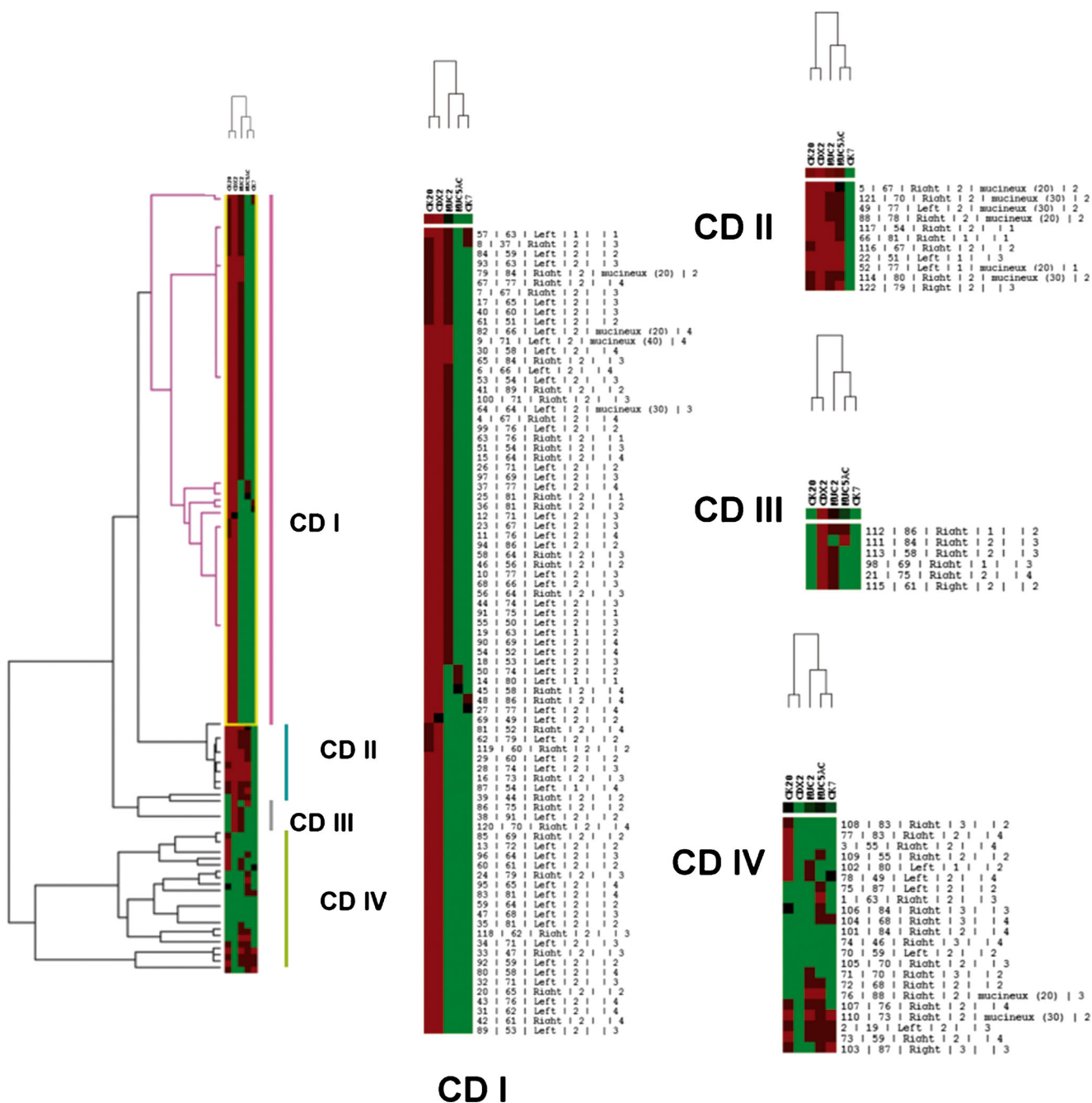


Fig. 2 Hierarchical clustering analysis of 122 colonic adenocarcinomas, NOS. In a first step, the tumors are grouped on the basis of the relatedness of immunostaining with CDX2, cytokeratin 20, MUC2, cytokeratin 7, and MUC5AC. Then, the patients are sorted out according to their degree of relatedness in the selected immunoprofile. The *dendrogram on the left side of the figure* shows two dimensions of clustering, i.e., the association between the markers of differentiation (top) and

association between the patients (*left hand side*). In this analysis, a strong positive score is represented as *bright red blocks*, a weak positive score as *black blocks*, and a negative score as *green blocks*. The *right side of the figure* detailed the four clusters of differentiation already identified on the left side of the figure (CD I, CD II, CD III, and CD IV)

(NOS) were MSI-H (18 %). Nineteen of the 39 tumors belonging to CD II, CD III, and CD IV (49 %) were MSI-H while 3 of the 83 tumors belonging to CD I were MSI-H (4 %), the difference being very significant (χ^2 test: $p < 0.0001$).

Survival analysis

Follow-up data were available for 116 patients with colonic adenocarcinoma (NOS). They showed that 31 patients died from their disease, 12 after 1 year (11 %), 24 after 3 years

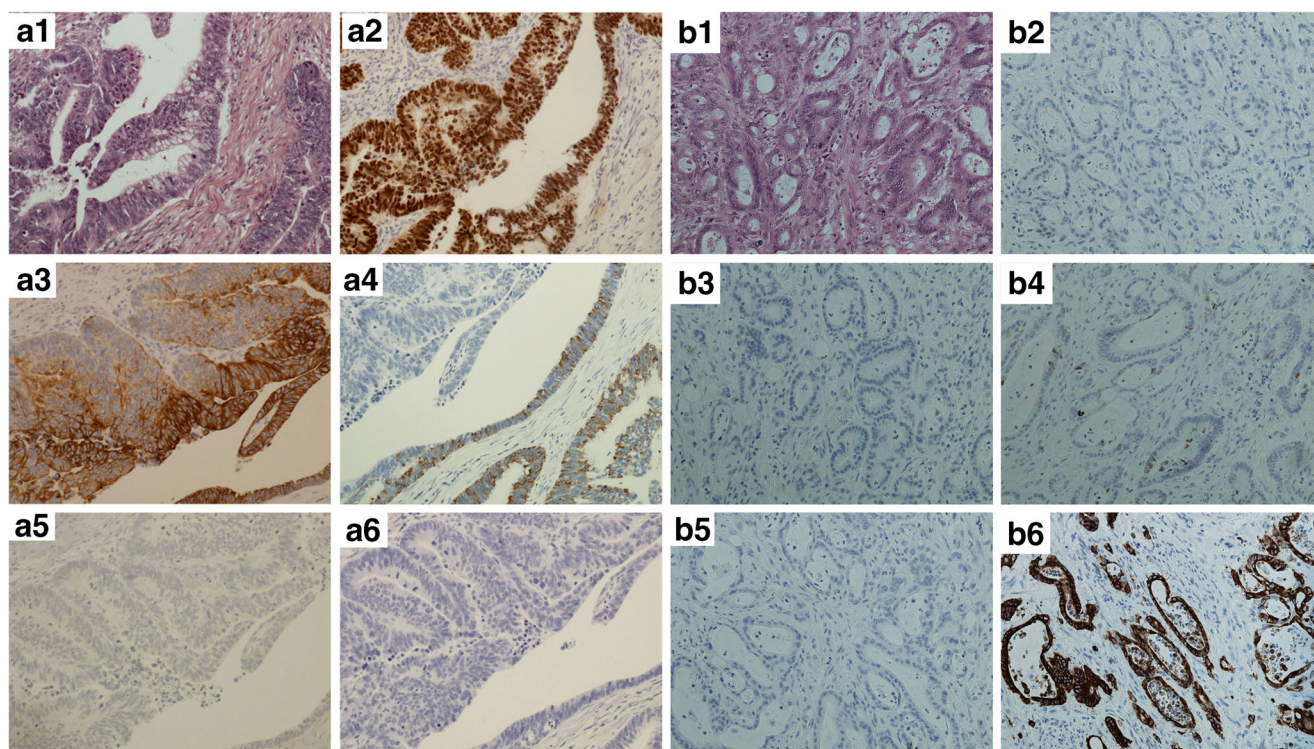


Fig. 3 Representative examples of adenocarcinomas belonging to clusters CD I and CD IV. Histological pattern of adenocarcinoma # ID65 belonging to cluster CD I (A1) and immunophenotypic characteristics: strong expression of CDX2 (A2), cytokeratin 20 (A3), MUC2 (A4), and absence of expression of cytokeratin 7 (A5) and

MUC5AC (A6). Histological pattern of adenocarcinoma # ID106 belonging to cluster CD IV (B1) and immunophenotypic characteristics: absence of expression of CDX2 (B2), cytokeratin 20 (B3), MUC2 (B4) and cytokeratin 7 (B5) and strong cytoplasmic expression of MUC5AC (B6)

(22 %), 28 after 5 years (30 %). Twenty-seven patients developed metachronous metastases including 18 cases of liver metastases, 4 cases of peritoneal metastases, 2 cases of duodenopancreatic metastases, and 1 case each of pulmonary, osseous, and distant lymph node metastases. The mean duration of follow-up was 36 months (range 1–123 months). The prognostic significance of (i) each single differentiation marker, (ii) clinico-pathologic parameters including MSI/MSS status, and finally (iii) the four clusters of differentiation was determined using univariate and multivariate statistical analysis.

In the univariate analysis, the parameters associated with a shorter survival of patients with colonic adenocarcinomas, NOS, were pT status ($p=0.0004$), pN+ status ($p=0.006$), and UICC-TNM stage ($p=0.006$). The prognostic value of the clusters of differentiation exhibited a tendency toward significance ($HR=1.77$; $CI\ 95\ %\ 0.78-4.01$, $p=0.17$) (Table 1). In the multivariate survival analysis using Cox's regression model, only advanced TNM stage had an independent negative association with prognosis ($HR=4.83$; $CI\ 95\ %\ 1.59-14.6$; $p=0.0054$).

In a second analysis of survival, we examined the value of immunophenotypical parameters as well as the MSI-H status in relation with each stage as defined by the UICC, using univariate and multivariate analyses. Expression of some individual markers of differentiation was associated with shorter overall survival of patients with stage III adenocarcinomas ($n=38$): (i)

Table 1 Impact of (i) isolated markers of differentiation, (ii) groups of differentiation determined by hierarchical clustering analysis, (iii) immunophenotypical and genotypical markers of oncogenesis, and (iv) tumor stage on the overall survival of patients with colonic adenocarcinomas, NOS, using univariate analysis Cox proportional hazards regression

	No.	HR	CI 95 %	<i>p</i> value	<i>p</i> value
CDX2 (>50 %)	115	0.54	0.21–1.35	NS	
CK20 (>0 %)	116	0.43	0.16–1.17	NS	
MUC5AC (>0 %)	116	1.06	0.36–3.13	NS	
Clusters CD II, III, or IV	116	1.77	0.78–4.01	NS	
MSI	116	1.61	0.67–3.90	NS	
pT stage	116				0.0004
1		0.00	0.00	NS	
2		0.00	0.00	NS	
3		0.17	0.07–0.38	<0.0001	
pN stage	114	5.47	1.63–18.4		0.0061
UICC stage	116				0.0066
1		0.00	0.00	NS	
2		0.15	0.05–0.47	0.0012	
3		0.32	0.13–0.80	0.0148	

HR hazard risk ratio, CI confidence interval, NS not significant

Table 2 Significance of (i) isolated markers of differentiation, (ii) groups of differentiation determined by hierarchical clustering analysis, and (iii) immunophenotypical and genotypical markers of oncogenesis on the overall survival of patients with colonic adenocarcinomas, NOS, UICC stage III ($n=38$), using univariate analysis Cox proportional hazards regression. CDX2 and cytokeratin 20 (CK20) have a positive impact (HR <1) whereas MUC5AC, clusters CD II, CD III, and CD IV, and MSI have a negative impact (HR >1)

	No.	HR	CI 95 %	<i>p</i> value
CDX2 (>50 %)	38	0.13	0.03–0.60	0.0089
CK20 (>0 %)	38	0.21	0.05–0.85	0.0289
MUC5AC (>0 %)	38	6.13	1.15–32.8	0.0339
Clusters CD II, CD III, or CD IV	38	9.50	2.33–38.7	0.0017
MSI	38	5.76	1.53–21.6	0.0095

HR hazard risk ratio, CI confidence interval, NS not significant

loss of CDX2 expression ≥ 50 % (HR=0.13; IC 95 % 0.03–0.6, $p=0.009$), (ii) loss of cytokeratin 20 expression (HR=0.21; CI 95 % 0.05–0.8, $p=0.03$), and finally (iii) the gain of MUC5AC expression (HR 6.13; CI 95 % 1.15–32.8, $p=0.03$). The MSI-H status was associated with a shorter overall survival (HR=5.76; IC 95 % 1.53–21.6, $p=0.009$). Finally, as shown in Fig. 4 and in Table 2, the overall survival was significantly better in adenocarcinomas of cluster CD I than in adenocarcinomas of clusters II, III, and IV in stage III adenocarcinomas, i.e., carcinomas with lymph nodes metastases (HR=9.5; CI 95 % 2.33–38.7, log-rank test, $p=0.0002$) (Fig. 4). A multivariate Cox's regression model was fed with parameters bearing significant value in univariate analysis. It took into account all individual markers of differentiation with significance in univariate analysis (i.e.,

CDX2 expression ≤ 50 %, loss of cytokeratin 20 expression, gain of MUC5AC expression) as well as MSI-H status. Assignment to clusters CD II, CD III, or CD IV (HR=13.5; CI 95 % 1.94–93.8, $p=0.0085$) was an independent adverse factor for overall survival of patients with stage III carcinoma. A note of caution should be added to this result as the number of cases is relatively small.

Discussion

It is implicitly admitted that colon adenocarcinomas (NOS) form a homogeneous group of tumors on the basis of their histopathological pattern. The rationale of our study was to try to sort out clusters with clinico-pathological relevance [17, 18]. Interestingly, the hierarchical clustering analysis software was adapted to manage findings from TMAs using a limited number of markers [14, 19–21].

In fact, hierarchical clustering analysis isolates a major cluster of homogeneous tumors characterized by the expression of three markers of differentiation of the normal colonic crypt. These three markers were CDX2, a homeobox gene involved in the development and maintenance of intestinal crypt [22, 23], CK20, a cytoskeletal protein of the mature epithelium [24], and MUC2, a secreted mucin restricted to the intestinal goblet cells [25]. Based on this immunoprofile, the term “Crypt-Like Adenocarcinoma” is proposed to identify the tumors belonging to the cluster CD I of differentiation. The biological consistence of this cluster is supported by several lines of investigation showing cross-talk between these

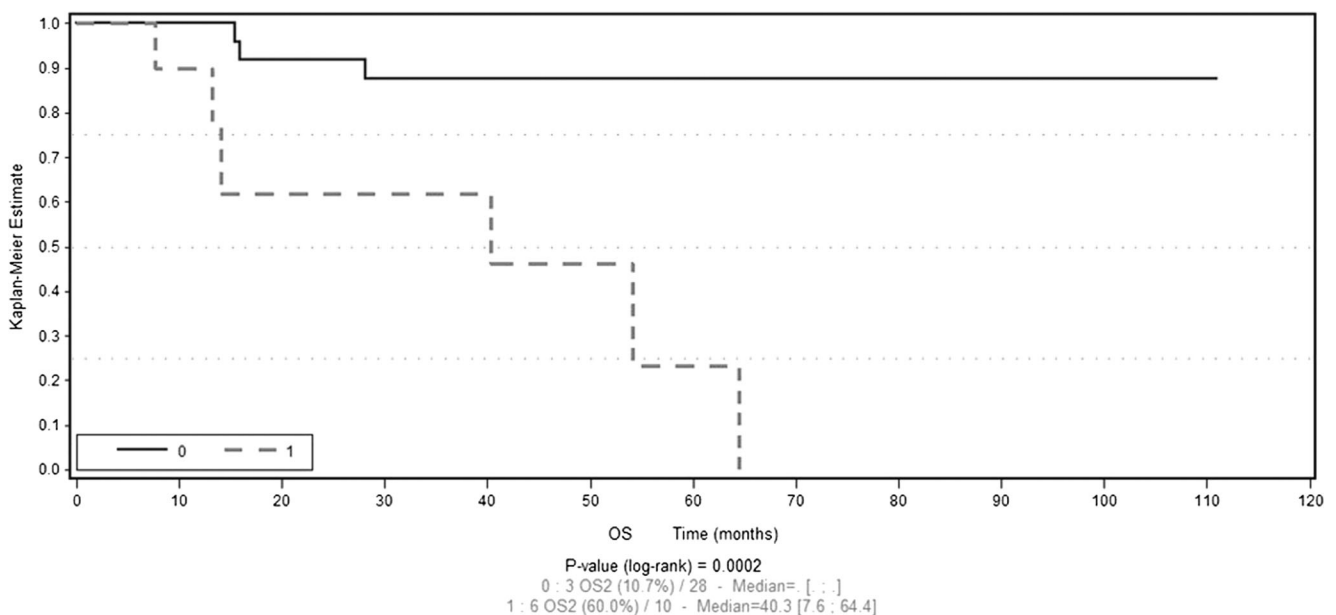


Fig. 4 Kaplan–Meier survival curve of patients with colonic adenocarcinomas, NOS, stage III. Survival of patients with adenocarcinoma belonging to cluster CD I is represented by a solid line

and of patients with adenocarcinoma belonging to clusters II, III, or IV with a dashed line. The difference between the two groups is significant (log-rank's test: $p=0.002$)

markers in regulating both development and homeostasis of the adult colonic crypt. For example, MUC2 expression by goblet cells is initiated by CDX2 and cytokeratin 20 expression is indirectly regulated by CDX2 [26].

Besides “Crypt-like adenocarcinoma,” hierarchical clustering analysis identified three other adenocarcinoma (NOS) clusters that illustrate the immunophenotypical heterogeneity of this group of tumors. These clusters were defined by the markers of differentiation cytokeratin 7, a cytoskeletal protein expressed in fetal stomach and during the intestinal healing process [27, 28], and MUC5AC, a secretory mucin normally restricted to gastric foveolar epithelium [29], together with CDX2, cytokeratin 20, and MUC2. These markers were chosen on the basis of the literature showing expression in relation with the tumor genotype (MSS/MSI) and the clinicopathological setting. In fact, studies have shown that the isolated loss of either CDX2 or cytokeratin 20 is significantly associated with MSI-H [30, 31]. Likewise, cytokeratin 7 is associated with inflammatory bowel disease-associated neoplasia [28] and MUC5AC expression is associated with the so-called serrated pathway of carcinogenesis [32, 33] or villous tumors [34]. However, none of these studies were conducted in a histologically homogeneous subtype of colonic adenocarcinomas. The three minor clusters of differentiation identified so far were characterized by (i) a mixed gastric and colonic immunophenotype, (ii) loss of colon crypt differentiation and acquisition of a gastric immunophenotype, and (iii) acquisition of a gastric immunophenotype including a fetal gastric marker of differentiation (cytokeratin 7) [27]. The immunohistochemical characteristics of these three minor clusters suggest that neoplasia is preceded by metaplasia.

Interestingly, crypt-like carcinomas (CLAs) were strongly associated with MSS status while the other clusters were significantly associated with MSI-H status. The positive predictive value of CLA was nearly 100 % for predicting MSS status. In fact, the inference of MSS status based on histopathological and immunophenotypical characteristics is clinically relevant for patient management because MSS is now considered predictive for the response to conventional 5-FU chemotherapy [35, 36]. Interestingly, Varadhachary et al. [37] who designed a chemotherapeutic approach of metastatic adenocarcinomas with unknown primary tumor observed that only those with a “colon carcinoma” immunoprofile expressing both cytokeratin 20 and CDX2 along with the lack of other markers such as cytokeratin 7, i.e., the markers defining in part our CLA cluster, had a response to 5-FU-based chemotherapy. In addition, in our study, the CLA immunoprofile had an independent prognostic value in stage III adenocarcinomas, traditionally treated with conventional chemotherapy.

In conclusion, this hierarchical clustering analysis isolates among a series of adenocarcinomas (NOS) a main cluster of tumors with a homogeneous immunohistochemical profile similar to that of the colonic crypt. This category of “crypt-

like adenocarcinoma” is diagnostically and prognostically relevant since it is highly predictive of MSS status. Its value as a biomarker of response to conventional chemotherapeutic agents deserves to be examined in comparison with MSS/MSI-H status in randomized therapeutic trial.

Acknowledgments This study is funded by a grant from “Projet Hospitalier de Recherche Clinique” PHRC BRD 07 from the French “Direction de l’Hospitalisation et l’Organisation des Soins” (DHOS). The authors thank the “Photologie” Department for their help.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Jass JR (2007) Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 50:113–130
- Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AM, Bos JL (1988) Genetic alterations during colorectal-tumor development. *N Engl J Med* 319:525–532
- Tomlinson I, Bodmer W (1999) Selection, the mutation rate and cancer: ensuring that the tail does not wag the dog. *Nat Med* 5:11–12
- Fodde R, Kuipers J, Rosenberg C, Smits R, Kielman M, Gaspar C, van Es JH, Breukel C, Wiegant J, Giles RH, Clevers H (2001) Mutations in the APC tumour suppressor gene cause chromosomal instability. *Nat Cell Biol* 3:433–438
- Aaltonen LA, Peltomäki P, Leach FS, Sistonen P, Pylkkänen L, Mecklin JP, Järvinen H, Powell SM, Jen J, Hamilton SR et al (1993) Clues to the pathogenesis of familial colorectal cancer. *Science* 260:812–816
- Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363:558–561
- Thibodeau SN, Bren G, Schaid D (1993) Microsatellite instability in cancer of the proximal colon. *Science* 260:816–819
- Cunningham JM, Christensen ER, Tester DJ, Kim CY, Roche PC, Burgart LJ, Thibodeau SN (1998) Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res* 58:3455–3460
- Gafà R, Maestri I, Matteuzzi M, Santini A, Ferretti S, Cavazzini L, Lanza G (2000) Sporadic colorectal adenocarcinomas with high-frequency microsatellite instability. *Cancer* 89:2025–2037
- Alexander J, Watanabe T, Wu TT, Rashid A, Li S, Hamilton SR (2001) Histopathological identification of colon cancer with microsatellite instability. *Am J Pathol* 158:527–535
- Greenon JK, Huang SC, Herron C, Moreno V, Bonner JD, Tomsho LP, Ben-Izhak O, Cohen HI, Trougouboff P, Bejhar J, Sova Y, Pinchev M, Rennert G, Gruber SB (2009) Pathologic predictors of microsatellite instability in colorectal cancer. *Am J Surg Pathol* 33: 126–133
- Hamilton SR, Bosman FT, Boffetta P, Ilyas M, Morreau H, Nakamura S-I, Quirke P, Riboli E, Sobin LH (2010) Carcinoma of the colon and rectum. In: Bosman FT, Carneiro F, Hruban RH, Theise ND (eds) WHO classification of tumours of the digestive system. IARC Press, Lyon, pp 134–146
- Sobin LH, Brierley J (2009) Colon and rectum. In: Sobin LH, Gospodarowicz MK, Wittekind C (eds) TNM classification of malignant tumours, 7th edn. Wiley-Blackwell, Oxford, pp 100–105

14. Liu CL, Prapong W, Natkunam Y, Alizadeh A, Montgomery K, Gilks CB, van de Rijn M (2002) Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. *Am J Pathol* 161:1557–1565
15. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95:14863–14868
16. Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, Seruca R, Iacopetta B, Hamelin R (2002) Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 123:1804–1811
17. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–511
18. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lønning PE, Børresen-Dale AL (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98:10869–10874
19. Makretsov NA, Huntsman DG, Nielsen TO, Yorida E, Peacock M, Cheang MC, Dunn SE, Hayes M, van de Rijn M, Bajdik C, Gilks CB (2004) Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically significant groups of breast carcinoma. *Clin Cancer Res* 10:6143–6151
20. Jacquemier J, Ginestier C, Rougemont J, Bardou VJ, Charafe-Jauffret E, Geneix J, Adélaïde J, Koki A, Houvenaeghel G, Hassoun J, Maraninchi D, Viens P, Birnbaum D, Bertucci F (2005) Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. *Cancer Res* 65:767–779
21. Alkushi A, Clarke BA, Akbari M, Makretsov N, Lim P, Miller D, Magliocco A, Coldman A, van de Rijn M, Huntsman D, Parker R, Gilks CB (2007) Identification of prognostically relevant and reproducible subsets of endometrial adenocarcinoma based on clustering analysis of immunostaining data. *Mod Pathol* 20:1156–1165
22. Lorentz O, Duluc I, Arcangelis AD, Simon-Assmann P, Kedinger M, Freund JN (1997) Key role of the Cdx2 homeobox gene in extracellular matrix-mediated intestinal cell differentiation. *J Cell Biol* 139:1553–1565
23. Silberg DG, Swain GP, Suh ER, Traber PG (2000) Cdx1 and cdx2 expression during intestinal development. *Gastroenterology* 119:961–971
24. Moll R, Löwe A, Laufer J, Franke WW (1992) Cytokeratin 20 in human carcinomas. A new histodiagnostic marker detected by monoclonal antibodies. *Am J Pathol* 140:427–447
25. Chang SK, Dohrman AF, Basbaum CB, Ho SB, Tsuda T, Toribara NW, Gum JR, Kim YS (1994) Localization of mucin (MUC2 and MUC3) messenger RNA and peptide expression in human normal intestine and colon cancer. *Gastroenterology* 107:28–36
26. Uesaka T, Kageyama N, Watanabe H (2004) Identifying target genes regulated downstream of Cdx2 by microarray analysis. *J Mol Biol* 337:647–660
27. Stosiek P, Bräutigam E, Kasper M (1991) Expression of cytokeratin 7 in human glandular epithelium of fetal stomach. *Acta Histochem* 91:21–23
28. Tatsumi N, Kushima R, Vieth M, Mukaisho K, Kakinoki R, Okabe H, Borchard F, Stolte M, Okanoue T, Hattori T (2006) Cytokeratin 7/20 and mucin core protein expression in ulcerative colitis-associated colorectal neoplasms. *Virchows Arch* 448:756–762
29. Forgue-Lafitte ME, Fabiani B, Levy PP, Maurin N, Fléjou JF, Bara J (2007) Abnormal expression of M1/MUC5AC mucin in distal colon of patients with diverticulitis, ulcerative colitis and cancer. *Int J Cancer* 121:1543–1549
30. Hinoi T, Tani M, Lucas PC, Caca K, Dunn RL, Macri E, Loda M, Appelman HD, Cho KR, Fearon ER (2001) Loss of CDX2 expression and microsatellite instability are prominent features of large cell minimally differentiated carcinomas of the colon. *Am J Pathol* 159:2239–2248
31. McGregor DK, Wu TT, Rashid A, Luthra R, Hamilton SR (2004) Reduced expression of cytokeratin 20 in colorectal carcinomas with high levels of microsatellite instability. *Am J Surg Pathol* 28:712–718
32. Biemer-Hüttmann AE, Walsh MD, McGuckin MA, Simms LA, Young J, Leggett BA, Jass JR (2000) Mucin core protein expression in colorectal cancers with high levels of microsatellite instability indicates a novel pathway of morphogenesis. *Clin Cancer Res* 6:1909–1916
33. Yao T, Nishiyama KI, Oya M, Kouzuki T, Kajiwara M, Tsuneyoshi M (2000) Multiple ‘serrated adenocarcinomas’ of the colon with a cell lineage common to metaplastic polyp and serrated adenoma. Case report of a new subtype of colonic adenocarcinoma with gastric differentiation. *J Pathol* 190:444–449
34. Buisine MP, Janin A, Maunoury V, Audié JP, Delescaut MP, Copin MC, Colombel JF, Degand P, Aubert JP, Porchet N (1996) Aberrant expression of a human mucin gene (MUC5AC) in rectosigmoid villous adenoma. *Gastroenterology* 110:84–91
35. Iacopetta B, Kawakami K, Watanabe T (2008) Predicting clinical outcome of 5-fluorouracil-based chemotherapy for colon cancer patients: is the CpG island methylator phenotype the 5-fluorouracil-responsive subgroup? *Int J Clin Oncol* 13:498–503
36. Boland CR, Goel A (2010) Microsatellite instability in colorectal cancer. *Gastroenterology* 138:2073–2087
37. Varadhachary GR, Raber MN, Matamoros A, Abbruzzese JL (2008) Carcinoma of unknown primary with a colon-cancer profile-changing paradigm and emerging definitions. *Lancet Oncol* 9:596–599