Genomic and Epigenomic Integration Identifies a Prognostic Signature in Colon Cancer

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

Clinical Cancer Research

Imaging, Diagnosis, Prognosis

See commentary p. 1215

Genomic and Epigenomic Integration Identifies a Prognostic Signature in Colon Cancer

Joo Mi Yi[1,9], Mashaal Dhir[2], Leander Van Neste[4,5], Stephanie R. Downing[2], Jana Jeschke[2], Sabine C. Glöckner[2], Marilia de Freitas Calmon[2], Craig M. Hooker[1], Juan M. Funes[4,6], Chris Boshoff[4,6], Kim M. Smits[7], Manon van Engeland[7], Matty P. Weijenberg[7], Christine A. Iacobuzio-Donahue[3], James G. Herman[1], Kornel E. Schuebel[8], Stephen B. Baylin[1], and Nita Ahuja[1,2]

Abstract

Purpose: The importance of genetic and epigenetic alterations maybe in their aggregate role in altering core pathways in tumorigenesis.

Experimental Design: Merging genome-wide genomic and epigenomic alterations, we identify key genes and pathways altered in colorectal cancers (CRC). DNA methylation analysis was tested for predicting survival in CRC patients using Cox proportional hazard model.

Results: We identified 29 low frequency-mutated genes that are also inactivated by epigenetic mechanisms in CRC. Pathway analysis showed the extracellular matrix (ECM) remodeling pathway is silenced in CRC. Six ECM pathway genes were tested for their prognostic potential in large CRC cohorts ($n = 777$). DNA methylation of *IGFBP3* and *EVL* predicted for poor survival (*IGFBP3*: HR = 2.58, 95% CI: 1.37–4.87, $P = 0.004$; *EVL*: HR = 2.48, 95% CI: 1.07–5.74, $P = 0.034$) and simultaneous methylation of multiple genes predicted significantly worse survival (HR = 8.61, 95% CI: 2.16–34.36, $P < 0.001$ for methylation of *IGFBP3*, *EVL*, *CD109*, and *FLNC*). DNA methylation of *IGFBP3* and *EVL* was validated as a prognostic marker in an independent contemporary-matched cohort (*IGFBP3* HR = 2.06, 95% CI: 1.04–4.09, $P = 0.038$; *EVL* HR = 2.23, 95% CI: 1.00–5.0, $P = 0.05$) and *EVL* DNA methylation remained significant in a secondary historical validation cohort (HR = 1.41, 95% CI: 1.05–1.89, $P = 0.022$). Moreover, DNA methylation of selected ECM genes helps to stratify the high-risk stage 2 colon cancers patients who would benefit from adjuvant chemotherapy (HR: 5.85, 95% CI: 2.03–16.83, $P = 0.001$ for simultaneous methylation of *IGFBP3*, *EVL*, and *CD109*).

Conclusions: CRC that have silenced genes in ECM pathway components show worse survival suggesting that our finding provides novel prognostic biomarkers for CRC and reflects the high importance of integrative analyses linking genetic and epigenetic abnormalities with pathway disruption in cancer. *Clin Cancer Res; 17(6); 1535–45. ©2011 AACR.*

Introduction

Recent efforts in genome-wide sequencing of multiple human cancers have identified a large number of mutations that cluster within a small number of intracellular pathways such as cellular senescence bypass, invasion, and metastasis (1, 2). Epigenetic abnormalities can also operate with genetic alterations to effect aberrant gene function in cancer (3). DNA hypermethylation of promoter- associated CpG islands is a frequent and early mechanism of tumor-suppressor gene inactivation in many cancers and may have great promise for identifying biomarkers (4–6). We have previously emphasized that DNA hypermethylation can affect many of the new genes found to be mutated in both colorectal (CRC) and breast cancer cells (7, 8) However, an understanding of such genome wide alterations, by itself, has not led to many clinically applicable biomarkers in predicting tumor behavior. We hypothesized that integration of the genome and hypermethylome may provide

Authors' Affiliations: Departments of [1]Oncology, [2]Surgery and [3]Pathology, Johns Hopkins University, Baltimore, Maryland; [4]Faculty of Bioscience Engineering, Department of Molecular Biotechnology, Ghent University, Ghent, Belgium; [5]MDxHealth, Ghent, Belgium [6]Cancer Research UK Viral Oncology Group, University College London, London, United Kingdom; [7]Department of Pathology, GROW School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, the Netherlands; [8]Laboratory of Neurogenetics, NIAAA, NIH, Bethesda, Maryland; [9]Research Center, Dongnam Institute of Radiological & Research Medical Sciences (DIRAMS), Busan, South Korea

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Note: J.M. Yi, M. Dhir, and L. Van Neste contributed equally to this study.

Corresponding Authors: Nita Ahuja, Department of Surgery and Oncology, Johns Hopkins University, Baltimore, MD 21231. Phone: 1-410-502-6135; Fax: 1-410-502-0987. E-mail: nahuja@jhmi.edu or Stephen B. Baylin, Department of Oncology, Johns Hopkins University, Baltimore, MD 21231; Kornel Schuebel, Laboratory of Neurogenetics, NIAAA, NIH, Bethesda, MD 20892-8110

doi: 10.1158/1078-0432.CCR-10-2509

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

Translational Relevance

We have used an exciting integrative approach of multiple whole genome analysis (genetic and epigenetic) to identify an important pathway, the extracellular matrix (ECM) maintenance, and remodeling pathway, which is silenced in all colon cancers. Colon cancers that have silenced multiple genes in this pathway have a poor prognosis based on analyses of large cohort of patients.

DNA methylation of these genes is particularly useful in stratifying the low- versus high-risk stage 2 colon cancer patients. Colorectal cancers that have silenced multiple genes in the ECM pathway appear to show a significantly worse survival in adjusted multivariate analysis of a large cohort of colon cancer patients. Our study has important ramifications in the clinical management of colon cancer patients as 30% to 40% of early stage colon cancers will present with recurrence/metastases but current strategies are not helpful in stratifying this high-risk cohort. We show that our prognostic markers are the most efficacious in stratifying the high-risk subset of stage 2 colon cancers who may then benefit from adjuvant chemotherapy.

insights into the key pathways altered in human cancers and lead to insights into relevant biomarkers. In the present study, we have continued to globally search for epigenetic alterations of genes also mutated in colon cancers to narrow down the key pathways altered in CRC.

We now find that many genes in the extracellular matrix (ECM) maintenance and remodeling pathway, critical for many normal developmental processes and for tumorigenesis (9), are epigenetically altered in nearly every human colon cancer. ECM can be remodeled by many processes, including changes in synthesis, contraction, and proteolytic degradation (9). Recently, misregulated ECM environments have also been implicated in the capacity of tumors to undergo epithelial-mesenchymal transition (EMT) and in the acquisition of metastatic potential in tumors as cells lose cell-cell and cell-ECM adherence along with profound changes in their cystoskeleton architecture (10–12). We now show that ECM pathways alterations are frequent in all colon cancers and DNA methylation-induced silencing of these genes can occur at multiple genes in the ECM pathway.

Importantly, we also now show that the simultaneous DNA hypermethylation of a subset of these ECM genes is also associated significantly with poor survival in adjusted analyses of a large cohort of colon cancer patients. Our data highlight the importance of integrative approaches in obtaining a better understanding of carcinogenesis as well as identifying novel biomarkers.

Materials and Methods

Patients

Johns Hopkins (JHU) training and validation cohort samples for the colorectal cancer (CRC) studies were prepared from formalin-fixed and paraffin-embedded (FFPE) colon cancer tissue samples from the pathology archives of the JHU Hospital in accordance with the Institutional Review Board (IRB) and as per HIPAA compliance. The JHU training cohort consisted of 147 tissue samples from stage 1 to 4 colon cancer patients who underwent primary surgery from 1995 to 2005 (median follow-up of 6.4 years). The JHU validation cohort consisted of 72 tissue samples from stage 1 to 4 colon cancer patients who underwent primary surgery from 1995 through 2005 (median follow-up of 7.5 years). Patients with neoadjuvant chemotherapy, including all rectal cancers, were excluded from the current study. Patients in both JHU cohorts were similar with respect to age, sex, race, proportion of cases with lymphovascular invasion, pathologic grade, location, and proportion of cases with recurrence (Table 1).

Netherlands Cohort study on Diet and Cancer (NLCS) Validation cohort were prepared from 558 FFPE colon cancers from stages 1 to 4 (median follow-up of 7.2 years). The NLCS has been previously described in detail (13). A total of 925 incident CRC cases were identified from 1989 to 1994, and DNA was available for analyses on our current subset (Table 1).

DNA methylation and gene expression analyses

Primer pairs were preferentially designed near the putative transcriptional start site (TSS) in the 5′-CpG islands of the genes. Primer sequences for methylation-specific PCR (MSP) analysis were designed using MSP Primer (7). All primer sequences are listed in Supplementary Table S5. For expression studies using reverse transcription PCR (RT-PCR), primers were designed using the open access program Primer3.

For MSP analysis, DNA was extracted using the standard phenol–chloroform extraction method. Bisulfite modification of genomic DNA was carried out using the EZ DNA methylation Kit (Zymo Research). DNA methylation analysis was performed using MSP primer pairs located close to the putative transcriptional start site in the 5′-CpG island with 2 μL of bisulfite-treated DNA as template and JumpStart Red Taq DNA Polymerase (Sigma) for amplification as previously described (7).

Total RNA was extracted from cell lines using the RNeasy Mini Kit (Qiagen), treated with DNase (Qiagen). For RT reaction, 1 μg of total RNA was subjected to the first strand cDNA synthesis using Superscript III first strand cDNA synthesis kit (Invitrogen) according to the manufacturer's instructions. Expression analysis was performed by RT-PCR using 1 μL of cDNA as template and JumpStart Red Taq DNA Polymerase (Sigma) for amplification.

Quantitative methylation analysis

Quantitative MSP (qMSP) was performed after sodium bisulfite modification for selected genes including *IGFBP3*, *EVL*, *CD109*, *FLNC*, and *FBN2* on a small series of normal colons and colon cancers ($n = 5$, each; Supplementary Fig. S2). qMSP analysis helped us to confirm whether our candidate genes showed cancer-specific DNA methylation

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

Table 1. Demographics of training and validation cohorts in this study

	JHU		NLCS set (the Netherlands)
	Training cohort	Validation cohorts	
Time period	1998–2005	1998–2005	1986–1991
N	*n* = 147	*n* = 72	*n* = 558
Median age, y	66	70	64
5-y survival	67.4%	65.3%	54.8%
Stage I	83.3%	83.3%	78.4%
Stage II	67.4%	70.4%	70.2%
Stage III	75.7%	60.0%	42.0%
Stage IV	17.0%	14.3%	1.6%
Gender			
Male	84 (57%)	30 (42%)	296 (53%)
Female	63 (43%)	42 (58%)	262 (47%)
Location			
Right colon	79 (54%)	40 (56%)	239 (43%)
Left colon	66 (45%)	32 (44%)	310 (56%)
Unknown			9 (1%)
Stage			
Stage I	36 (24.4%)	18 (25.0%)	111 (19%)
Stage II	49 (33.3%)	27 (37.5%)	201 (36%)
Stage III	37 (25.2%)	20 (27.8%)	143 (25%)
Stage IV	25 (17.1%)	7 (9.7%)	62 (11%)
Unknown			41 (9%)

patterns consistent with the qualitative MSP analysis. For quantitative real-time analyses, the Power SYBR Green PCR kit (Applied Biosystems) was used and the amplification conditions consisted of an initial 10-minute denaturation step at 95°C, followed by 40 cycles of denaturation at 95°C for 15 s and annealing and extension for 30 s and 60 s, respectively. An ABI StepOnePlus Real-Time PCR System was used (Applied Biosystems), and for quantitation the comparative cycle threshold (Ct) method was used, normalizing the Ct values for the indicated gene to the Ct values of unmethylated reaction relative to a methylated reaction sample.

Cell culture and treatment

Cancer cell lines (CRC cell lines: HCT116, SW480, RKO, HT29, Caco-2, Lovo, COLO 320, COLO 205, DLD1, SW48, and SW620; breast cancer cell lines: MCF7, T47D, MDA-MB-231, and MDA-MB-468) were obtained from American Type Culture Collection (ATCC) and cultured in appropriate media and under conditions described by ATCC, with media obtained from Invitrogen, supplemented with 10% fetal bovine serum (Gemini Bio-Products) and 1% penicillin/streptomycin (Invitrogen). DKO cells (HCT116 cells with genetic disruption of *DNMT1* and *DNMT3b*) were cultured as described previously (7).

Signaling pathway analysis of methylated genes

We used a highly curated database (Metacore) that includes human protein–protein interactions, signal transduction, and metabolic pathways, and a variety of cellular functions and processes for signaling pathway analysis. Functional ontology enrichment selected the significantly altered canonical pathways, networks, and gene ontology terms. Analysis was run to detect deregulated interaction networks in CRC, using a subset of genes that are both mutated (1, 2) and hypermethylated in CRC tissue. Significant pathways were selected and intersected with those found by Wood and colleagues (2). While the article by Wood and colleagues (2) was focusing on those networks targeted by mutations alone, the combination of both approaches was designed to zoom in on those pathways that are important to CRC as a whole through alterations of these molecular interaction networks, be it by gene mutations or promoter hypermethylation.

Survival and statistical analysis

Univariate and multivariate Cox regression was used to identify genes that could significantly predict poor overall survival. All ties were handled by the Breslow method, and the proportional hazards assumption was verified by examination of residual plots. Adjusted Kaplan–Meier survival plots were created to demonstrate the ability of the genes to predict poor overall survival. Statistical significance in this study was set at 0.050. To control for multiple testing, Sidak adjustment method was used to correct for family-wise error rate. Results of all models are reported as odds ratios with 95% CIs. All data were analyzed using StataTM v9.2 statistical analyses software (Stata Corporation).

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

Results

DNA methylation analysis of mutated genes

Using the recently compiled database of human cancer mutations in CRC and breast cancers, termed the candidate cancer genes (or CAN genes; refs 1, 2), we have now identified 119 genes which, in colon and breast cancer cell cultures, are also identified by our recently published gene discovery approaches (7, 8) as candidate genes for promoter DNA hypermethylation. As the majority of these genes were found to be hypermethylated in colon tumors, we focused the remainder of our studies on colon cancers. These 119 genes then represent the intersection of the consensus-mutated and DNA-hypermethylated genes in colon cancer. Seventy-five of these 119 genes contained promoter CpG islands (14) and were then first experimentally validated for (1) expression in normal colon tissues by RT-PCR ($n = 65$ genes); Of these, 45 genes showed (2) correlation between loss of gene expression by RT-PCR and promoter CpG island DNA methylation by MSP analysis in colon cancer cell lines (HCT116, RKO, SW480, and Colo320; Fig. 1A, Supplementary Table S1). To filter down to genes that only showed cancer-specific DNA methylation, we next looked at DNA methylation frequencies of these 45 genes in a series of primary colon cancers ($n = 21$, stages 2 and 3) as well as normal colon from cancer-free individuals ($n = 20$). (3) Thirty-four genes showed more than 5% frequency of DNA methylation in a series of primary colon cancers and (4) 29 of these genes showed cancer-specific DNA methylation with the absence of DNA methylation in normal colon (Supplementary Table S1 for schema and genes; Supplementary Table S2 for primary colon cancer sample information). Cancer-specific DNA methylation was also confirmed using qMSP for a subset of genes and showed identical results to the gel-based MSP assay (Supplementary Fig. S2).

A subset of these genes had been previously identified by us during genome-wide screens (7, 8) and were included in the current integrative approaches to provide a more comprehensive view of frequent epigenetic alterations in colon cancer. Among the 29 genes, 15 genes are newly identified by this integrative approach as potential targets of both mutation and DNA methylation (*FLNC*, *SH3TC1*, *ZNF569*, *SLC22A15*, *HAPLN1*, *PRKD1*, *IGFBP3*, *TCERG1L*, *LAMA1*, *FBN2*, *CPAMD8*, *NTNG1*, *EYA4*, *GRID1*, and *PPM1E*) in this study with 5% to 100% DNA methylation frequencies in CRC (Fig. 1B), whereas the remaining 14 genes were also previously identified (7). A direct comparison of DNA methylation and known mutation frequencies of these 29 genes reveals that they are more often epigenetically than genetically altered in colon cancer (Supplementary Fig. S1). DNA methylation frequencies of the 29 genes show variable frequencies in the primary colon cancers ranging from 5% to 100% (Supplementary Fig. S1). 17 out of 29 genes (*GUCY1A2*, *NRCAM*, *PTPRD*, *PPM1E*, *STARD8*, *GRID1*, *EVL*, *EYA4*, *NTNG1*, *CPAMD8*, *FBN2*, *MMP2*, *SYNE1*, *LAMA1*, *TCERG1L*, *APC2*, and *GPNMB*) displayed high level ($\geq$50%) hypermethylation in colon cancers, whereas 12 genes (*SLC22A15*, *ICAM5*, *HAPLN1*, *CHD5*, *RET*, *LGR6*, *PRKD1*, *IGFBP3*, *FLNC*, *CD109*, *ZNF569*, and *SH3TC1*) showed lower frequency of DNA methylation (<50%) in colon cancers (Fig. 1B).

Pathway analysis identifies global disruption of the ECM pathway in CRC

A growing body of research is defining that, particularly with respect to new genes being discovered in cancer resequencing efforts and those genes which are genetically altered with low frequency, alterations of pathways rather than of individual genes may be the important issue for affecting the course of tumorigenesis (1, 2). These mutated genes, then, belong to a handful of important signaling pathways in breast, colorectal, and other cancer types (2, 15). We queried an extensive gene ontology set (MetaCore; GeneGo) and nearly half of the genes that we identified to be both mutated and hypermethylated (13 of 29, 44.8%) mapped to the ECM pathway in CRC (Fig. 2). ECM remodeling is critical for many developmental processes, and recently, abnormally remodeled ECM has been shown to contribute to the neoplastic process (9).

We, therefore, focused the remainder of our studies to understanding the functional implications of the dysregulation of the ECM pathway in CRC. In our analyses of these ECM pathway genes, 8 of the genes (*NRCAM*, *EVL*, *NTNG1*, *CPAMD8*, *FBN2*, *MMP2*, *LAMA1*, and *GPNMB*) show high frequency (>50%) of DNA methylation in CRC and are amongst those mapping to the ECM pathway (refs. 10, 16, 17; Fig. 2), and the remainder show lower frequency of DNA methylation in CRC (*ICAM5*, *HAPLN1*, *IGFBP3*, *FLNC*, and *CD109*; refs 18–20).

In addition, we also noted that genes such as *TIMP2* and *TIMP3* that have previously been identified by ourselves and others as being hypermethylated in colorectal and other cancers also belong to the ECM pathways (21, 22). Based on this observation, we also investigated, for potential epigenetic silencing in CRC, other genes in the ECM remodeling pathway which have previously been identified to show hypermethylation in other cancers by us and others. Six other genes, in addition to *TIMP2* and *TIMP3*, were identified, including *PLAU*, *Nidogen* (*NID1*), *Osteonectin*, *CD44*, *MMP9*, and *Laminin 5* (*LAMA5*; refs 23–26) and these 8 genes were then investigated for potential DNA methylation in our current set of primary CRC. Six of these 8 ECM genes, including *TIMP2*, *TIMP3*, *MMP9*, *PLAU*, *Nidogen*, and *Osteonectin* were found to be methylated in CRC with variable DNA methylation frequencies, ranging from 5% to 95% (Supplementary Table S3). All 8 additional ECM genes also showed no evidence of DNA methylation in normal colon (data not shown; $n = 5$). The *TIMPs* (*TIMP2* and *TIMP3*) and *MMPs* (*MMP9*) have long been thought to be an integral components of the ECM pathway and their expression has been critical in primary tumor growth, angiogenesis, invasion, and metastasis in different type of cancers (18). *Nidogen* forms a complex with laminins and

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

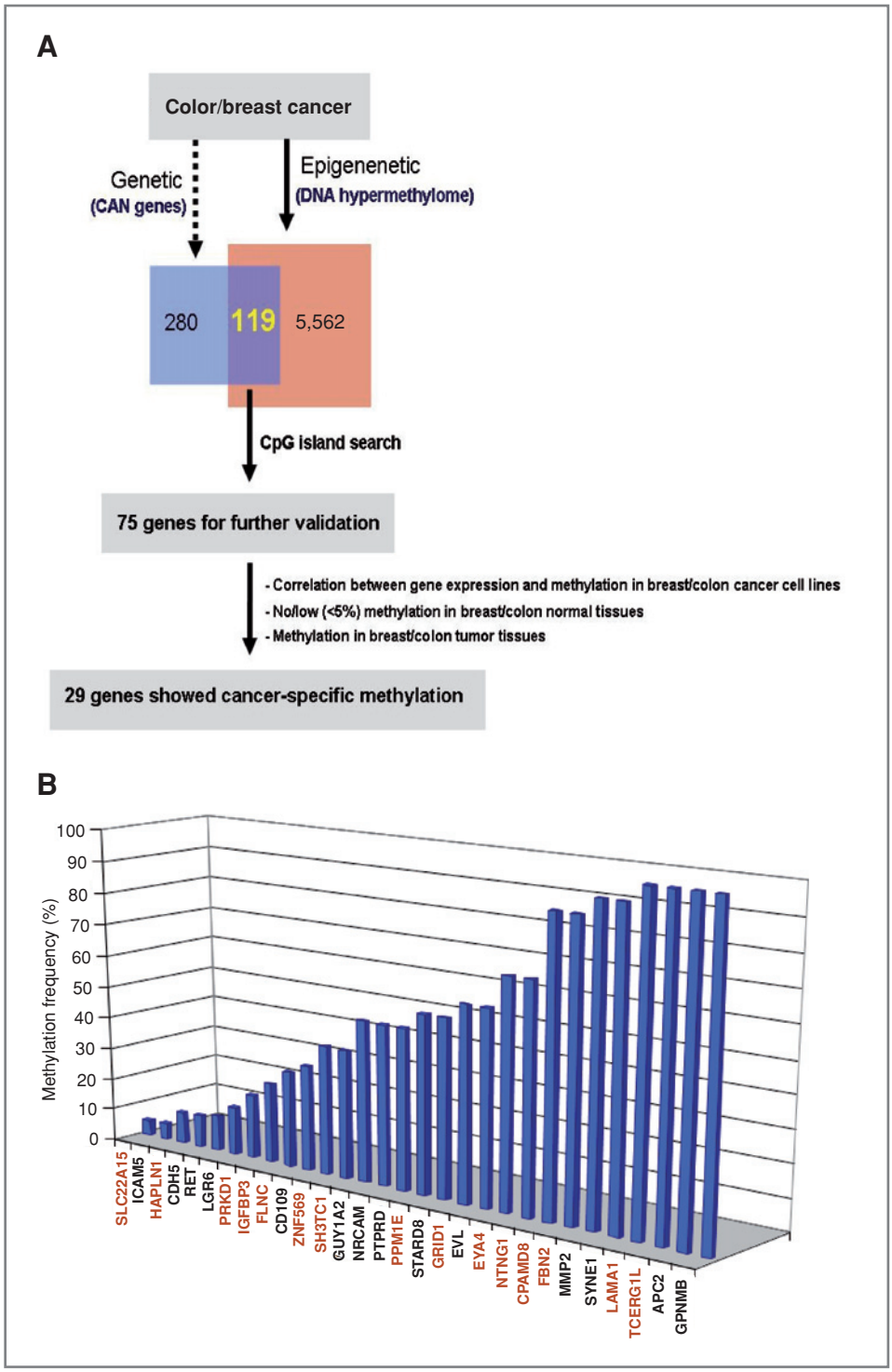
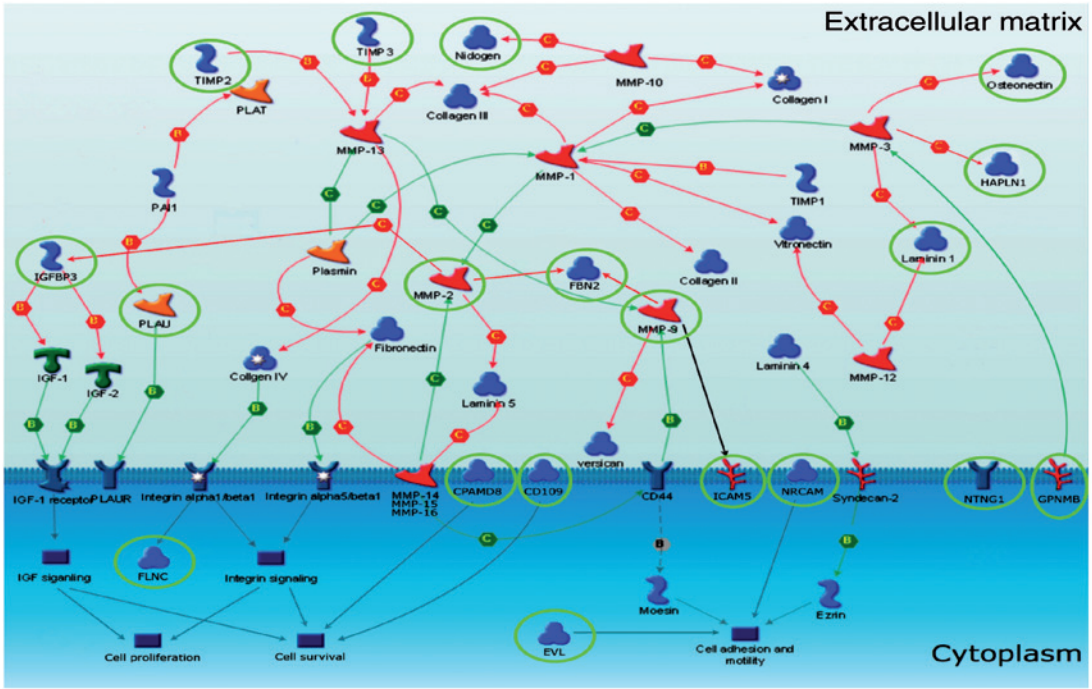


Figure 1. Identification of mutated and DNA hypermethylated genes in colon cancers. A, schematic representation of the intersection of consensus-mutated and DNA hypermethylated genes in colon cancer (see also Supplementary Table S1 for schema). B, DNA methylation frequencies of the 29 low-frequency-mutated genes in CRC patients ($n = 20$). A subset of genes (black letters) have been noted in our previous studies (7, 8) and newly identified genes in this study are indicated with red letters.

collagen IV and helps to stabilize the basement membrane structure as well as being important for cell adhesion by establishing contact with integrins (24). *Osteonectin* (also known as *SPARC*) and *PLAU* are involved in wound repair (25), cell migration, and differentiation but their role in tumorigenesis has not yet been defined.

Together, all our data show that at least 19 genes in the ECM pathway are hypermethylated in CRC (Fig. 2 and Supplementary Table S3), and strongly suggests that the ECM pathway is globally and universally altered in CRC by mutation or epigenetic silencing but much more commonly by the latter abnormality.

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509



Gene name	RefSeq ID	Description
CPAMD8	NM_015692	Complement 3 D8; transmembrane protein that binds heparin
CD109	NM_133493	GPI anchored member of a2 macroglobulin C3/4/5 family; involved in RET signaling; linked to TGFb signaling and cytoskeleton maintenance
EVL	NM_016337	Ena/ Vasp-like; links profilin with actin cytoskeleton via GTPase Rap1; involved in cell spreading and lamellopodia formation
FBN2	NM_001999	Fibrillin; colocalizes with laminin, collagen, fibronectin, versican, and others
FLNC	NM_001458	Filamin c; regulates actin cytoskeleton via interaction with fibronectin
GPNMB	NM_001005340	Osteoactivin; regulates MMP3 activity
HAPLN1	NM_001884	Hyaluronan and proteoglycan link protein 1
ICAM5	NM_003259	Binds ERM (ezrin/moesin/radixin) proteins to regulate actin cytoskeleton
IGFBP3	NM_000598	Regulates IGF1 activity through direct interaction
LAMA1	NM_005559	Laminin alpha 1; basement membrane protein
MMP2	NM_004530	Matrix metalloproteinase 2; enhances cell–cell adhesion cleavage of ECM proteins fibronectin and vitronectin into fragments and facilitates binding to β-integrin receptors
NTNG1	NM_001113226	Netrin G1 transmembrane receptor mediates cell matrix adhesion
NRCAM	NM_001037132	Neuron related cell adhesion molecule; β-catenin target
TIMP3	NM_000362	Tissue inhibitor of metalloproteinase 3; inhibitory role against metalloproteinases
TIMP2	NM_003255	TIMP metallopeptidase inhibitor 2 precursor, inhibitory role against metalloproteinases
Nidogen	NM_002508	Basement membrane glycoproteins; play a role in cell interactions with the extracellular matrix.
Osteonectin	NM_003118	Secreted protein acidic and rich in cysteine/osteonectin/BM40, or SPARC
MMP9	NM_004994	Matrix metalloproteinase 9 preproprotein; secreted as inactive proproteins which are activated when cleaved by extracellular proteinases
PLAU	NM_002658	Urokinase plasminogen activator preproprotein; involved in degradation of the extracellular matrix and possibly tumor cell migration and proliferation

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

CRC with ECM dysregulation demonstrate worse survival

We next wanted to study the clinical correlates of our gene silencing data in terms of the previously observed juxtaposition of embryonic gene expression patterns and EMT with such tumors having the most primitive and aggressive phenotypes (27). In this respect, an exploratory clinical outcome analysis in our initial small cohort of colon cancer patients ($n = 21$ patients) was performed for hypothesis-generation and showed that DNA methylation of selected ECM genes (*IGFBP3* and *EVL*) was associated with significantly worse outcomes on unadjusted analysis ($P = 0.02$ and $P = 0.03$, respectively), whereas DNA methylation of other genes (*CD109*, *NRCAM*, *NTNG*, and *FLNC*) showed a trend toward poor survival which was not statistically significant. To explore the potential clinical significance of these data, we tested these above 6 ECM pathway genes for their prognostic potential in a large training cohort of patients with colon cancers and well-annotated clinical data that could be correlated with survival and gene methylation status. Strikingly, in an analyses of tumors from 147 patients with stage 1 to 4 colon cancers (JHU training cohort; Table 1), we observed a statistically significant increased risk for mortality, when either individual genes or combinations of genes were methylated after adjusting for other prognostic variables such as age and stage of cancers (Figs. 3A and 4A). For example, adjusted Kaplan–Meier survival curves for 2 of the individual ECM genes, *IGFBP3* and *EVL*, show that DNA methylation of each is associated with decreased survival (adjusted for age and stage; Fig. 3A). Multivariate HRs (Fig. 4A) for survival depicted in a forest plot show the HRs for DNA methylation of these 2 genes predicted risk for poor survival (*IGFBP3* HR = 2.58, 95% CI, 1.37–4.87, $P = 0.004$; *EVL* HR = 2.48, 95% CI, 1.07–5.74, $P = 0.034$), whereas another gene *CD109* shows a trend toward statistical significance (HR = 1.81, 95% CI, 0.94–3.50, P = NS). Moreover, in addition, in patients with simultaneous DNA methylation of multiple genes in the ECM pathway, significantly worse survival is predicted in adjusted analysis. For example, for DNA methylation of *IGFBP3*, *EVL*, *CD109*, and *FLNC*, the odds of dying are increased to greater than eight-fold after adjusting for age and stage of cancer compared with patients who lack DNA methylation of these genes (HR: 8.61, 95% CI: 2.16–34.36, $P < 0.001$) (Fig. 4A).

We further examined DNA methylation status of *IGFBP3*, *EVL*, *CD109*, *NRCAM*, and *FLNC* in an independent validation cohort of 72 patients with stage 1 to 4 colon cancers (Validation cohort JHU; Table 1 and Supplementary Table S4). In multivariate-adjusted analysis, both *IGFBP3* (HR: 2.06, 95% CI: 1.04–4.09, $P = 0.038$) and *EVL* (HR: 2.23, 95% CI: 1.00–5.06, $P = 0.050$) remained statistically significant as markers associated with worse survival. *CD109* showed a trend toward worse survival outcome that was not statistically significant (HR: 1.49, 95% CI: 0.72–3.08, P = NS), whereas *FLNC* and *NRCAM* had no prognostic value. Adjusted Kaplan–Meier survival curves (adjusted for age and stage) for 2 of the individual ECM genes, *IGFBP3* and *EVL*, again show that DNA methylation of each is associated with decreased survival in our validation cohort (Fig. 3A). Multivariate HRs (Fig. 4B) for survival depicted in a forest plot of both the training and validation cohort shows that simultaneous DNA methylation of multiple genes in the ECM pathway is associated with significantly worse survival. These findings support our hypothesis that DNA methylation of ECM pathway genes may have an important role during cancer progression and metastasis.

We also looked at the prognostic significance of these ECM genes (*IGFBP3*, *EVL*, *CD109*, *NRCAM*, and *FLNC)* in a secondary validation cohort NLCS (13) of 558 patients with stage 1 to 4 colon cancers (Table 1). Of note, the NLCS cohort differs significantly from the JHU training and validation in being a historical cohort (study period 1986–1991) with marked differences in survival of colon cancers, a preponderance of distal colon cancers and markedly worse outcomes (see Table 1). *EVL* (HR: 1.41, 95% CI: 1.05–1.89, $P = 0.022$) again is statistically significant as a marker associated with worse survival and *CD109* again shows a trend toward statistical significance (HR: 1.19; 95% CI: 0.95–1.48, P = NS), whereas *IGFBP3* methylation was not significant in this historical cohort.

ECM biomarkers in stage 2 colon cancers

In the clinical management of colon cancer patients, there is an urgent need for identification of biomarkers that may stratify the high risk stage 2 colon cancers patients who are at the highest risk of recurrence. Previous large randomized trials have not shown a consistent benefit of adjuvant chemotherapy (28) but approximately 30% to 40% of these patients will nevertheless recur/metastasize (29). Identification of this high-risk cohort of stage 2 colon cancer patients would be useful as this subset may derive a benefit from adjuvant chemotherapy. Therefore, we investigated the biomarker potential of the ECM gene methylation in the stage 2 colon cancers in the combined JHU training and validation cohort ($n = 219$; Supplementary Table S6). Both *IGFBP3* and *CD109* are associated with worse survival in these stage 2 colon cancers (*IGFBP3* HR: 3.02; 95% CI: 1.42–6.4;

Figure 2. Widespread ECM pathway silencing by DNA hypermethylation in colon cancer. A, hypermethylated genes in colon cancer (green circled) are highlighted on a map of the ECM pathway. The network shows interaction of the genes with each other within the ECM. Circled (green) genes include 13 hypermethylated genes in CRC derived from our gene discovery approach (*LAMA1*, *MMP2*, *FBN2*, *CPAMD8*, *NTNG1*, *NRCAM*, *EVL* and *GPNMB*, *FLNC*, *CD109*, *ICAM5*, *HAPLN1*, and *IGFBP3*). Arrows indicate incoming (←) or outcoming (→) connector between proteins. Colors (green/red/black) of line describe positive, negative, and unspecified effect of functional interaction (www.genego.com). Additional 6 genes including *TIMP2*, *TIMP3*, *MMP9*, *Nidogen*, *Osteonectin*, and *PLAU* were previously identified to be downregulated in association with DNA methylation in other cancer types, and now are shown to be similarly altered in colon cancers in the present study. The functional gene ontology analysis was performed using the MetaCore database (GeneGo). The table below the schematic figure shows the putative role of all of these 19 genes in the ECM pathway.

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

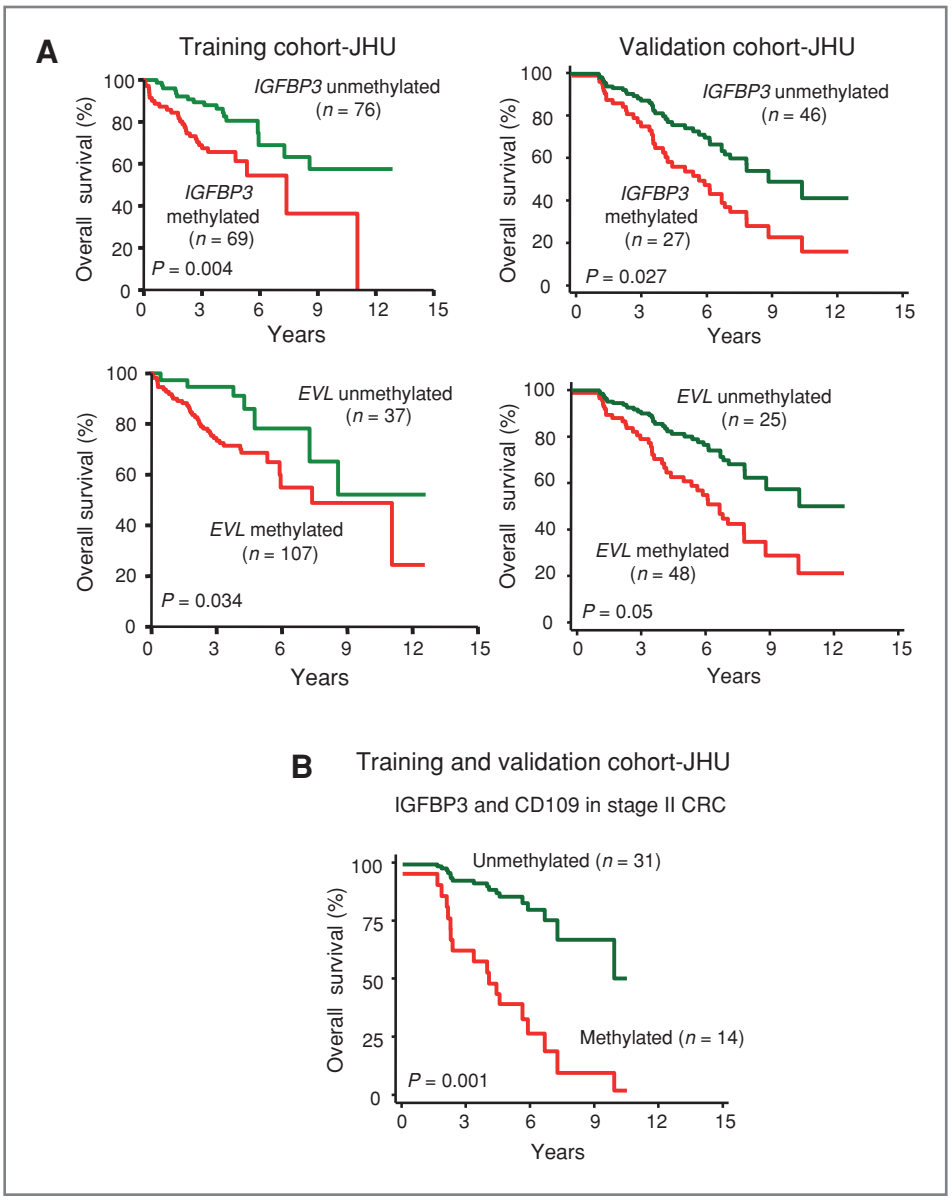


Figure 3. Kaplan–Meier estimates of overall survival according to *IGFBP3* and *EVL* methylation status (adjusted for age and stage). *IGFBP3* and *EVL* methylation was tested in (A) 147 colon cancer patients (stages 1–4; JHU training cohort) and (B) 72 CRC patients (stages 1–4; JHU validation cohort)

$P = 0.004$ and *CD109* HR: 2.41; 95% CI: 1.14–5.10; $P = 0.021$, respectively). Similarly, simultaneous DNA methylation of multiple genes in the ECM pathway is again associated with significantly worse survival for the stage 2 colon cancers (HR: 5.85, 95% CI: 2.03–16.83, $P = 0.001$ for DNA methylation of both *IGFBP3* and *CD109*; Fig. 3B).

Discussion

Our data highlight the important role that integrative approaches may have in understanding cancer biology and implications of such in the translational arena. Our current approach of analyzing the genomic and epigenomic changes in colon cancer identified the role of the ECM pathway in colorectal carcinogenesis. These findings also are potentially useful in finding novel prognostic biomarkers as highlighted by our finding of DNA methylation of *IGFBP3*, *EVL*, and *CD109* being associated with worse survival. As mentioned in the results, this may have direct clinical implications in stratifying the high-risk stage 2 colon cancers patients who have worse outcomes and who may benefit from adjuvant chemotherapy, although our findings would need to be validated in other retrospective studies of carefully clinically annotated tissues or a well-designed prospective study. Our findings were validated in an internal validation cohort (JHU validation) where DNA methylation of both *IGFBP3* and *EVL* again showed the association with worse survival, whereas *CD109* showed a trend toward worse survival. We then validated our findings in another large external NLCS for DNA methylation of

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

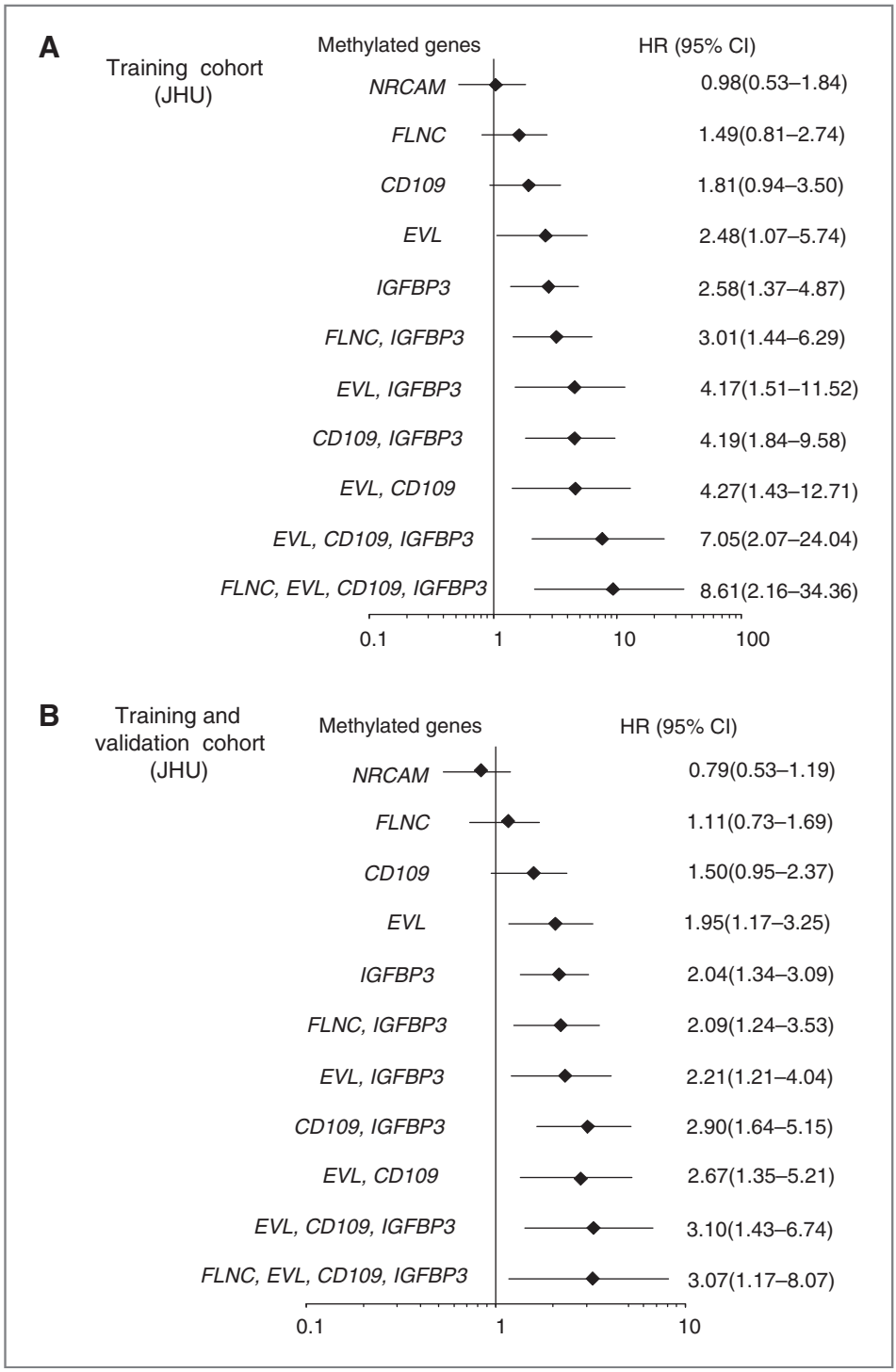


Figure 4. Odds ratios for ECM genes in CRC patients: (A) Forest plot depicting multivariate HRs and corresponding 95% CIs for overall mortality risk associated with DNA methylation of individual, or combinations of genes ($n = 147$ colon cancer samples; JHU training cohort) and (B) combined JHU training and validation cohorts ($n = 219$ colon cancer samples). Selected statistically significant gene combinations are shown. Multivariate Cox regression analysis was performed for 6 ECM genes (*NRCAM*, *FLNC*, *CD109*, *EVL*, and *IGFBP3*) which showed either significant or trends toward significance for HRs on bivariate analyses. The prognostic value of each gene was adjusted for stage and age and then graphed on the Forest plot.

EVL and *CD109* associated with worse outcomes. However, *IGFBP3* methylation was not associated with worse survival in this cohort. Although at present, it is unclear why *IGFBP3* methylation was not validated in the NLCS cohort, there are multiple possible explanations for this. This may simply reflect the different biology in the two cohorts. In fact, epigenetic alterations have been shown to vary in different populations and races and potentially affect outcomes (30, 31). Netherlands also has a lack of colon cancer screening program and this is reflected in the markedly different 5-year survival of this cohort (54.8% NLCS vs. 67.4% JHU training cohort or 65.3% JHU validation cohort) and the biases associated with screening (32). Moreover, the NLCS cohort has a preponderance of left-sided cancers, whereas

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

the JHU cohorts reflect a distribution and survival similar to what is seen in the Surveillance, Epidemiology, and End Results database (a large national cancer database in the United States; ref. 33). The location of colon cancer has been shown to affect both genetic and epigenetic alterations (34–36) and survival (33, 37). In addition, adjuvant chemotherapy was not offered routinely for high risk patients in the NLCS cohort. Thus lack of validation of *IGFBP3* may also reflect the fact that this is a predictive rather than a prognostic marker. Further validation studies in comparable modern-day cohorts similar to the JHU cohort may clarify this. In particular, our data on *IGFBP3* and survival is intriguing as previous studies have shown that *IGFBP3* induces apoptosis and inhibits DNA synthesis in breast, prostate, and non-small–cell lung cancer (38). Moreover, prior studies have also reported that high levels of circulating *IGFBP3* decrease the risk of recurrent colorectal adenoma formation (39). Further studies are needed to confirm these findings particularly the risk stratification in stage 2 colon cancers.

Previous studies have also alluded to the potential of epigenetic changes, especially DNA hypermethylation of promoter-associated CpG islands, in serving as prognostic biomarkers. Brock and colleagues showed that hypermethylation of 2 genes, *p16* and *H-cad*, in the primary lung tumor and the mediastinal lymph nodes could predict recurrence for early stage lung cancers (5). In another study, DNA methylation profiles of multiple methylated-in tumor (MINT) loci predicted distal recurrence in a small cohort of rectal cancer patients (40), while *ID4* methylation has also been associated with worse outcomes (41). However most of these studies used candidate genes from the literature and tested them for prognostic significance in a cohort of patients. In our study, we have not only identified a unique pathway, the ECM pathway, that is altered by DNA methylation of many genes in CRC, but then show that dysregulation of this core pathway has prognostic significance as well. However, our current studies do not clearly distinguish between the DNA hypermethylation arising from the tumorigenic cells or from the non-tumorigenic stromal cells. This will need to be investigated in future studies.

What are the functional consequences of loss of function for the above ECM genes in CRC? We also considered, from several perspectives, our findings in another context—namely, that they might represent an example of properties which reflect the primitive cell status increasingly attributed to multiple tumor types (42). First, multiple previous studies (3) have stressed that some 50% of genes that become DNA hypermethylated in colon cancers are marked by a chromatin pattern, termed bivalent chromatin, that keeps these same genes in a low expression pattern in ESC and embryonic mesenchymal (EM) progenitor cells (43). Indeed, a key chromatin component of bivalent chromatin, the presence of polycomb group (PcG) constituents around gene start sites, is present in ESC and EM cells in the promoter regions of several of the ECM genes (*IGFBP3*, *NRCAM*, *CPAMD8*, *FBN2*, *MMP2*, *TIMP3*, *MMP9*), based on our database analyses (unpublished data).

In summary, our results suggest that CRC may harbor defects in key ECM genes, induced infrequently by gene mutations, and even more frequently by gene promoter DNA hypermethylation. This loss of function for such genes appears to correlate with loss of cellular differentiation and to be associated with aspects of primitive, stem/progenitor cell, features of CRC (Supplementary Fig. S3). Loss of ECM gene function could foster metastatic behavior by altering events which require ECM remodeling such as allowing cells to invade through the basement membrane, migrate into the lymphovascular space, and establish metastatic foci in a distant organ. Our study suggests that this ECM alteration by hypermethylated genes may contribute to carcinogenesis, to some degree in virtually every CRC, through silencing of selected ECM pathway genes by both genetic and epigenetic alterations. Finally, our data indicate that detection of DNA hypermethylation of selected of the genes we have studied may provide a biomarker strategy for predicting the clinical behavior of CRC. Our data emphasize the importance of integrating cancer gene mutations and DNA hypermethylation changes to uncover the molecular events disrupting cell signaling in CRC and other cancers.

Disclosure of Potential Conflicts of Interest

S.B. Baylin and J.G. Herman are consultants of MDx Health and receive research support from MDx Health. The other authors disclosed no potential conflicts of interest.

Acknowledgments

We thank Dr. Steven Goodman for Biostatistical discussion. We also thank Sharon Metzger-Gaud and the Johns Hopkins Cancer Registry for assistance with the primary cancer databases, Iris van Vlodrop for technical assistance, and Kathy Bender for administrative support.

Grant Support

The study was supported by the National Institute of Environmental Health Sciences (grant number: ES011858), the NIH (grant number: K23CA127141), the Jeannik M. Littlefield-AACR Grant in Metastatic Colon Cancer, the Mary Kay Ash Foundation, the Susan G. Komen Foundation, and the American College of Surgeons/Society of University Surgeons Career Development Award.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 17, 2010; revised November 23, 2010; accepted December 15, 2010; published OnlineFirst January 28, 2011.

References

1. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, et al. The consensus coding sequences of human breast and colorectal cancers. Science 2006;314:268–74.

2. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. Science 2007;318:1108–13.

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

3. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007;128: 683–92.
4. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 2003;349: 2042–54.
5. Brock MV, Hooker CM, Ota-Machida E, Han Y, Guo M, Ames S, et al. DNA methylation markers and early recurrence in stage I lung cancer. N Engl J Med 2008;358:1118–28.
6. Kim MS LJ, Sidransky D. DNA methylation markers in colorectal cancer. Cancer Metastasis Rev 2010;29:181–206.
7. Schuebel KE, Chen W, Cope L, Glöckner SC, Suzuki H, Yi JM, et al. Comparing the DNA hypermethylome with gene mutations in human colorectal cancer. PLoS Genet 2007;3:1709–23.
8. Chan TA, Glöckner SC, Yi JM, Chen W, Van Neste L, Cope L, et al. Convergence of mutation and epigenetic alterations identifies common genes in cancer that predict for poor prognosis. PLoS Med 2008;5:823–38.
9. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer 2002;2:442–54.
10. DeClerck YA, Mercurio AM, Stack MS, Chapman HA, Zutter MM, Muschel RJ, et al. Proteases, extracellular matrix, and cancer: a workshop of the path B study section. Am J Pathol 2004;164:1131–9.
11. Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Migrating cancer stem cells—an integrated concept of malignant tumour progression. Nat Rev Cancer 2005;5:744–9.
12. Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Dev Cell 2008;14: 818–29.
13. Van Den Brandt PA, Goldbohm RA, Van 'T Veer P, Volovics A, Hermus RJJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in the Netherlands. J Clin Epidemiol 1990;43:285–95.
14. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22.Proc Natl Acad Sci U S A2002;99: 3740–5.
15. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, et al. An integrated genomic analysis of human glioblastoma multiforme. Science 2008;321:1807–12.
16. Lane RP, Chen X-N, Yamakawa K, Vielmetter J, Korenberg JR, Dreyer WJ. Characterization of a highly conserved human homolog to the chicken neural cell surface protein bravo/Nr-CAM that maps to chromosome band 7q31. Genomics 1996;35:456–65.
17. Beck K, Hunter I, Engel J. Structure and function of laminin: anatomy of a multidomain glycoprotein. FASEB J 1990;4:148–60.
18. Chambers AF, Matrisian LM. Changing views of the role of matrix metalloproteinases in metastasis. J Natl Can Inst 1997;89:1260–70.
19. Finnson KW TB, Liu K, Marcoux A, Lepage P, Roy S, Bizet AA, Philip A. Identification of CD109 as part of the TGF-beta receptor system in human keratinocytes. FASEB J 2006;20:1525–7.
20. Tian L SM, Ning L, Van Lint P, Nyman-Huttunen H, Libert C, Itohara S, et al. Activation of NMDA receptors promotes dendritic spine development through MMP-mediated ICAM-5 cleavage. J Cell Biol 2007;178:687–700.
21. Bachman KE, Herman JG, Corn PG, Merlo A, Costello JF, Cavenee WK, et al. Methylation-associated silencing of the tissue inhibitor of metalloproteinase-3 gene suggest a suppressor role in kidney, brain, and other human cancers. Cancer Res 1999;59:798–802.
22. Galm O, Suzuki H, Akiyama Y, Esteller M, Brock MV, Osieka R, et al. Inactivation of the tissue inhibitor of metalloproteinases-2 gene by promoter hypermethylation in lymphoid malignancies. Oncogene 2005;24:4799–805.
23. Ibanez de Cac eres I, Dulaimi E, Hoffman AM, Al-Saleem T, Uzzo RG, Cairns P. Identification of novel target genes by an epigenetic reactivation screen of renal cancer. Cancer Res 2006;66:5021–8.
24. Ulazzi L SS, Miotto E, Veronese A, Angusti A, Gafà R, Manfredini S, et al. Nidogen 1 and 2 gene promoters are aberrantly methylated in human gastrointestinal cancer. Mol Cancer 2007;6:17.
25. Cheetham S, Tang MJ, Mesak F, Kennecke H, Owen D, Tai IT. SPARC promoter hypermethylation in colorectal cancers can be reversed by 5-Aza-2′-deoxycytidine to increase SPARC expression and improve therapy response. Br J Cancer 2008;98:1810–9.
26. Chicoine E, Esteve P-O, Robledo O, Van Themsche Ci, Potworowski EF, St-Pierre Y. Evidence for the role of promoter methylation in the regulation of MMP-9 gene expression. Biochem Biophys Res Commun 2002;297:765–72.
27. Mani SA, Guo W, Liao M-J, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133:704–15.
28. Andre T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med 2004;350:2343–51.
29. Kuebler JP, Wieand HS, O'Connell MJ, Smith RE, Colangelo LH, Yothers G, et al. Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: results from NSABP C-07. J Clin Oncol 2007;25:2198–204.
30. Mehrotra J, Ganpat MM, Kanaan Y, Fackler MJ, McVeigh M, Lahti-Domenici J, et al. Estrogen receptor/progesterone receptor-negative breast cancers of young African-American women have a higher frequency of methylation of multiple genes than those of Caucasian women. Clin Cancer Res 2004;10:2052–7.
31. Chan AO, Soliman AS, Zhang Q, Rashid A, Bedeir A, Houlihan PS, et al. Differing DNA methylation patterns and gene mutation frequencies in colorectal carcinomas from Middle Eastern countries. Clin Cancer Res 2005;11:8281–7.
32. Pelikan S, Moskowifz M. Effects of lead time, length bias, and false-negative assurance on screening for breast cancer. Cancer 1998;71:1998–2005.
33. Meguid RA SM, Wolfgang CL, Chang DC, Ahuja N. Is there a difference in survival between right- versus left-sided colon cancers? Ann Surg Oncol 2008;15:2388–94.
34. Barry I. Are there two sides to colorectal cancer? Int J Cancer 2002;101:403–8.
35. Annika L. Different mechanisms in the tumorigenesis of proximal and distal colon cancers. Curr Opin Oncol 2001;13:63–9.
36. Sugai T, Habano W, Jiao Y-F, Tsukahara M, Takeda Y, Otsuka K, et al. Analysis of molecular alterations in left- and right-sided colorectal carcinomas reveals distinct pathways of carcinogenesis: proposal for new molecular profile of colorectal carcinomas. J Mol Diagn 2006;8:193–201.
37. Benedix F KR, Meyer F, Schmidt U, Gastinger I, Lippert H; Colon/rectum carcinomas (primary tumor) study group. Comparison of 17,641 patients with right- and left-sided colon cancer: differences in epidemiology, perioperative course, histology, and survival. Dis Colon Rectum 2010;53:57–64.
38. Yamada PM, Lee K-W. Perspectives in mammalian IGFBP-3 biology: local vs. systemic action. Am J Physiol Cell Physiol 2009:296; C954–76.
39. Flood A MV, Pfeiffer R, Kahle L, Rosen CJ, Lanza E, Schatzkin A. Serum concentrations of insulin-like growth factor and insulin-like growth factor binding protein 3 and recurrent colorectal adenomas. Cancer Epidemiol Biomarkers Prev 2008;17: 1493–8.
40. de Maat MFG, van de Velde CJH, van der Werff MPJ, Putter H, Umetani N, Klein-Kranenbarg EM, et al. Quantitative analysis of methylation of genomic loci in early-stage rectal cancer predicts distant recurrence. J Clin Oncol 2008;26:2327–35.
41. Umetani N, Takeuchi H, Fujimoto A, Shinozaki M, Bilchik AJ, Hoon DSB, Epigenetic inactivation of ID4 in colorectal carcinomas correlates with poor differentiation and unfavorable prognosis. Clin Cancer Res 2004;10:7475–83.
42. Ben-Porath I,Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet 2008; 40:499–507.
43. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell 2007;128:669–81.

Published OnlineFirst January 28, 2011; DOI: 10.1158/1078-0432.CCR-10-2509

Clinical Cancer Research

Genomic and Epigenomic Integration Identifies a Prognostic Signature in Colon Cancer

Joo Mi Yi, Mashaal Dhir, Leander Van Neste, et al.

Clin Cancer Res 2011;17:1535-1545. Published OnlineFirst January 28, 2011.