



Epigenetic Homogeneity Within Colorectal Tumors Predicts Shorter Relapse-Free and Overall Survival Times for Patients With Locoregional Cancer

Anna Martínez-Cardús,^{1,*} Sebastian Moran,^{1,*} Eva Musulen,² Cátia Moutinho,¹ Jose L. Manzano,³ Eva Martinez-Balibrea,⁴ Montserrat Tierno,⁴ Elena Élez,⁵ Stefania Landolfi,⁶ Patricia Lorden,¹ Carles Arribas,¹ Fabian Müller,⁷ Christoph Bock,^{7,8} Josep Taberner,⁵ and Manel Esteller^{1,9,10}

¹Cancer Epigenetics and Biology Program, Bellvitge Biomedical Research Institute, L'Hospitalet, Barcelona, Catalonia, Spain; ²Pathology Department, Germans Trias i Pujol University Hospital, Badalona, Catalonia, Spain; ³Medical Oncology Department, Germans Trias i Pujol University Hospital, Badalona, Catalonia, Spain; ⁴Catalan Institute of Oncology, Health Sciences Research Institute of the Germans Trias i Pujol Foundation, Barcelona, Catalonia, Spain; ⁵Medical Oncology Department, Vall d'Hebron University Hospital and Institute of Oncology, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Center affiliated with the Spanish Cancer Research Network (Institute of Health Carlos III), Spain; ⁶Pathology Department, Vall d'Hebron University Hospital, Barcelona, Catalonia, Spain; ⁷Max Planck Institute for Informatics, Saarbrücken, Germany; ⁸Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria; ⁹Department of Laboratory Medicine, Medical University of Vienna, 1090 Vienna, Austria; ¹⁰Department of Physiological Sciences II, School of Medicine, University of Barcelona, Barcelona, Catalonia, Spain; and ¹⁰Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia, Spain

BACKGROUND & AIMS: There are few validated biomarkers that can be used to predict outcomes for patients with colorectal cancer. Part of the challenge is the genetic and molecular heterogeneity of colorectal tumors not only among patients, but also within tumors. We have explored intratumor heterogeneity at the epigenetic level, due to its dynamic nature. We analyzed DNA methylation profiles of the digestive tract surface and the central bulk and invasive front regions of colorectal tumors. **METHODS:** We determined the DNA methylation profiles of >450,000 CpG sites in 3 macrodissected regions of 79 colorectal tumors and 23 associated liver metastases, obtained from 2 hospitals in Spain. We also analyzed samples for KRAS and BRAF mutations, 499,170 single nucleotide polymorphisms, and performed immunohistochemical analyses. **RESULTS:** We observed differences in DNA methylation among the 3 tumor sections; regions of tumor–host interface differed the most from the other tumor sections. Interestingly, tumor samples collected from areas closer to the gastrointestinal transit most frequently shared methylation events with metastases. When we calculated individual coefficients to quantify heterogeneity, we found that epigenetic homogeneity was significantly associated with short time of relapse-free survival (log-rank $P = .037$) and short time of overall survival (log-rank $P = .026$) in patients with locoregional colorectal cancer. **CONCLUSIONS:** In an analysis of 79 colorectal tumors, we found significant heterogeneity in patterns of DNA methylation within each tumor; the level of heterogeneity correlates with times of relapse-free and overall survival.

Keywords: Colorectal Cancer; Epigenetics; DNA Methylation; Heterogeneity.

pathways underlying CRC, few prognostic biomarkers have been established to estimate survival.² This is partly due to the heterogeneous nature of the disease. At the intertumoral level, different genetic³ and epigenetic⁴ alterations could define several subtypes of colorectal primary tumors. However, the lack of knowledge about the molecular defects that drive metastatic disease,⁵ which causes of 90% of cancer morbidity, represents a major drawback in our ability to define accurate prognostic markers. Interestingly, it is widely accepted that most tumors operate through clonal sweeps and branched evolution, produced from sequential selection of acquired phenotypic traits,⁶ and it is also known that intratumoral genetic diversity takes place.^{7,8} But there are important constraints to intratumoral genetic diversity that are less evident with epigenetic changes, which are plastic in nature.⁹ For this reason, shifts in the DNA methylation patterns have been involved in metastasis formation¹⁰ or chemoresistance.¹¹ We studied the extent of intratumoral epigenetic heterogeneity in CRC and whether it has any impact in patient prognosis.

*Authors share co-first authorship.

Abbreviations used in this paper: CB, central bulk; CRC, colorectal cancer; EH, epigenetic homogeneity; FFPE, formalin-fixed paraffin-embedded; HR, hazard ratio; IF, invasive front; MSI, microsatellite instability; SNP, single nucleotide polymorphism; TS, tract surface.

Most current article

© 2016 by the AGA Institute. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

0016-5085

<http://dx.doi.org/10.1053/j.gastro.2016.08.001>

Colorectal cancer (CRC) is the second most common cancer cause of death in developed countries.¹ Despite increased understanding of the molecular

Material and Methods

Clinical, Molecular, and Histopathologic Characteristics of Colorectal Cancer Patients

This was a retrospective study of 79 CRC patients. Formalin-fixed paraffin-embedded (FFPE) tissue was obtained by surgical resection in 2 different Spanish hospitals (Hospital Germans Trias i Pujol, Catalan Institute of Oncology, Badalona and Hospital Vall d'Hebron, Barcelona). Signed informed consent was obtained from each patient. Twenty-three of the 79 primary tumors had tissue available from paired liver metastases: 10 synchronous and 13 metachronous. Clinical and histopathologic characteristics of the patients are shown in Table 1. Briefly, 80% of patients presented local disease (stage I, II, or III) at diagnosis, while the remainder had metastases. Microsatellite instability (MSI) was present in 14% (11 of 79) of tumors, all of which were from patients with local disease. Complete clinical data necessary for survival analysis were available in 71 of 79 patients, 11% of which had MSI (8 of 71) and 18% of which (13 of 71) had advanced disease. Patients who received adjuvant treatment were treated with fluoropyrimidines and oxaliplatin based-therapy. An additional set of stage IV CRC tumors (n = 17) was also included in the study. Histopathologic parameters were assessed in the Pathology

Unit of Hospital Germans Trias i Pujol (Badalona) using H&E staining and optic microscope.

Delimitation of Primary Intratumoral Regions

All 79 FFPE primary tumors underwent macrodissection (Figure 1A). Three different tumor regions were assessed at the Pathology Unit of Hospital Germans Trias i Pujol (Badalona), using H&E staining and optic microscope. All cancer specimens contained at least 80% of carcinoma cells for each intratumoral section. Stromal and inflammatory cells were also counted for each region. Three regions were defined: the region nearest the digestive tract surface (TS), the central bulk (CB), and the invasive front (IF).

DNA Extraction and Quality Control

After deparaffinization by Xylene and tissue digestion by proteinase K (Qiagen, Venlo, The Netherlands), DNA was extracted using E.Z.N.A FFPE DNA extraction kit (Omega Bio-Tek Inc, Norcross, GA) following the manufacturer's instructions. DNA was quantified by fluorometric method (Quant-iT PicoGreen dsDNA Assay; Life Technologies, Carlsbad, CA) and checked for its suitability for FFPE restoration, as indicated on Infinium HD FFPE QC Assay (Illumina Inc., San Diego, CA), by performing a quantitative polymerase chain reaction with 2 ng FFPE DNA. ΔC_q value was calculated by subtracting the mean C_q value of the interrogated sample to the C_q value of a standard provided by the manufacturer. All samples resulted in $\Delta C_q < 5$, which is the threshold recommendation for suitability of samples for FFPE restoration.

Table 1. Clinical and Histopathologic Characteristics of Colorectal Cancer Cohort (n = 79)

Characteristic	n (%)	Parameter	n (%)
Sex		Survival	
Male	62 (78)	Relapse	
Female	17 (22)	Yes	26 (33)
Age		No	45 (57)
<60 y	16 (20)	Unknown	8 (10)
>60 y	48 (61)	Survival status	
Unknown	15 (19)	Alive	58 (73)
Stage		Death	13 (17)
Locoregional disease	63 (80)	Unknown	8 (10)
Advanced disease	16 (20)	Histopathologic	
Localization		Histology subtype	
Right colon	31 (39)	ADC	63 (80)
Left colon	48 (60)	ADC mucinous or PD	14 (18)
MSI status		Unknown/NA	2 (2)
MSS	68 (86)	Type of IF	
MSI+	11 (14)	Expansive	39 (49)
KRAS status		Infiltrative	36 (46)
Wild-type	56 (71)	Unknown/NA	4 (5)
Mutated	23 (29)	Tumor budding	
BRAF status		No/low	54 (69)
Wild-type	68 (86)	Medium/high	20 (25)
Mutated	11 (14)	Unknown/NA	5 (6)
Liver metastasis		Tumor inflammation	
Metachronous	13 (16)	No/low	45 (57)
Synchronous	10 (13)	Medium/high	32 (41)
No metastasis/NA	56 (71)	Unknown/NA	2 (2)
Adjuvant treatment		Tumor necrosis	
Yes	58 (74)	No/low	60 (76)
No	13 (16)	Medium/high	14 (18)
Unknown	8 (10)	Unknown/NA	5 (6)

ADC, adenocarcinoma; MSS, stable tumors; NA, not available; PD, poorly differentiated.

DNA Methylation Analysis

DNA bisulfite conversion was carried out using 300 ng DNA randomized on a 96-well plate using EZ-96 DNA Methylation kit (Zymo Research Corp, Irvine, CA) following the manufacturer's recommendations for Infinium assays. Bisulfite-converted DNA was processed with the Infinium FFPE restoration process and then hybridized on Infinium Human-Methylation450 BeadChip array following the manufacturer's instructions for automated processing of the arrays with a liquid handler (Illumina Infinium HD Methylation Assay Experienced User Card, Automated Protocol 15019521 v01). We normalized raw-intensity data using manufacturer's method (GenomeStudio) implemented on minfi package (pre-processillumina; minfi, version 1.12), as this method relies on the fact that fixed amounts of the different control oligonucleotides have been added to the different reagents used during array sample processing. As the amounts of each control oligonucleotides are the same for all samples and beadchips used, normalization method makes the necessary adjustments so control intensities are comparable among beadchips and samples within it. Normalized intensities were transformed into β values by computing per marker the quotient between methylated intensity divided by the sum of methylated and unmethylated intensities and a offset of 100 (to avoid dividing with small values). The obtained values were uploaded to GEO repository under accession numbers GSE69550, GSE81359, GSE81362, and GSE81472. After the normalization step, probes related to X and Y chromosomes were removed, as well as those probes with 10 bases nearer the interrogated site that contained a single nucleotide polymorphism (SNP), as

annotated on the product description file (Human-Methylation450, version 1.2 manifest file). In addition, CpG sites with associated P values from channel detection $>.01$ were discarded for the analysis. To assert that the 3 different tumoral types per patient were being used correctly, probes designed within the microarray to target 65 SNPs were used to ensure proper comparison of patient samples. Technical replicates were obtained for 12 intratumoral samples.

Unsupervised heatmap representation of the 3 different tumoral regions (TS, CB, and IF) for the whole cohort of patients was generated using randomly selected 1% of the CpG sites and organizing them by applying a hierarchical clustering method on Manhattan distances aggregated by Ward's linkage. Similarly, per patient, hierarchical clustering of Manhattan distances aggregated by ward's linkage over the top 1% of CpG sites, with the greatest differences on methylation values among intratumoral regions were considered. The robustness of the identified groups was determined using multiscale bootstrap resampling ($n = 1000$).

Coefficient of Epigenetic Homogeneity

In order to estimate the level of intratumoral epigenetic homogeneity, we calculated a coefficient based on the epigenetic changes among the 3 regions. SD mean promoters coefficient was used to compute the SD across the 3 intratumoral regions (TS, CB, and IF) for each patient and each marker that passed our filtering procedures, which consisted of the following exclusion criteria: probes outside promoters (according to promoter regions of Ensembl genes, version Ensembl Genes 75), probes located in sex chromosomes, probes with detection P values $>.01$, and probes interrogating non-CpG sites. The score for a patient was computed according to the equation:

$$SD \text{ mean promoters} = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{3-1} \sum_{j=1}^3 (\beta_{j,i} - \bar{\beta}_i)^2}$$

where N is total number of filtered CpG sites interrogated in the microarray, i is one of the filtered CpG sites, j is 1 of the 3 intratumoral regions (TS, CB, and IF), and β is the methylation value. A similar approach has been used to quantify intratumoral genetic heterogeneity.¹² The lowest value indicates the highest level of homogeneity, meaning a lower heterogeneity.

Immunohistochemistry

Immunohistochemistry was performed using EnVision + Dual Link System-HRP, DAB+ (Dako, Carpinteria, CA) according to manufacturer's instructions, using the PRDM16 (ab118573; Abcam, Cambridge, MA, dilution 1:1000; epidermis as positive control tissue) and OPCML (ab100923, Abcam, dilution 1:200; fetal brain as external positive control) antibodies. Samples were additionally counterstained with hematoxylin.

Mutational Analysis

KRAS and BRAF mutational status was determined through standard Sanger sequencing using specific primers for genomic DNA (available upon request). The mutational analysis included codons 12 and 13 for KRAS and the V600E mutation for BRAF for the 3 different intratumoral regions for each tumor.

Single Nucleotide Polymorphism Microarray

SNP genotypes were determined using the Illumina Infinium OncoArray-500K. FFPE genomic material (200 ng) was restored using Illumina Restoration Kit. Restored DNA was treated following manufacturer's protocol for Infinium HTS assay. Genotypes, SNPs GeneTrain scores, as well as genotyping call scores, were extracted from GenomeStudio (V2011.1; Genotyping module V1.9.4) using manufacturers cluster file specific for OncoArray-500K. Obtained data were processed under R statistical environment (version 3.1.1), filtering out unreliable genotypes (GeneTrain score ≤ 0.6 ; genotype call score ≤ 0.6). Patient genotype concordance rates were computed by counting the number of markers matching in all 3 intratumoral regions.

Calculation of Z-Scores

Using the 20 normal colon samples included in our study, mean ($\mu_i^{[N]}$) and SD ($\sigma_i^{[N]}$) per probe (i) were computed. The z-score of each probe and intratumoral region (j) was computed according to the following formula:

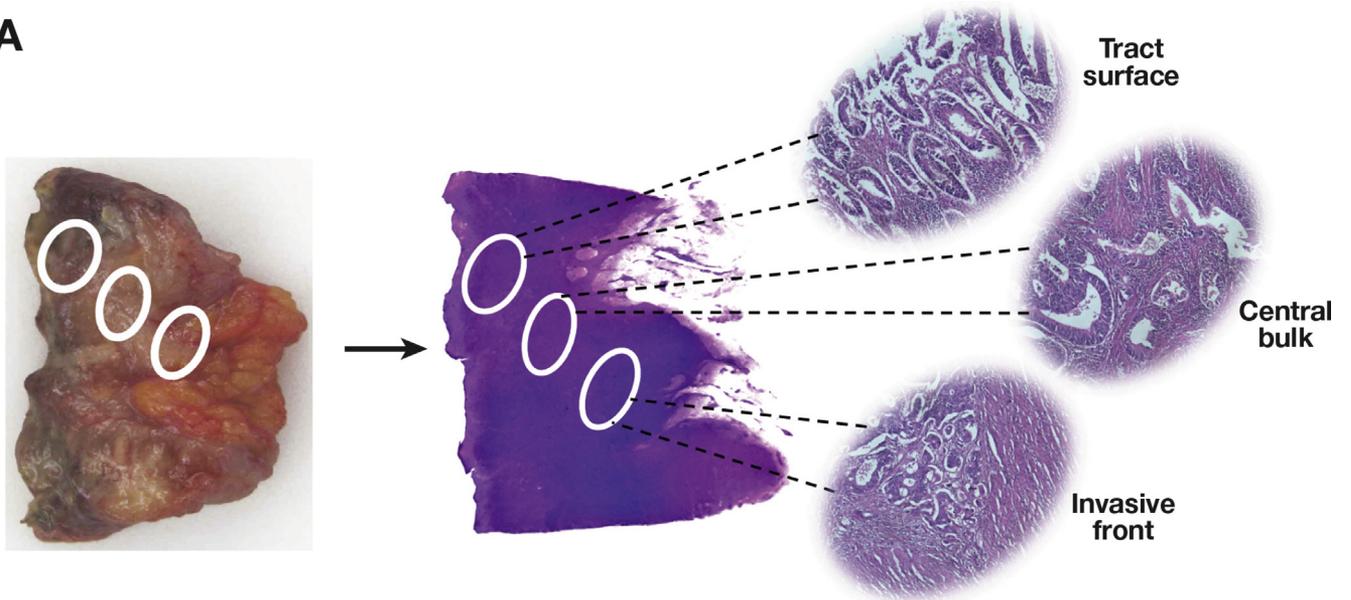
$$Z_{ij} = \frac{|\beta_{ij} - \mu_i^{[N]}|}{\sigma_i^{[N]} - \xi}$$

where a regularization parameter, $\xi = .02$ was chosen after considering the distribution of probe SDs in normal samples. This regularization parameter was introduced in the z-score formula to prevent very large values from occurring. Per intratumoral region, the mean of all z-scores of probes located on promoters were calculated and used to analyze whether an intratumoral region showed more or less differences with normal tissue than others.

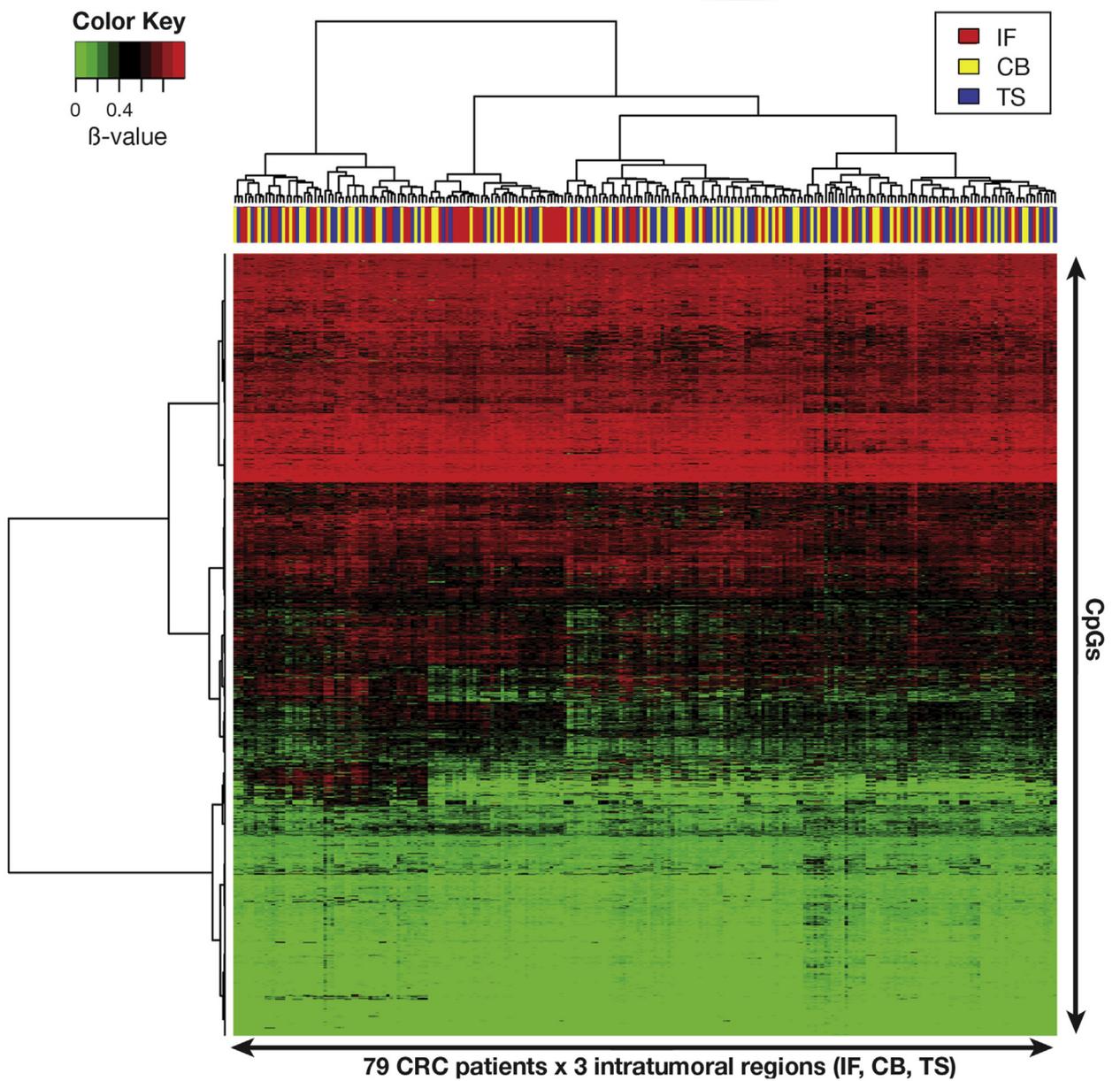
Statistical Analyses

Significant differences in percentages distribution were assessed using the χ^2 goodness-of-fit test. The χ^2 goodness-of-fit test is applied when you have one categorical variable from a single population. The test determines whether a distribution of data or scores for one nominal (categorical) variable matches theoretical expectations for that distribution. In our case, we used the test to evaluate which percentages, calculated for all cohorts according to which was the most divergent intratumoral region, were statistically significant with respect to theoretical expectations for that distribution. In order to evaluate the impact of the level of intratumoral epigenetic homogeneity (EH) in CRC prognosis, we performed survival analysis based on Kaplan-Meier plots, estimating the effect of EH level in relapse-free survival and overall survival. Epigenetic homogeneity level was calculated using the SD mean promoters coefficient described. Two groups were established (high and low homogeneity level) using the quartile 50 (median) of SD mean promoters coefficient as threshold value. Kaplan-Meier plots were drawn and compared with the 2-sided log-rank test. To study the role of EH level in survival considering the effect of other clinical, histopathologic and molecular covariables, hazard ratios (HRs) and their 95% confidence intervals were estimated using multivariate Cox proportional hazards regression models, considering as independent prognostic factor covariables with $P < .05$. Results were further validated by

A



B



BASIC AND
TRANSITIONAL AT

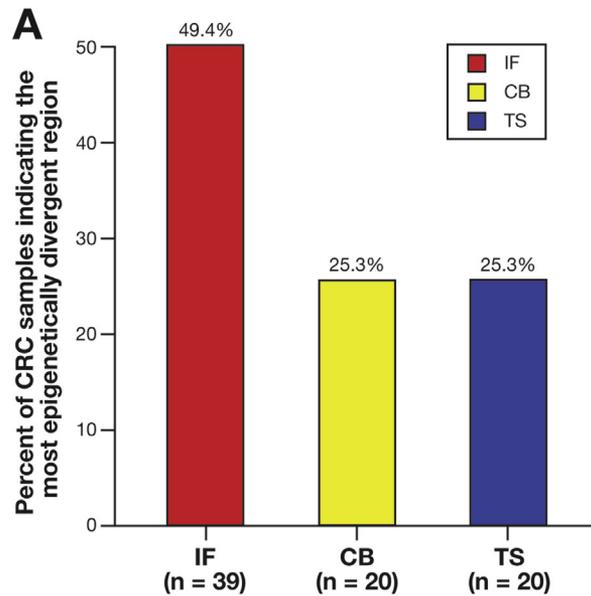
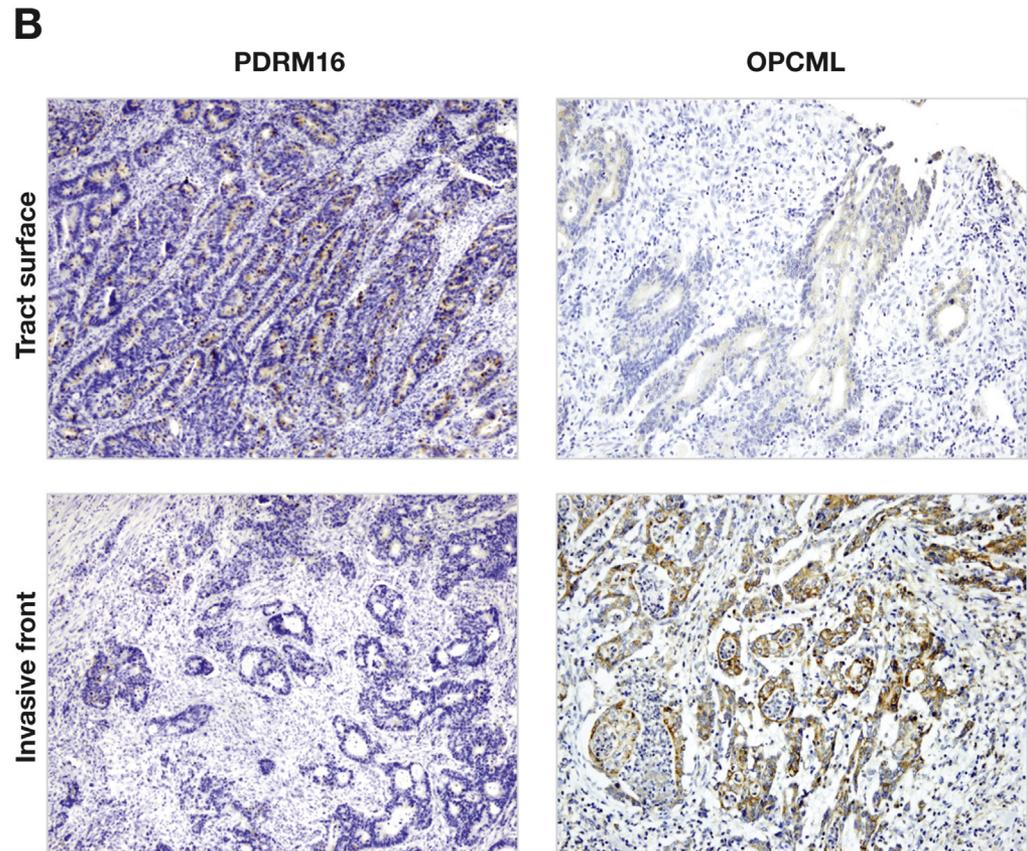


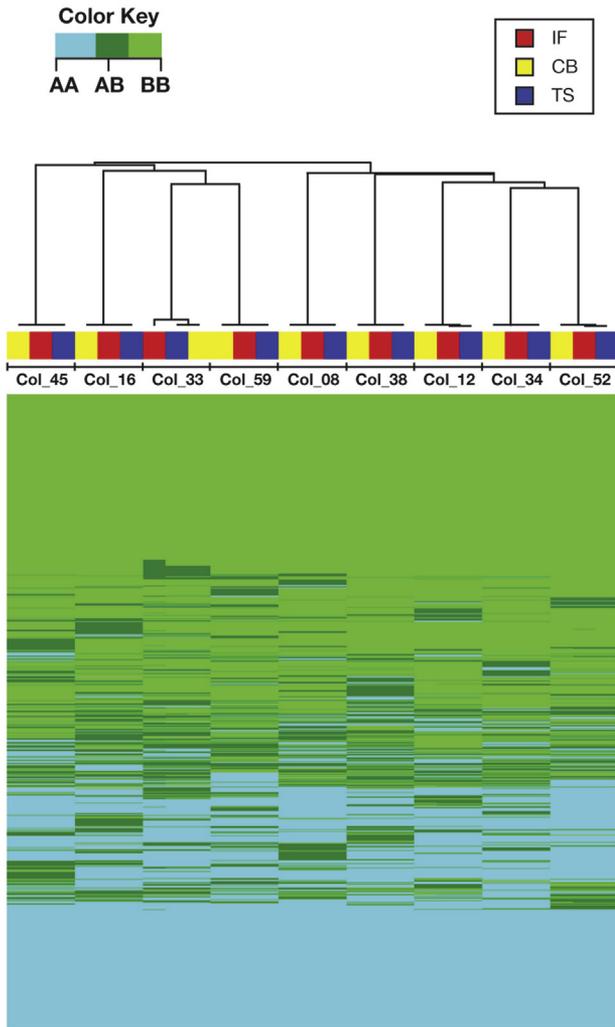
Figure 2. Epigenetically divergent regions and immunohistochemistry validation. (A) Percentage of patients classified according to which is the most DNA methylation divergent intratumoral region when patients were assessed individually: TS; CB; and IF. Divergence was calculated by SD among intratumoral regions, using supervised clustering analysis, comparing all CpGs in promoter sequences. (B) Immunohistochemistry staining in 2 independent colorectal tumors to show the correlation between methylation levels and protein expression of 2 candidate genes. Protein expression (*brown staining*) of PRDM16 is observed in the TS that was found unmethylated for the PRDM16 promoter, whereas the invasive front was hypermethylated and PRDM16 expression was not detected. For the OPCML gene, the opposite pattern is observed in the second tumor: expression is found in the unmethylated IF (*brown staining*), whereas loss of staining is found in the hypermethylated TS.



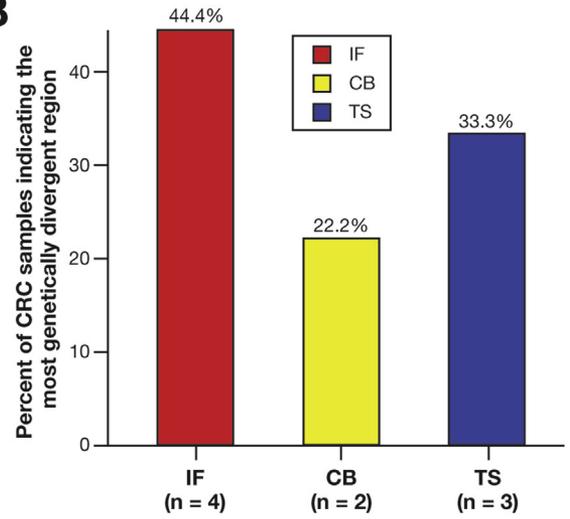
BASIC AND TRANSLATIONAL AT

Figure 1. Analysis of intratumoral DNA methylation heterogeneity. (A) Example of a tumor macrodissection of FFPE tissue from CRC patients. Different intratumoral regions were assessed using H&E staining and optic microscope and were described as region nearest digestive TS, CB, and IF. (B) Unsupervised clustering analysis of whole CRC primary intratumoral regions (n = 237, corresponding to the 3 intratumoral regions from 79 patients) using hierarchical ward agglomeration method for the Manhattan distances of the interrogated CpGs (a 1% randomly selected subset of CpGs from the DNA methylation microarray is shown). Color-coded regions as blue (TS), yellow (CB), and red (IF).

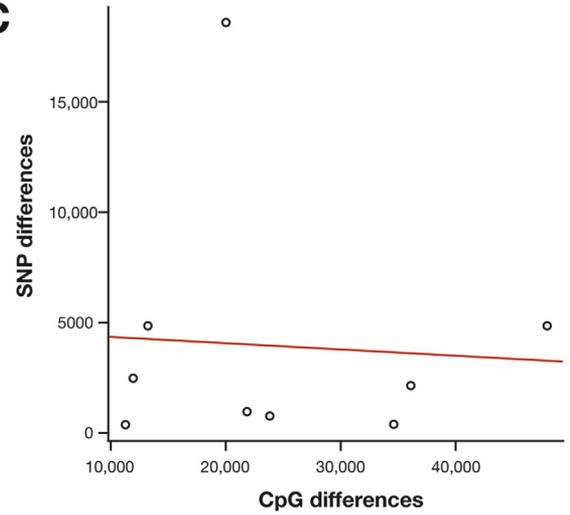
A



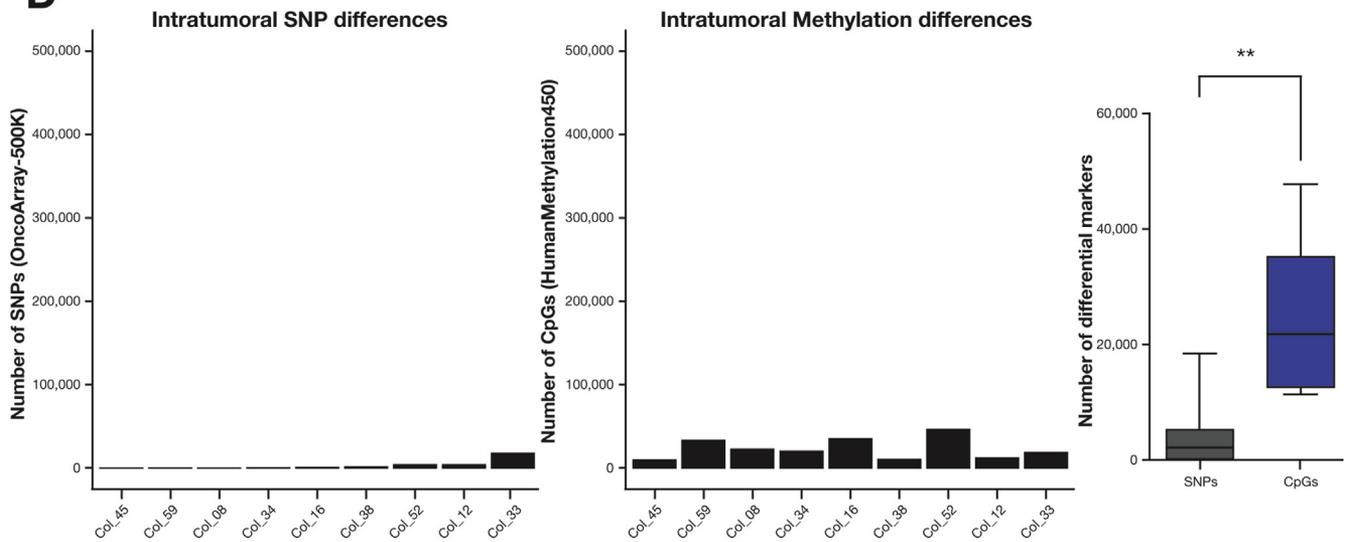
B



C



D



performing a bootstrap resampling procedure among 1000 new models. Kaplan-Meier plots and log-rank test were also applied for the determination of the effect of which was most divergent intratumoral region (IF, CB, or TS) in survival. The associations between the categorical covariables and low or high epigenetic homogeneity level were assessed by using the χ^2 or Fisher's exact test, as required, considering a statistical significant association when $P < .05$. Differences between categorical variables and EH level were assessed using Mann-Whitney U nonparametric test in cases where SD mean promoters were applied as a continuous (quantitative) variable. Pearson's test was calculated to establish the correlation between continuous variables (number of differences in genotype vs in methylation), considering $r > 0.5$ and $P < .05$ as positive correlation.

Results

Intertumoral Epigenetic Heterogeneity Is Higher Than Intratumoral Diversity

We analyzed the DNA methylation profiles of >450,000 CpG sites¹³ of 79 colorectal tumors (Table 1) in 3 macrodissected regions for each sample. These 3 intratumoral sections corresponded to a region more exposed to the gastrointestinal transit, the digestive TS; a region corresponding to the middle of the surgical tumoral piece, the CB, and a region close to the tumor–host interface, the IF (Figure 1A). We observed first in the unsupervised analysis that each of the 3 tumor sections preferentially clustered within each CRC patient (65.8%; 52 of 79) (multiscale bootstrap resampling, $n = 1000$, $P < .05$), thus, intertumoral epigenetic heterogeneity was significantly higher than intratumoral DNA methylation diversity (χ^2 goodness-of-fit test: $P = .005$) (Figure 1B). Technical replicates of a subset of the experimental tissue samples (12 intratumoral regions from 4 CRC patients) showed high reproducibility (Pearson's coefficient $r = 0.98$ [0.97–0.99]; $P = 2.22 \times 10^{-16}$) (Supplementary Figure 1). If we only require that at least 2 of the 3 intratumoral regions cluster together, the percentage reached 100% of cases (Figure 1B). Principal component analysis also showed that the same intratumoral section from different patients did not cluster together (Supplementary Figure 2). In this regard, colorectal tumors undergo promoter CpG island methylation of tumor

suppressor genes that can differ between patients.^{14,15} For example, MSI was present in 14% (11 of 79) of tumors, and these cases showed an enrichment in a specific DNA methylation pattern, which has been previously named as CpG island methylator phenotype,¹⁶ which distinguished them from the microsatellite stable tumors (86%, 68 of 79) (Supplementary Figure 3). Examples of hypermethylated CpG island methylator phenotype–associated genes in our samples included CRABP1 and NEUROG1.¹⁷ These data confirm that there is more intertumoral than intratumoral heterogeneity present, as described previously.¹⁸ Unsupervised clustering analysis of 20 paired normal colorectal mucosa samples showed that most of them cluster together, not with any intratumoral section of the corresponding patient (Supplementary Figure 4). To further prove that we were measuring tumor-specific methylation differences and not underlying normal tissue epigenetic variation for these patients, we used the 20 available normal matched tissues to calculate a probe z-score statistic, as described previously.¹⁹ None of the different intratumoral regions showed statistically significant differences according to this factor (pairwise t -test $P > .05$) (Supplementary Figure 5).

Invasive Front Is the Most Epigenetically Divergent Region

We then proceeded to develop individualized unsupervised clustering analyses for each CRC patient. This approach showed that there was a common DNA methylation profile among the 3 studied locations where 90.9% (range, 70.5%–99.4%) of the CpGs had identical methylation levels (SD of CpG sites calculated per patient within the 3 intratumoral regions < 0.1) in a given case, but 9.1% (range, 0.6%–29.5%) of methylation sites were distinct (SD > 0.1) between the described regions among the interrogated 485,577 CpG dinucleotides. The most epigenetically divergent region was the IF (χ^2 goodness-of-fit test, $P = .01$), a phenomenon occurring in half of the cases (49.4% [39 of 79]), while for the remaining cases, the most different section was either the CB (25.3% [20 of 79]) or the digestive TS (25.3% [20 of 79]) (Figure 2A). This observation fits with the idea that the IF represents a critical interface for molecular changes in colorectal carcinoma that includes changes in adhesion molecules and

Figure 3. Analysis of genetic vs epigenetic intratumoral heterogeneity. (A) Unsupervised clustering analysis of genotyping differences among primary intratumoral regions (TS, CB, and IF) in a set of CRC cases ($n = 9$), using hierarchical ward agglomeration method for the Manhattan distances of the SNPs interrogated with Illumina Infinium OncoArray-500K. Obtained data were filtered out of unreliable genotypes (GeneTrain score ≤ 0.6 ; genotype call score ≤ 0.6). Color coded genotypes were *light blue* (AA = 0), *dark green* (AB = 1) and *light green* (BB = 2). Color coded intratumoral regions were *red* (IF), *yellow* (CB) and *dark blue* (TS). (B) Bar plot representation of which is the most frequently genetic divergent intratumoral region in a set of patients assessed individually ($n = 9$): TS; CB; and IF. Patient genotype discordances were computed by counting the number of markers for which at least one intratumoral region was discordant with the remaining ones. (C) Scatter plot for CRC patients, using the number of differentially methylated CpGs and the number of SNPs showing genotype differences among intratumoral regions (Pearson's correlation $r = -0.062$; $P = .56$). (D) Graphical representation of individual genotype (*left*) and epigenetic (*middle*) differences in a set of samples ($n = 9$), comparing for each patient the 3 intratumoral regions. Patient genotype discordances were computed counting the number of markers for which at least one intratumoral region was discordant with the remaining ones. Methylation differences were considered when SD among 2 intratumoral regions was < 0.05 , and their difference with the third intratumoral region was $> 10\%$. In *right, box plots* representing median values of differences for CpGs and SNPs in the studied 9 patients (2-sided t -test, $P = .0014$).

epithelial–mesenchymal transition (EMT).²⁰ We have not observed any association between the most and the least divergent regions within each patient (χ^2 goodness-of-fit; $P = .84$). Overall, according to analysis of variance after conducting pairwise comparisons across the 3 intratumoral regions from combined probe data in all patients, we observed 3002 CpGs (Supplementary Table 1) with statistically significant changes (false-discovery rate–adjusted $P < .01$, $\Delta\beta > .10$).

For further validation of our assay, we used immunohistochemistry, which is one of the most accepted approaches to studying molecular heterogeneity in a tissue. Our DNA methylation analyses identified, among other genomic sites, genes with differential DNA methylation status of their promoters, which are associated with loss of expression. We selected 2 genes with distinct DNA methylation patterns among the 3 intratumoral sections and tested their expression by immunohistochemistry. We found that the DNA methylation status in each region was associated with the expected expression patterns for the PRDM16 ($n = 11$) and OPCML ($n = 5$) proteins: hypermethylation was linked to diminished expression and an unmethylated status was associated with high expression. Illustrative examples for the 2 genes are shown in Figure 2B.

From a genetic standpoint, we detected the presence of KRAS or BRAF mutation in 29% and 14% of the studied colorectal tumors, respectively, that were present throughout the 3 intratumoral regions (Supplementary Figure 6). We studied the contribution of the genetic component in the generation of heterogeneity by interrogating 499,170 SNPs in the 3 intratumoral sections. The unsupervised clustering analysis of the SNP microarray in 9 CRC patients from our studied set showed that each of the 3 tumor sections clustered within each CRC patient (100%; 9 of 9) (multiscale bootstrap resampling, $n = 1000$; $P < .05$). Thus, as it also occurred with the epigenomic assessment, intertumoral heterogeneity was higher than intratumoral diversity at the SNP array level (Figure 3A). Interestingly, the most SNP-divergent intratumoral section was the IF (χ^2 goodness-of-fit test; $P = .01$), a phenomenon occurring in close to half of the studied cases (44.4% [4 of 9]) (Figure 3B). This is the same region that we found most divergent at the epigenetic level, however, the SNPs associated with heterogeneity (Supplementary Table 2) did not correlate with the CpG dinucleotides displaying methylation heterogeneity (Supplementary Table 1) as shown in Figure 3C (Pearson's correlation $r = 0.062$; $P = .56$). Most important, the number of CpGs with intratumoral methylation heterogeneity was significantly higher than the number of SNPs that were distinct among the regions (Figure 3D) (2-sided t -test; $P = .0014$), suggesting a relevant and independent contribution of the epigenetic landscape to intratumoral heterogeneity. Although for all intratumoral sections, the tumoral component was $\geq 80\%$ of the content of the analyzed section, we wondered whether the observed epigenetic heterogeneity could come from a different cellular composition. Supplementary Table 3 shows the percentage of the tumoral, stromal, and inflammatory components for each 1 of the 3 studied intratumoral

sections. The assignment of the most epigenetically divergent intratumoral region did not correlate with the most different section, according to the percentage of cellular components (χ^2 test: $P = .56$).

Liver Metastasis and its Epigenetic Resemblance to Intratumoral Sections

We investigated which intratumoral region was most similar to 23 liver metastases from the studied cases. Using unsupervised analysis for each intratumoral section within a given case and including its corresponding metastasis sample, we observed that the most frequent region of the primary CRC that shared more methylation events with the metastasis was the digestive TS (43.5%), followed by the IF (30.4%) and CB (26.1%) (Supplementary Figure 7). Similar results were obtained when we analyzed metachronous ($n = 13$) and synchronous ($n = 10$) metastases separately (Supplementary Figure 7). These results are interesting because the TS constitutes the “oldest” part of the primary tumor,²¹ suggesting that dissemination, in addition to the most advanced stages, might also occur early in the carcinogenesis process. In this regard, the fact that the available metastasis samples can be closely related to the digestive tract region, which is thought to represent ancestral lineages, supports the existence of a proposed parallel progression model.²²

Higher Epigenetic Homogeneity and Poor Outcomes in Locoregional Disease

Finally, we studied whether the level of epigenetic heterogeneity in CRC had any impact in the outcomes in these patients. The effects of heterogeneity in tumor progression have been addressed using nucleotide²³ and copy number data,²⁴ but it remains largely unexplored at the DNA methylation level. For the 71 samples for which we had complete clinical information (relapse and survival status, Table 1), we observed from a qualitative manner that knowing which intratumoral region was the more epigenetically divergent did not exhibit prognostic value (Supplementary Figure 8A). Then, we reanalyzed this issue from a quantitative manner by determining the degree of heterogeneity within each tumor, obtaining a coefficient of homogeneity based on the computation of the SD across the 3 regions in each patient for each CpG methylation site located in a gene promoter region. Thus, the highest coefficient of homogeneity indicated the lower heterogeneity. We did not observe any correlation between the obtained coefficient of epigenetic homogeneity and the differences in the cellular components (tumoral, stromal, and inflammatory) in each intratumoral region (χ^2 test: $P = .74$). In this regard, 47% (7 of 15) of the low epigenetic homogeneity cases presented the same percentage of cellular composition in each one of their intratumoral regions, whereas these last cases constituted 52% (13 of 25) of the tumors in the high epigenetic homogeneity group.

For the whole population with clinical data ($n = 71$), a nonsignificant trend between epigenetic homogeneity and reduced relapse-free survival and overall survival was observed (log-rank $P = .11$ and $P = .08$, respectively)

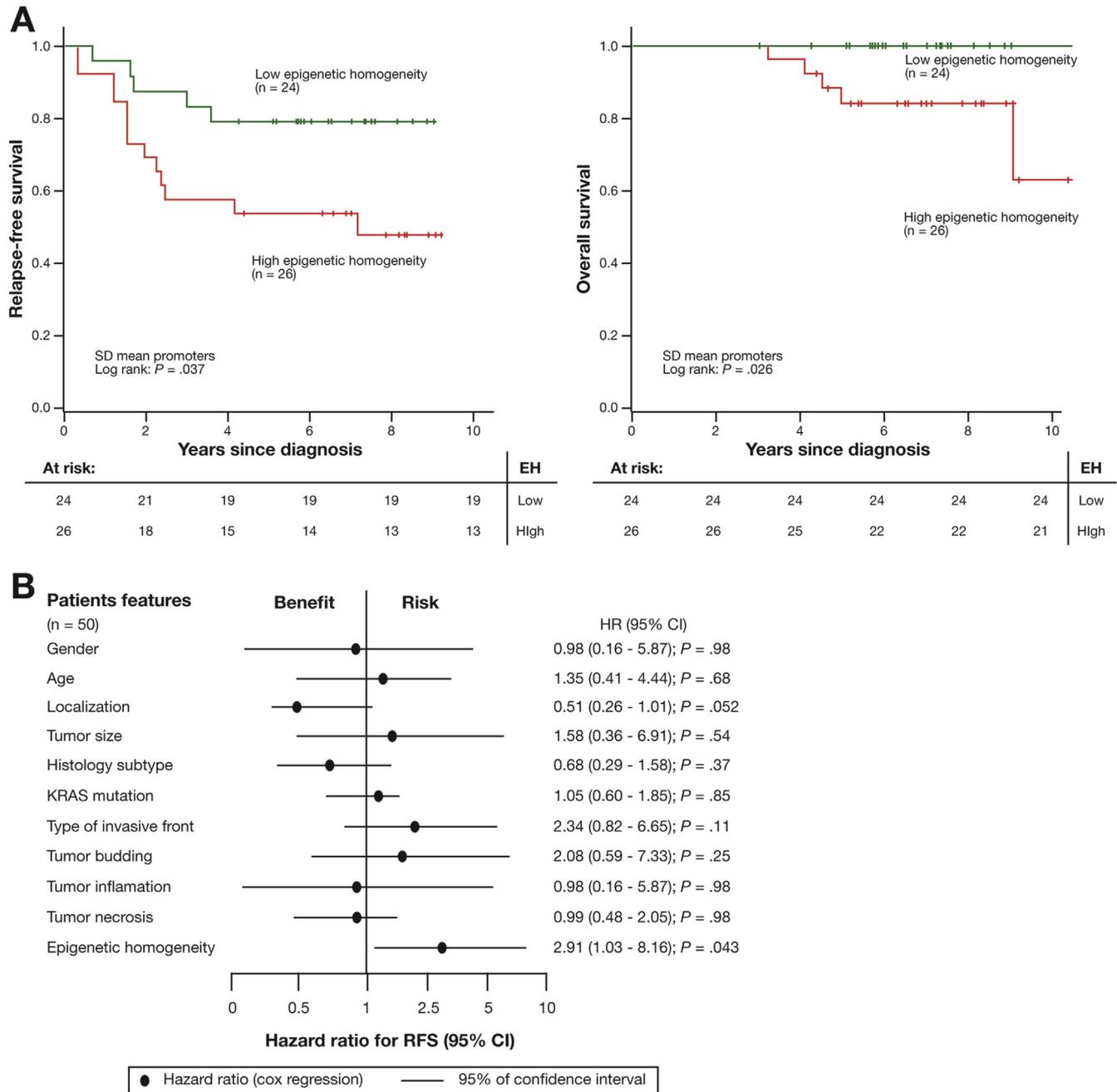


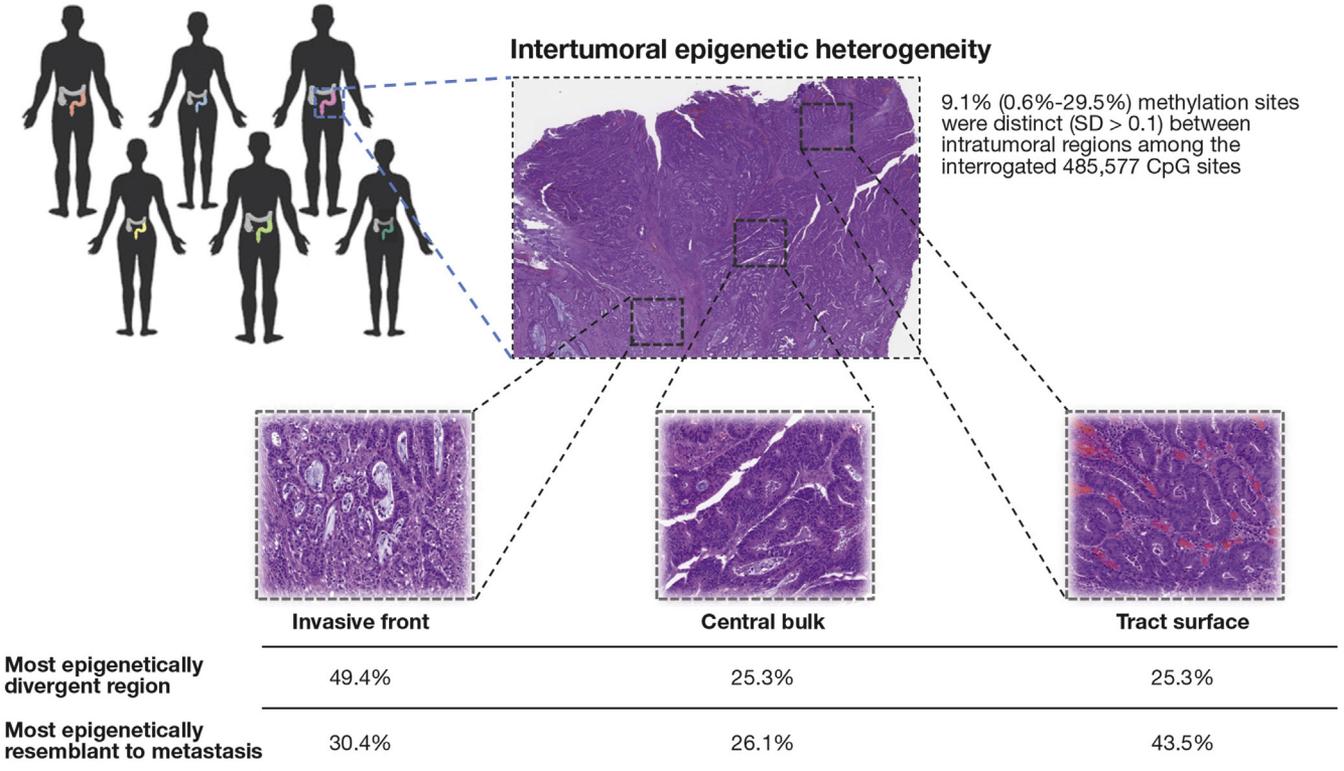
Figure 4. Impact of epigenetic heterogeneity in CRC outcome. (A) Kaplan-Meier analysis for relapse-free survival and overall survival according to tumor EH level. High and low EH groups were established using the coefficient SD mean promoters, considering the median as threshold value between both groups of homogeneity. The *P* value corresponds to the log-rank 2-sided test. Number of events (relapse-free survival [RFS]: relapse; overall survival: death) are shown from 2 to 10 years in high and low epigenetic homogeneity groups. (B) Forest plot representation of Cox proportional hazards regression analysis, taking into account different clinical and histopathologic covariables in CRC cohort of patients with locoregional disease. The *P* values correspond to HR. Concerning overall survival, Cox Regression is not assessed since there were no events (death) in the low epigenetic homogeneity group.

(Supplementary Figure 8B) that was also present in the Cox regression multivariate analysis (*P* = .12 and *P* = .09, respectively) (Supplementary Figure 8C). In this cohort, among all clinical variables collected for each patient, advanced stage was the only independent prognostic factor (HR, 2.41; 95% CI, 1.01–5.81; *P* = .048) (Supplementary Figure 8C). Therefore, in order to evaluate the effect of

epigenetic homogeneity in a more homogeneous CRC population, we stratified the cohort according to this parameter, leaving 58 patients. We also excluded MSI+ cases (n = 8) due to the recognized association of this parameter with good prognosis.^{25,26} Thus, we were able to determine the influence of tumor epigenetic homogeneity in prognosis in a set of 50 patients with locoregional disease (stages I, II, and III)

Intertumoral epigenetic heterogeneity

Significantly higher than intratumoral DNA methylation diversity (Chi-Square goodness of fit test: $P = .005$)



Epigenetic homogeneity

Relapse-free and overall survival times

Figure 5. Summary of our findings in CRC epigenetics highlighting that intertumoral epigenetic heterogeneity is higher than intratumoral diversity; the identification of the invasive front as the most epigenetically divergent region; the epigenetic resemblance of liver metastasis to intratumoral sections; and the impact of epigenetic homogeneity on poor clinical outcomes.

(Supplementary Table 4). We found that CRC patients with locoregional disease and a higher epigenetic homogeneity coefficient showed lower relapse-free survival (log-rank $P = .037$) and reduced overall survival (log-rank $P = .026$) (Figure 4A). According to Cox proportional hazard regression models, higher epigenetic homogeneity was found to be an independent prognostic factor for worse relapse-free survival (HR, 2.91; 95% CI, 1.03–8.16; $P = .043$) (Figure 4B). Our results were further validated by performing a bootstrap resampling procedure. Among 1000 new models, higher epigenetic homogeneity remained an independent prognostic factor for worse relapse-free survival (HR, 3.05; 95% CI, 1.04–8.90; $P = .042$).

We also wondered whether the presence of selective external pressures, such as cytotoxic drugs, could reverse the value of higher epigenetic homogeneity to predict poor outcomes. To address this issue, we included a new set of stage IV CRC tumors for which patients received FOLFIRI (folinic acid + 5-fluorouracil + irinotecan) as first-line treatment ($n = 17$) (Supplementary Table 5) after the

determination of the degree of epigenetic heterogeneity in the surgery specimens with the 3 described intratumoral sections. Coefficient of homogeneity “SDmean promoter” was calculated and patients were divided in low and high epigenetic homogeneity groups, using median as threshold point, as we did. The association between level of homogeneity and clinical covariables was determined using χ^2 tests or Fisher’s exact test whenever required; $P < .05$ was considered as statistically significant (Supplementary Table 5). Interestingly, and in concordance with our previous results in the CRC cohort with locoregional disease, high epigenetic homogeneity was associated with worse response rate to treatment (85.7% of nonresponder cases belonged to high epigenetic homogeneity group; Fisher’s exact test; $P = .049$) and poor survival in these new sets of patients treated with FOLFIRI (75% of exitus cases belonged to high epigenetic homogeneity group, whereas for the low epigenetic homogeneity group, a 100% of the patients remained alive; Fisher’s exact test; $P = .009$) (Supplementary Table 5). Overall, these results could

pinpoint that those tumors with a lower heterogeneity are driven by powerful clonal events that are associated with expansive growth and more aggressive clinical behavior.

Discussion

Intertumoral differences in CRC at the genetic³ and epigenetic⁴ levels occur and have made it possible to establish a consensed molecular classification of potential subtypes of CRC.²⁶ However, the existence of intratumoral molecular diversity within each CRC patient and its potential impact on outcomes, generation of metastasis, and development of chemoresistance, have been explored to a lesser degree. In this regard, genomic sequencing studies that addressed the presence of distinct mutational profiles between different intratumoral regions of primary and metastases indicate that molecular heterogeneity is a common finding.^{27,28} Even distinct metastases from the same patient can show a certain degree of genetic diversity.^{27,28} These different molecular setting can have a relevant role in outcomes and response to targeted therapies.²⁹ For example, the therapeutic use of epidermal growth factor receptor blockade in CRC patients could induce the outgrowth of intratumoral regions carrying KRAS mutations or carrying other molecular events associated with activation of mitogen-associated protein kinases.^{30,31}

The contribution of epigenetic changes, such as DNA methylation aberrations, to intratumoral heterogeneity have been less studied despite the well-recognized contribution to carcinogenesis of these chemical modifications of the genome. For CRC, relevant tumor suppressor genes are inactivated by promoter CpG island hypermethylation³² and some genomic sites acquire hypermethylation frequently, mostly in the MSI+ context, configuring the *CpG island hypermethylation phenotype*.¹⁶ However, despite that reports reflect the existence of intratumoral epigenetic heterogeneity for other malignancies such as prostate³³ and breast³⁴ cancers, we know extremely little about this phenomenon in CRC. In this study, the determination of the DNA methylation landscape of 288 intratumoral sections corresponding to 96 CRC patients contributes to fill the knowledge gap in the field. Our findings are summarized in [Figure 5](#). Our results demonstrates that intertumoral DNA methylation profiles are always more distinct than intratumoral profiles. However, we also show that in CRC primary tumors, intratumoral DNA methylation differences occur and, most of the time, are more extensive than the analyzed genetic diversity. Among the 3 studied intratumoral sections, the invasive front was the most significantly epigenetically divergent. These data agree with the proposed role of the invasive front as a critical region in CRC that undergoes expression changes in adhesion and EMT proteins,²⁰ some of them in a transient manner.³⁵ This fits with the dynamic nature of the epigenetic changes. Our results also show that the corresponding metastasis can resemble any intratumoral region, but were most often similar to the region closest to the gastrointestinal transit, the oldest part of the tumor. This provides a potential explanation for the early occurrence of metastasis, even for

those more localized and small carcinomas. Finally, we also found that those locoregional CRC tumors more homogeneous at the epigenetic level show poor clinical outcomes. The same effect is observed in more advanced cases that received the standard chemotherapy FOLFIRI treatment as first-line therapy. Prospective studies are now necessary to demonstrate whether the use of the newly developed targeted therapies will have an impact on shifting the described epigenetically divergent regions toward ones that harbor an enriched population of resistant cells.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at <http://dx.doi.org/10.1053/j.gastro.2016.08.001>.

References

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. *CA Cancer J Clin* 2015;65:87–108.
2. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet* 2014;383:1490–1502.
3. Linnekamp JF, Wang X, Medema JP, et al. Colorectal cancer heterogeneity and targeted therapy: a case for molecular disease subtypes. *Cancer Res* 2015;75:245–249.
4. Okugawa Y, Grady WM, Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology* 2015;149:1204–1225.e12.
5. Gupta GP, Massagué J. Cancer metastasis: building a framework. *Cell* 2006;127:679–695.
6. Nowell PC. The clonal evolution of tumor cell populations. *Science* 1976;194(4260):23–28.
7. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;366:883–892.
8. Yates LR, Campbell PJ. Evolution of the cancer genome. *Nat Rev Genet* 2012;13:795–806.
9. Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat Rev Cancer* 2013;13:497–510.
10. Vizoso M, Ferreira HJ, Lopez-Serra P, et al. Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. *Nat Med* 2015;21:741–750.
11. Moutinho C, Martinez-Cardus A, Santos C, et al. Epigenetic inactivation of the BRCA1 interactor SRBC and resistance to oxaliplatin in colorectal cancer. *J Natl Cancer Inst* 2014;106:djt322.
12. Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol* 2013;49:211–215.
13. Sandoval J, Heyn H, Moran S, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 2011;6:692–702.
14. Heyn H, Esteller M. DNA methylation profiling in the clinic: applications and challenges. *Nat Rev Genet* 2012;13:679–692.

15. Berman BP, Weisenberger DJ, Aman JF, et al. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet* 2011;44:40–46.
16. Toyota M, Ahuja N, Ohe-Toyota M, et al. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A* 1999;96:8681–8686.
17. Ogino S, Kawasaki T, Kirkner GJ, et al. Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. *J Mol Diagn* 2007;9:305–314.
18. Lawrence MS, Stojanov P, Mermel C, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 2014;505:495–501.
19. Bartlett TE, Zaikin A, Olhede SC, et al. Corruption of the intra-gene DNA methylation architecture is a hallmark of cancer. *PLoS One* 2013;8:e68285.
20. Dawson H, Lugli A. Molecular and pathogenetic aspects of tumor budding in colorectal cancer. *Front Med (Lausanne)* 2015;2:11.
21. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61:759–767.
22. Naxerova K, Jain RK. Using tumour phylogenetics to identify the roots of metastasis in humans. *Nat Rev Clin Oncol* 2015;12:258–272.
23. Park SY, Gönen M, Kim HJ, et al. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *J Clin Invest* 2010;120:636–644.
24. Shivakumar BM, Rotti H, Vasudevan TG, et al. Copy number variations are progressively associated with the pathogenesis of colorectal cancer in ulcerative colitis. *World J Gastroenterol* 2015;21:616–622.
25. Tejpar S, Zacharenia S, Delorenzi M, et al. Microsatellite instability, prognosis and drug sensitivity of stage II and III colorectal cancer: more complexity to the puzzle. *J Natl Cancer Inst* 2011;103:841–844.
26. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;21:1350–1356.
27. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science* 2013;339:1546–1558.
28. **Hiley C, de Bruin EC, McGranahan N**, et al. Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. *Genome Biol* 2014;15:453–462.
29. Fearon ER, Carethers JM. Molecular subtyping of colorectal cancer: time to explore both intertumoral and intratumoral heterogeneity to evaluate patient outcome. *Gastroenterology* 2015;148:10–13.
30. Diaz LA Jr, Williams RT, Wu J, et al. The molecular evolution of acquired resistance to targeted EGFR blockage in colorectal cancers. *Nature* 2012;486:537–540.
31. **Betgeowda C, Sausen M, Leary RJ**, et al. Detection of circulating tumor DNA in early and late-stage human malignancies. *Sci Transl Med* 2014;6:224ra24.
32. Okugawa Y, Grady WM, Goel A. Epigenetic Alterations in Colorectal Cancer: Emerging Biomarkers. *Gastroenterology* 2015;149:1204–1225.
33. Brocks D, Assenov Y, Minner S, et al. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep* 2014;8:798–806.
34. **Su Y, Subedee A**, Bloushtain-Qimron N, et al. Somatic cell fusions reveal extensive heterogeneity in basal-like breast cancer. *Cell Rep* 2015;11:1549–1563.
35. **Spaderna S, Schmalhofer O**, Hlubek F, et al. A transient, EMT-linked loss of basement membranes indicates metastasis and poor survival in colorectal cancer. *Gastroenterology* 2006;131:830–840.

Author names in bold designate shared co-first authorship.

Received October 8, 2015. Accepted August 2, 2016.

Reprint requests

Address requests for reprints to: Manel Esteller, MD, PhD, Cancer Epigenetics and Biology Program, Bellvitge Biomedical Research Institute, 3rd Floor, Hospital Duran i Reynals, Avda. Gran Via de L'Hospitalet 199-203, 08908 L'Hospitalet, Barcelona, Catalonia, Spain. e-mail: mesteller@idibell.cat; fax: (+34) (93) 2607219.

Conflicts of interest

The authors disclose no conflicts.

Funding

The research leading to these results has received funding from European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement HEALTH-F2-2011-259015–COLTHERES and HEALTH-F5-2010-258236–SYSCOL, European Research Council Advanced Grant #268626–EPINORC project, Ministerio de Ciencia e Innovacion grant number SAF2011-22803, Institute of Health Carlos III under the Integrated Project of Excellence no. PIE13/00022 (ONCOPROFILE) and the Spanish Cancer Research Network no. RD12/0036/0039, Cellex Foundation, Olga Torres Foundation and the Catalan Government (AGAUR-2014SGR633). Manel Esteller is an Institució Catalana de Recerca i Estudis Avançats Research Professor.