Novel Epigenetic 12-gene Signature Predictive of Poor Prognosis and MSI-like Phenotype in Human Metastatic Colorectal Carcinomas

Valentina Condelli
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Giovanni Calice
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Alessandra Cassano
Policlinico A Gemelli: Policlinico Universitario Agostino Gemelli

Michele Basso
Policlinico A Gemelli: Policlinico Universitario Agostino Gemelli

Maria Grazia Rodriquenz
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Angela Zupa
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Francesca Maddalena
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Fabiana Crispo
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Michele Pietrafesa
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Michele Aieta
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Alessandro Sgambato
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata

Giampaolo Tortora
Policlinico A Gemelli: Policlinico Universitario Agostino Gemelli
Pietro Zoppoli
IRCCS CROB: Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata
Matteo Landriscina (matteo.landriscina@unifg.it)
Laboratory of Preclinical and Translational Research, Istituto di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata (IRCCS-CROB), Rionero in Vulture (Pz), Italy;
https://orcid.org/0000-0003-0591-9799

Research

Keywords: promoter methylation, colorectal carcinoma, gene signature, prognosis, MSI-like signature, CIMP status

Posted Date: September 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-74746/v1

Abstract

Background. Epigenetic remodeling is responsible for tumor progression and drug resistance in human colorectal carcinoma (CRC). A subgroup of human CRCs exhibits the CIMP status, with extensive hypermethylation events in promoter regions of several genes, even though the prognostic significance of CIMP is controversial. This study addressed the hypothesis that DNA methylation profiling may identify metastatic CRC (mCRC) subtypes with different clinical behavior.

Methods. Global methylation profile was comparatively analyzed between 24 first-line primary-resistant and 12 drug-sensitive mCRCs (in-house cohort), two subgroups of tumors with significantly different outcome. Methylation and gene expression data from 33 mCRCs of the TCGA COAD dataset (TCGA COAD cohort) were used to identify, among differentially methylated genes, a prognostic signature of functionally methylated genes. Clusters of mCRCs with different methylation patterns were further characterized for DNA mutational load, gene copy number and gene expression profiles. Human CRC HT29 and HCT116 cell lines were adapted to growth in presence of oxaliplatin and irinotecan and used as *in vitro* model to validate gene expression data.

Results. Twelve functionally methylated genes yielded a hierarchical clustering of patients in two well-defined clusters with hypermethylated tumors characterized by a significantly worse relapse-free and overall survival compared to hypomethylated cancers and this was reproduced in both the in-house and the TCGA COAD cohorts. Interestingly, the hypermethylated poor prognosis cluster was enriched of CIMP-high and MSI-like cases. Furthermore, methylation events were enriched in genes located on q-arm of chromosomes 13 and 20, two chromosomal regions with gain/loss alterations strongly associated with adenoma-to-carcinoma progression. Finally, the expression of the 12-genes signature and MSI-enriching genes was confirmed in two independent oxaliplatin- and irinotecan-resistant CRC cell lines.

Conclusions. These data represent the proof of concept that the hypermethylation of specific sets of genes may provide prognostic information being able to identify a subgroup of mCRCs with poor prognosis.

Background

Colorectal carcinoma (CRC) is among the most frequent causes of cancer-related death in Western countries (1) and, besides significant improvements in treatment strategies, the prognosis of metastatic CRC (mCRC) remains poor (2). First-line therapy for mCRC includes either chemotherapeutics (i.e., fluoropyrimidines, oxaliplatin, irinotecan) or molecular-targeted agents and standard regimens are based on doublet- or triplet-chemotherapy regimens (i.e., FOLFOX, XELOX, FOLFIRI and FOLFOXIRI) combined with anti-angiogenic (i.e., bevacizumab) or anti-EGFR (i.e., cetuximab or panitumumab) monoclonals. However, the main cause of treatment failure is drug resistance and currently a major clinical issue is tumor molecular profiling to improve our capacity to predict patients' prognosis and design personalized treatments. At present, among several proposed biomarkers, NRAS, KRAS and BRAF mutational status

and microsatellite instability (MSI) are the most reliable tools in clinical setting, allowing the selection of RAS/BRAF wild type tumors that are more likely to respond to anti-EGFR agents (3, 4) and MSI tumors that are more likely to respond to immune checkpoint inhibitors. No biomarkers are available to predict resistance/sensitivity to first-line chemotherapy and antiangiogenic agents.

For long time, genetic aberrations and mutations in oncogenes and tumor suppressor genes have been considered the only molecular events driving tumor initiation and progression. Nowadays, epigenetic alterations gained consideration as additional crucial events in the multi-step carcinogenetic process (5, 6). Indeed, the emerging leaning suggests a crosstalk between gene mutations and epigenetic alterations (5) and this interplay is responsible for the activation of signaling pathways regulating cancer hallmarks with an impact on clinical outcomes. Particularly, the majority of human cancers is characterized by mutations in enzymes (i.e., writers, readers, and erasers) involved in chromatin organization; hence, tumor cells are triggered by epigenetic alterations (7, 8) and this results in loss and gain of functions in genes correlated with tumorigenesis (9), drug resistance, stem cell differentiation (10). DNA methylation is the first epigenetic mechanism reported in humans (11–13) and the evaluation of DNA methylation of CpG island promoters represents the starting point of many cancer studies in this field. Moreover, since methylation remodeling is a rapid event compared to genetic mutations, it is likely that cancer cells preferentially use this mechanism to rapidly adapt to unfavorable conditions and trigger survival pathways and this is particularly relevant in acquired and de novo resistance to anticancer agents (14). Hence, a novel frontier for biomarker development is the identification of gene methylation patterns to predict clinical outcome, thus driving the selection of patients who may benefit from specific anticancer treatments. In such a context, this study examined the DNA methylation pattern of a cohort of primary-resistant mCRCs in comparison with drug-sensitive tumors treated with 1st-line FOLFOX or FOLFIRI backbone chemotherapy to identify epigenetic modifications able to predict patient's prognosis.

Methods

Patients and samples collection.

Twenty-four primary resistant mCRCs treated with 1st-line FOLFOX (8 patients) or FOLFIRI (16 patients) chemotherapy in combination or not with bevacizumab or anti-EGFR agents and 12 drug-sensitive mCRCs (4 treated with FOLFOX and 8 with FOLFIRI combined with molecular targeted agents) were selected for this study. Patients' characteristics are described in Additional Table 1. Tumors were selected based on the evidence of tumor progression (primary-resistant) or partial/complete response (drug-sensitive) at the first radiological assessment after 2–3 months of first-line therapy. Patients were enrolled at the Medical Oncology Units of the IRCCS-CROB (Rionero in Vulture, Italy) and the Fondazione Policlinico Universitario "A. Gemelli" (Rome, Italy) and were called "in-house" cohort.

Cell Lines and in vitro drug-resistant models.

Human HT29 and HCT116 CRC oxaliplatin (Oxa) and irinotecan (Iri) resistant (R) cell lines were obtained in our laboratory as described in Additional data. The surviving chronically resistant cells were named

HT29-OxaR and HCT116-OxaR and HT29-IriR and HCT116-IriR cells. Experiments were carried out at 70% cell confluence and confirmed at least in three independent replicates. Cell cultures were routinely screened for mycoplasma contamination.

Array-based DNA methylation profiling.

DNA was isolated from formalin-fixed, paraffin-embedded (FFPE) CRC specimens (Additional data). Five hundred ng of total gDNA were treated with sodium bisulfite using the Zymo EZ DNA Methylation Kit (Zymo Research) according to the Infinium HD Methylation Assay protocol. The bisulfite converted gDNA was hybridized on the Infinium Human Methylation 850 BeadChip array (Illumina Inc), following manufacturer's instructions. After washing and staining procedures, chips were scanned by the Illumina HiScanSQ system.

Bioinformatics analysis.

To identify a prognostic signature of functionally methylated (fMET) genes, a multistep analysis was performed as described in Additional Fig. 1. In the first step, global DNA methylation profiles were obtained from primary-resistant and drug-sensitive in-house tumor specimens, as described in Additional data. In a subsequent step, since gene expression data from in-house tumors were not available, fMET genes were defined based on gene expression and methylation data from The Cancer Genome Atlas COlon ADenocarcinoma (TCGA COAD) data collection. To this purpose, gene expression, methylation, DNA sequencing, gene copy number and clinical data of 33 patients with mCRC from the TCGA COAD database were downloaded using the TCGA biolink package (Additional Table 2). After normalization of expression data, any further analysis was restricted to Differentially Methylated Genes (DMGs) with an absolute methylation delta greater than 0.1 and a p-value < 0.05 (6672 genes from TCGA COAD). Among in-house DMGs, we defined fMET genes, those showing a significant (FALSE DISCOVERY RATE, FDR, adjusted p-value < 0.05) inverse correlation (Rsquared > 0.1) between the promoter's methylation and the gene expression profile. Only genes present in both in-house DMGs and TCGA COAD datasets were further analyzed. Hierarchical clustering was performed using Ward's linkage and Euclidean distance.

To evaluate differences in prognosis, Kaplan-Meier and log-rank test were applied to both TCGA COAD and in-house datasets. To identify biological and clinical differences between poor and good prognosis clusters, Gene Set Analysis (GSA) was performed on all gene sets collections of the mSigDB repository and DNA sequencing data were comparatively analyzed.

RNA extraction and Real Time RT-PCR analysis.

Total RNA was extracted using the TRIzol Reagent (Invitrogen) from parental and drug-resistant strains of HCT116 and HT29 CRC cells lines. In specific experiments, RNA was obtained from cell lines exposed to 5-Azacytidine (5-Aza-dC) (Sigma-Aldrich), an inhibitor of DNA methyltransferase 1 (DNMT1), at a final concentration of 10 uM for 72 h. Real Time PCR analysis was described in Additional data. Primers are reported in Additional Table 3.

Apoptosis assay.

Parental and drug-resistant cell lines were seeded on day 0 in 6-well plates in triplicate and incubated on day 1 in normal medium or exposed to 10 μM 5-Aza-dC for 72 h. After 48 h, cells were further treated with 3 μM Oxa or 2 μM Iri for 24 h. Apoptosis was evaluated by cytofluorimetric analysis (Additional data).

Results

DNA methylation profile is remodeled in primary-resistant mCRCs.

Primary-resistant mCRCs were selected for this study as representative colorectal malignancies with poor prognosis and poor response to anti-cancer agents (15). Thus, in order to identify epigenetic alterations with prognostic relevance, global DNA methylation was assessed on 24 mCRCs primary-resistant to 1st-line FOLFOX (8 patients) or FOLFIRI (16 patients) chemotherapy combined or not with molecular targeted agents. Twelve drug-sensitive mCRCs (4 treated with FOLFOX and 8 with FOLFIRI combined with molecular targeted agents) were used as controls to obtain the differential methylation profile between primary-resistant and drug-sensitive tumors (in-house cohort; Additional Table 1). Differential methylation profiles were analyzed in a multistep process, as described in Additional Fig. 1. Indeed, 74843 and 36876 probes were significantly differentially methylated in FOLFOX and FOLFIRI datasets (p-value < 0.05), respectively (Fig. 1A-B) and these were widely distributed between different genomic regions (Fig. 1C-D).

Epigenetic alterations predict prognosis in human mCRCs.

Since it is well established that promoter hypo/hypermethylation is the main mark resulting in gene expression modifications (16), only genes with methylation modifications in promoter regions (with a p-value < 0.05 and an absolute difference of Beta value > 0.1) were used in subsequent analyses. Using this more restrictive cut-off, 7432 probes, corresponding to 4533 genes, for patients treated with 1st-line FOLFOX and 5005 probes, corresponding to 3803 genes, for patients treated with 1st-line FOLFIRI resulted differentially methylated between drug-resistant and drug-sensitive tumors. These two lists were named, respectively, FOLFOX_DMGs and FOLFIRI_DMGs (Additional Dataset). We next quested whether these DMGs were also fMET, with a methylation profile consistent with the gene expression profile. Since we could not obtain gene expression data from in-house colorectal tumor samples due to the poor amount and quality of RNA purified from paraffin-embedded specimens, this issue was addressed using a cohort of 33 mCRCs obtained from the TCGA COAD database, which provides gene expression, DNA methylation, DNA sequencing and clinical data for each patient (Additional Table 2). From the analysis of the TCGA COAD database, we obtained 741 fMET genes defined as COAD fMET genes. Among these 741 TCGA COAD fMET genes, 542 resulted DMGs in FOLFOX dataset and 248 in FOLFIRI dataset. Applying more stringent filters (p-value < 0.01 and absolute difference of Beta value > 0.2) on the FOLFOX_DMGs and the FOLFIRI_DMGs datasets and the COAD fMET genes (R2 > 0.25), this list was restricted to 8 fMET genes for FOLFOX and 7 fMET genes for FOLFIRI datasets, with 3 of them common to both datasets (Additional Table 4). Interestingly, hierarchical clustering on these sets of genes (using both expression and methylation data) allowed to separate the TCGA COAD samples into two quite homogeneous

clusters characterized by over or under expression (Fig. 2A and D) and hypo or hyper methylation (Fig. 2B and E) of, respectively, the above 8 and 7 fMET genes. A similar separation was obtained in our in-house FOLFOX and FOLFIRI cohorts upon hierarchical clustering of methylation data using the same gene sets (Fig. 2C and F). In order to evaluate the prognostic relevance of these 8 and 7 fMET gene signatures, log-rank test was performed on both the TCGA COAD and the in-house cohorts using the two previously obtained clusters. Noteworthy, a significant (p-value < 0.05) separation for both Relapse Free Survival (RFS) and Overall Survival (OS) was observed between hypermethylated/underexpressed tumors, characterized by worst prognosis, and hypomethylated/overexpressed tumors using both FOLFOX (Additional Fig. 2A-D) and FOLFIRI (Additional Fig. 3A-D) 8 and 7 genes signatures. Log-rank test performed on in-house FOLFOX (Additional Fig. 4A-B) and FOLFIRI (Additional Fig. 4C-D) cohorts resulted in a significant (p-value < 0.05) separation of the two clusters considering RFS and a non-significant separation considering OS. Based on this evidence, the two cohorts were labeled as “good” and “poor” prognosis clusters.

As a next step, we combined the 8 and 7 fMET gene signatures from FOLFOX and FOLFIRI datasets obtaining a new signature of 12 fMET genes, being 3 of them common to both datasets (Fig. 3). Hierarchical clustering on such signature separated TCGA COAD patients into two well-defined cohorts (22 hypo and 11 hypermethylated tumors, Fig. 3A-B and Additional Table 2). A similar clustering was obtained in our in-house FOLFOX dataset (Fig. 3C) and partially in the FOLFIRI dataset (Fig. 3D). Furthermore, upon combination of in-house FOLFOX and FOLFIRI datasets, the 12-genes signature obtained a more defined separation between 6 hypo and 30 hypermethylated tumors (Fig. 3E). Log-rank test confirmed that the cohort with hypermethylation of the 12-genes signature is characterized by significantly shorter survival (both OS and RFS; p-value < 0.05) compared to the cohort with hypomethylation of these genes in both TCGA COAD and in-house cohorts (Fig. 3F-G). Altogether, these data suggest that this 12-fMET genes signature discriminates between mCRC patients with good (hypomethylated tumors) and poor (hypermethylated tumors) prognosis.

The poor prognosis hypermethylated cluster is characterized by a MSI-like phenotype and is enriched of CIMP-high tumors.

The poor prognosis hypermethylated and the good prognosis hypomethylated clusters were further characterized respect their clinical and biological profiles using gene expression and DNA sequencing and gene copy number data from the TCGA COAD database. No major differences were observed between the poor and good prognosis clusters respect to T and N categories and sites of primary tumor (right versus left colon) (Additional Table 2). Similarly, no major differences were observed respect to the tumor mutational load, with the exception of two hypermutated cases in the poor prognosis cluster (Additional Fig. 5, insert). Interestingly, specific gene mutations were differently distributed between the two subgroups, being mutations in SRGAP2B, AC007682.1, AC104820.2, AF121898.3 and MROH5 genes enriched in the good prognosis cluster and mutations in GRP98, NRXN2, HDN1 and TTC40 genes in the poor prognosis cluster (Additional Fig. 5). Consistently, several gene aberrations were significantly more abundant in the good prognosis cluster (Additional Fig. 6).

A differential gene expression comparison between the two clusters yielded 116 Differentially Expressed Genes (DEGs) (false discovery rate, FDR, adjusted p-value < 0.05 and abs(logFC) > 0.58). Among these, 80 genes were downregulated and 36 of them hypermethylated in the poor prognosis cluster and conversely upregulated/hypomethylated in the good prognosis subgroup (Additional Fig. 7). Gene Set Analysis (GSA) was performed on the gene set collection of the mSigDB repository, obtaining significant enrichments for signaling pathways and positional collections. Among different signaling pathways (Fig. 4A), GSA identified the Watanabe gene set, which includes genes discriminating between MSI and MSS (microsatellite instability/stability) colorectal cancers (17). The statistical analysis identified 7 genes in our list of DEGs, which enrich Watanabe gene dataset and whose expression profile is consistent with a separation of the TCGA cohort in good and poor prognosis clusters (Fig. 4B). Noteworthy, 5/7 genes enriching Watanabe gene set were significantly downregulated due to hypermethylation. Since this observation suggests that the 12-genes hypermethylated signature identifies a subgroup of mCRCs with a MSI-like phenotype, an independent MSI-like gene expression signature was evaluated for the capacity to reproduce the separation of the TCGA COAD cohort in the same good and poor prognosis clusters, according to the 12-genes signature. Noteworthy, the MSI-like gene expression signature of Pačínková et al (18) mirrored the separation of the 33 mCRCs TCGA cohort in the same clusters as obtained by our 12-genes signature (Fig. 4C). Consistently, five genes from the Pačínková signature (i.e., PLAGL2, ACSL6, ARID3A, NKD1, TNNT1) were characterized by an expression profile consistent with the expression profile of the MSI-like poor prognosis TCGA cluster.

Interestingly, 15–20% of human CRCs are characterized by the CpG-island methylator phenotype (CIMP), subdivided in CIMP-high (CIMP-H) and CIMP-low (CIMP-L) and this correlates with the MSI phenotype (19). Thus, the relationship between our 12-genes hypermethylated signature and CIMP status was evaluated in TCGA COAD dataset and in our in-house cohort according to Hinoue et al (20). Interestingly, the poor prognosis TCGA hypermethylated cohort was enriched of CIMP-H cases, being 5 out of 6 CIMP-H samples classified as poor prognosis patients, whereas the hypomethylated good prognosis cohort was enriched on no-CIMP tumors. CIMP-L cases were distributed between the two subgroups, being classified in 7 out of 11 cases as good prognosis patients (Fig. 4D). This difference between the groups was statistically significant by two-sided Fisher exact test (p-value < 1e-02). Consistently, the vast majority of CIMP-H tumors were characterized by the hypermethylation of the 12-genes signature in our in-house cohort (Additional Fig. 8A). Thus, the prognostic relevance of our 12-genes signature was compared to CIMP status and, noteworthy, the RFS and the OS analyses showed a better capacity of our 12-genes hypermethylated signature to predict poor prognosis in both the TCGA COAD (Fig. 4E-F) and the in-house cohorts (Additional Fig. 8B-C). Altogether, these observations strongly support the conclusion that the 12-genes hypermethylated signature correlates with a MSI-like phenotype and is characterized by a better capacity to predict prognosis compared to CIMP status.

Hypermethylated genes are enriched on arms q of chromosomes 13 and 20.

Among positional collections, GSA identified enrichments of chromosome 13 arm q and chromosome 20 arm q gene sets (Fig. 5A). Interestingly, 23/80 downregulated genes in our list of DEGs are located on

chromosome 13 arm q and 8 of them are hypermethylated. Consistently, 12/80 downregulated genes are located on chromosome 20 arm q. Noteworthy, the expression profile of each of these gene sets (Additional Fig. 9A-B) and of their combination (Fig. 5B) mirrored the separation of the TCGA COAD cohort in the good and poor prognosis clusters, obtained according to the 12-genes signature. These data suggest an enrichment of methylation events in genes located in specific chromosomal regions in mCRCs with poor prognosis. In such a context, the enrichment analysis for signaling pathways identified the Ding lung cancer expression by copy number and the Nikolsky breast cancer 20q12-13 amplicon gene sets (Fig. 4A). These authors reported respectively a correlation between the copy number variation and the expression of 26 genes in lung cancers (21) and the identification of 149 genes in amplicon 20q12-13 in breast tumors (22). As expected, all the genes enriching the Nikolsky gene set overlap with our chromosome 20 arm q genes and 5 out of 6 Ding genes with our chromosome 13 arm q genes. In both cases, the expression profile of these genes reproduced the separation of the TCGA dataset in good and poor prognosis clusters (Additional Fig. 9C-D). Altogether, these data highlight the relevance of expression/methylation modifications of genes located on chromosomes 13 and 20 in human colorectal cancer.

Epigenetic modifications are reproduced in drug-resistant cell models.

To validate epigenetic data obtained from FOLFOX and FOLFIRI primary-resistant mCRCs, we generated in vitro drug-resistant cellular models chronically adapted to oxaliplatin (Oxa; HCT116-OxaR and HT29-OxaR cells) or irinotecan (Iri; HCT116-IriR and HT29-IriR cells). In preliminary experiments, apoptotic cell death was evaluated in drug-sensitive and drug-resistant cell lines in response to Oxa or Iri in combination or not with the demethylating agent 5-Aza-dC. These experiments confirmed that cell lines chronically exposed to chemotherapeutics are indeed poorly sensitive to Oxa and Iri and that drug resistance is reverted upon treatment with 5-Aza-dC (Fig. 6A-B). Since these data support the hypothesis that methylation modifications are responsible for resistance to Oxa and Iri in these CRC cell lines, in subsequent experiments, drug-resistant cell lines were used to validate the expression profiles of the 12-fMET genes signature. Real Time RT-PCR analysis of the 12-genes confirmed a significant downregulation of 10 genes in HCT116-OxaR and HT29-OxaR cell lines (Fig. 6C-D) and 9 and 7 genes in, respectively, HCT116-IriR and HT29-IriR cell lines (Fig. 6E-F) compared to the respective drug-sensitive cell lines. NROB2 was undetectable in both cell lines (data not shown). Noteworthy, the pretreatment of drug-resistant cell lines by 5-Aza-dC resulted in a significant upregulation of the majority of the downregulated genes: 7 genes in, respectively, HCT116-OxaR and HT29-OxaR cells (Fig. 6C-D) and 6 genes in, respectively, HCT116-IriR and HT29-IriR cell lines (Fig. 6E-F). These data suggest that the downregulation of genes belonging to the 12-genes signature correlates with onset of drug resistance in CRC cell lines and that this downregulation is likely mediated by methylation events.

In parallel analyses, starting from GSA results suggesting an enrichment of a MSI-like phenotype in hypermethylated poor-prognosis tumors, 7 genes enriching Watanabe pathway and 5 representative genes belonging to the mismatch repair system (18) were evaluated in drug-resistant cell lines (Additional Fig. 10). Indeed, PCR data confirmed the downregulation of the majority of these genes in drug-resistant

cell lines with HT29-OxaR and HT29-IriR characterized by a more evident MSI-like phenotype (Additional Fig. 10).

Discussion

Molecular profiling is a major objective in mCRC in order to define prognostic homogeneous subgroups of patients and deliver personalized therapies (23). In such a context, genetic mutations and gene expression profiles have been proposed as potential predictive/prognostic biomarkers and some of them entered in daily clinical practice (24, 25). By contrast, while preclinical evidence is currently available on the role of epigenetic modifications in tumor progression (26), their significance as prognostic tools is so far mostly unexplored and/or results are conflicting (27, 28). Indeed, the majority of human CRCs are characterized by global hypomethylation and promoter-specific DNA methylation (29), whereas 15–20% of them exhibit the CIMP status, with extensive and co-ordinate patterns of hypermethylation events in numerous CpG islands surrounding the promoter regions of several genes whose transcriptional silencing contributes to onset and progression of CRC (30, 31). The prognostic significance of CIMP is controversial, with several studies suggesting that CIMP status is an independent prognostic factor of poor outcome (32, 33). However, this conclusion is still debated likely due to different definitions of CIMP among studies with respect to methylation loci and laboratory methods (32). Recently, in contrast with the conclusion that CIMP status predict poor prognosis, a prognostic score based on low methylation level of seven CpG sites was strongly associated with poor CRC survival (32). Thus, in order to better define the prognostic relevance of promoter hypermethylation events in human CRC, this study was designed to characterize epigenetic signatures helpful to identify mCRCs molecular subgroups with defined clinical behavior. In such a context, primary-resistant colorectal carcinomas were selected as cases representative of poor outcome (33), based on the evidence that methylation modifications are key events used by cancer cells to rapidly adapt to unfavorable environments and acquire drug resistance (34, 35). Our data suggest that the methylation profile of 12 functionally methylated genes is predictive of clinical outcome being able to clusterize two independent mCRC cohorts (i.e., the TCGA COAD and the in-house datasets) in two well-defined clusters with hypermethylated tumors characterized by worse prognosis and a MSI-like phenotype compared to hypomethylated tumors. As expected, the hypermethylated poor prognosis cluster is enriched of CIMP-H cases, but our 12-genes signature showed a better capacity to identify hypermethylated malignancies with poor prognosis compared to CIMP status.

Clinically relevant is the observation that the poor prognosis cohort with hypermethylation of the 12-genes signature is characterized by a MSI-like phenotype. Indeed, the CIMP-H status is frequently associated with methylation of MLH1 promoter region and consequent gene silencing (20), this resulting in acquisition of MSI and strong immune activation (36). In such a context, the relevance of the MSI phenotype is a controversial issue in CRCs. Indeed, TNM stage II colorectal tumors with deficient mismatch repair system/MSI-high phenotype are characterized by good prognosis, but do not benefit from 5-fluorouracile adjuvant chemotherapy (37). By contrast, there are several controversies about whether the MSI-high phenotype is a good prognostic factor in mCRC patients. Some studies proved that

MSI-high is a beneficial factor associated with a better outcome (38, 39), whereas, several others came to the opposite conclusion, indicating MSI-high as an adverse factor (40, 41). On the other hand, unlike MSS CRCs, MSI-high CRCs showed a much better response to immune checkpoint inhibitors (42, 43). More recently, several MSI-like gene expression signatures were also proposed with likewise controversial significance (18, 44, 45). Our data suggest that the 12-genes hypermethylated cohort of mCRCs is characterized by a MSI-like phenotype and that the methylation profile of the 12-genes signature may represent an alternative strategy to better define a subgroup of mCRCs with a CIMP-H status and a MSI-like phenotype, characterized by a poor clinical outcome. In such a context, a limitation of our study is the limited number of cases in our series, even though the significance of the 12-genes signature was proven in two independent mCRC datasets. Thus, its prognostic relevance needs to be further validated in a larger series to establish its wider use in a clinical setting and further studies are also needed to establish whether this gene signature may improve our capacity to select mCRCs amenable to immunotherapy.

In a biological perspective, it is noteworthy that two genes in our list, C13orf18 and LRRC2, are putative oncosuppressor genes. Indeed, C13orf18, a gene with a hypermethylation status common to both FOLFOX and FOLFIRI first-line datasets, is frequently hypermethylated and silenced in cervical cancer and its re-expression results in growth inhibition of cervical tumor cells (46). In addition, C13orf18 is significantly downregulated in our drug-resistant cell lines and its expression is reverted by the demethylating treatment. LRRC2 gene expression is impaired in renal carcinoma cell lines, this also suggesting a putative oncosuppressive gene function (47). Thus, the functional hypermethylation of both these genes in a cohort of mCRCs with poor outcome supports the hypothesis that these genes may play a role in colorectal carcinogenesis.

Finally, GSA enrichment analysis suggests that methylation events are enriched in genes located on arms q of chromosomes 13 and 20 in mCRCs with poor prognosis, supporting the hypothesis that epigenetic remodeling may not occur in a random manner during colorectal carcinogenesis, but rather it may be a coordinated process with hypo/hypermethylation of selective genomic regions. This hypothesis is consistent with the evidence that accumulation of gains/losses in 13q and 20q regions are strongly associated with adenoma-to-carcinoma progression (48) and that mutations in the same chromosomal regions are relevant in other human malignancies. Indeed, the GSA identified the Ding lung cancer expression by copy number and the Nikolsky breast cancer 20q12-13 amplicon gene sets. Ding et al reported a correlation between the copy number variation and the expression of 26 genes in lung cancers (21), whereas Nikolsky et al the identification of 149 genes in amplicon 20q12-13 in breast tumors (22). It is intriguing that genes enriching Ding and Nikolsky gene sets reproduced the clustering of TCGA mCRCs in the same cohorts as obtained by the 12-genes signature. Consistently, the vast majority of genes enriching the Nikolsky and the Ding gene sets overlap with our chromosome 20 arm q or 13 arm q genes.

In conclusion, this study provides the proof of concept that epigenetic profiling may represent a strategy to predict patients' prognosis in mCRC and that a novel 12-genes methylation signature may better define a poor prognosis subgroup of mCRCs with CIMP-H status and a MSI-like phenotype.

List Of Abbreviations

5-Aza-dC, 5-Azacytidine

CIMP, CpG-island methylator phenotype

CRC, colorectal carcinoma

DEG, Differentially Expressed Genes

DMG, Differentially Methylated Genes

DNMT, DNA methyltransferase

FDR, FALSE DISCOVERY RARE

fMET, functionally methylated

GSA, Gene Set Analysis

HM, Heatmap

Iri, Irinotecan

Iri-R, Irinotecan-Resistant

mCRC, metastatic colorectal carcinoma

MSI, microsatellite instability

MSS, microsatellite stability

OS, Overall Survival

Oxa, Oxaliplatin

Oxa-R, Oxaliplatin-Resistant

RFS, Relapse Free Survival

TCGA COAD, The Cancer Genome Atlas COlon ADenocarcinoma

Declarations

Ethics approval and consent to participate

All experiments were performed in accordance with protocols approved by Ethics Committee of IRCCS CROB (reference number 20120010288).

Consent for publications

Express written informed consent to use biological specimens for investigational procedures was obtained from all patients.

Availability of data and materials

DNA methylation data generated in this study have been deposited at the NCBI GENE expression Omnibus repository (GEO) and are accessible through the accession number GSE148766 (www.ncbi.nlm.nih.gov/geo/).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by Finanziamento Ricerca Corrente, Ministero della Salute, Italy.

Authors' contributions

VC, PZ, ML conceptualized the study. VC, FM, FC, MP performed experiments and analyzed data. AC, MB, MGR, MA, AZ, GT provided clinical resources for the project. GC, PZ performed biostatistical analysis. VC, ML, GC, PZ wrote the manuscript. ML and VC supervised the study. FM, FC, MP, AS, GT reviewed and edited the manuscript.

Acknowledgements

Not applicable.

References

1. Màrmol L, Sánchez-de-Diego C, Pradilla Dieste A, Cerrada E, Rodriguez Yoldi MJ. Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. Int J Mol Sci. 2017;18(1). pii:E197.
2. Chakedis J, Schmidt CR. Surgical Treatment of Metastatic Colorectal Cancer. Surg Oncol Clin N Am. 2018;27(2):377–99.
3. Elez E, Argilés G, Tabernero J. First-Line Treatment of Metastatic Colorectal Cancer: Interpreting FIRE-3, PEAK, and CALGB/SWOG 80405. Curr Treat Options Oncol. 2015;16(11):52.
4. Loupakis F, Depetris I, Biason P, Intini R, Prete AA, Leone F, et al. Prediction of Benefit from Checkpoint Inhibitors in Mismatch Repair Deficient Metastatic Colorectal Cancer: Role of Tumor Infiltrating

Lymphocytes. Oncologist. 2020;25:1–7.

5. Park JW, Han JW. Targeting epigenetics for cancer therapy. Arch Pharm Res. 2019;42(2):159–70.
6. Mummaneni P, Shord SS. Epigenetics and oncology. Pharmacotherapy. 2014;34(5):495–505.
7. Vacante M, Borzì AM, Basile F, Biondi A. Biomarkers in colorectal cancer: Current clinical utility and future perspectives. World J Clin Cases. 2018;6(15):869–81.
8. Jones PA, Issa JP, Baylin S. Targeting the cancer epigenome for therapy. Nat Rev Genet. 2016;17(10):630–41.
9. Condelli V, Crispo F, Pietrafesa M, Lettini G, Matassa DS, Esposito F, et al. HSP90 Molecular Chaperones, Metabolic Rewiring, and Epigenetics: Impact on Tumor Progression and Perspective for Anticancer Therapy. Cells. 2019;8(6). pii:E532.
10. Yang X, Lay F, Han H, Jones PA. Targeting DNA methylation for epigenetic therapy. Trends Pharmacol Sci. 2010;31(11):536–46.
11. Jones PA. At the tipping point for epigenetic therapies in cancer. J Clin Invest. 2014;124(1):14–6.
12. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012;150(1):12–27.
13. Heyn H, Esteller M. DNA methylation profiling in the clinic: applications and challenges. Nat Rev Genet. 2012;13(10):679–92.
14. Glasspool RM, Teodoridis JM, Brown R. Epigenetics as a mechanism driving polygenic clinical drug resistance. Br J Cancer. 2006;94(8):1087–92.
15. Wasan HS, Gibbs P, Sharma NK, Taieb J, Heinemann V, Ricke J, et al. First-line selective internal radiotherapy plus chemotherapy versus chemotherapy alone in patients with liver metastases from colorectal cancer (FOXFIRE, SIRFLOX, and FOXFIRE-Global): a combined analysis of three multicentre, randomised, phase 3 trials. Lancet Oncol. 2017;18(9):1159–71.
16. Luo C, Haikova P, Ecker JT. Dynamic DNA methylation: In the right place at the right time. Science. 2018;361(64099):1336–40.
17. Watanabe T, Kobunai T, Toda E, Yamamoto Y, Kanazawa T, Kazama Y, et al. Distal colorectal cancers with microsatellite instability (MSI) display distinct gene expression profiles that are different from proximal MSI cancers. Cancer Res. 2006;66(20):9804–8.
18. Pačinkovà A, Popovici V. Cross-platform Data Analysis Reveals a Generic Gene Expression Signature for Microsatellite Instability in Colorectal Cancer. Biomed Res Int. 2019; 6763596. doi:10.1155/2019/6763596.
19. Advani SM, Advani P, DeSantis SM, Brown D, VonVille HM, Lam M, et al. Clinical, Pathological, and Molecular Characteristics of CpG Island Methylator Phenotype in Colorectal Cancer: A Systematic Review and Meta-analysis. Trans Oncol. 2018;11(5):1188–201.
20. Hinoue T, Weisenberger DJ, Lange CP, Shen H, Byun HM, Van Der Berg D, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res. 2012;22(2):271–82.

21. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008;455(7216):1069–75.
22. Nikolsky Y, Sviridov E, Yao J, Dosymbekov D, Ustyansky V, Kaznacheev V, et al. Genome-wide functional synergy between amplified and mutated genes in human breast cancer. Cancer Res. 2008;68(22):9532–40.
23. Leichsenring J, Koppelle A, Reinacher-Schick A. Colorectal Cancer: Personalized Therapy. Gastrointest Tumors. 2014;1(4):209–20.
24. Miyamoto Y, Zhang W, Lenz HJ. Molecular Landscape and Treatment Options for Patients with Metastatic Colorectal Cancer. Indian J Surg Oncol. 2017;8(4):580–90.
25. Ciombor KK, Bekaii-Saab T. A Comprehensive Review of Sequencing and Combination Strategies of Targeted Agents in Metastatic Colorectal Cancer. Oncologist. 2018;23(1):25–34.
26. Fang F, Cardenas H, Huang H, Jiang G, Perkins SM, Zhang C, et al. Genomic and Epigenomic Signatures in Ovarian Cancer Associated with Resensitization to Platinum Drugs. Cancer Res. 2018;78(3):631–44.
27. Miranda Furtado CL, Dos Santos Luciano MC, Silva Santos RD, Furtado GP, Moraes MO, Pessoa C. Epidrugs: targeting epigenetic marks in cancer treatment. Epigenetics. 2019;14(12):1164–76.
28. Tse JWT, Jenkins LJ, Chionh F, Mariadason JM. Aberrant DNA Methylation in Colorectal Cancer: What Should We Target? Trends Cancer. 2017;3(10):698–712.
29. Okugawa Y, Grady WM, Goel A. Epigenetic Alterations in Colorectal Cancer: Emerging Biomarkers. Gastroenterology. 2015;149(5):1204–25.
30. Vilar E, Gruber SB. Microsatellite instability in colorectal cancer—the stable evidence. Nat Rev Clin Oncol. 2010;7(3):153–62.
31. Hashimoto Y, Zumwalt TJ, Goel A. DNA methylation patterns as noninvasive biomarkers and targets of epigenetic therapies in colorectal cancer. Epigenomics. 2016;8(5):685–703.
32. Jia M, Gao X, Zhang Y, Hoffmeister M, Brenner H. Different definitions of CpG island methylator phenotype and outcomes of colorectal cancer: a systematic review. Clin Epigenetics. 2016;8:25.
33. Chen KH, Lin LI, Tseng LH, Lin YL, Liau JY, Tsai JH, et al. CpG Island Methylator Phenotype May Predict Poor Overall Survival of Patients with Stage IV Colorectal Cancer. Oncology. 2019;96(3):156–63.
34. Chen CC, Lee KD, Pai MY, Chu PY, Hsu CC, Chiu CC, et al. Changes in DNA methylation are associated with the development of drug resistance in cervical cancer cells. Cancer Cell Int. 2015;15:98.
35. Xu T, Li HT, Wei J, Li M, Hsieh TC, Lu YT, et al. Epigenetic plasticity potentiates a rapid cyclical shift to and from an aggressive cancer phenotype. Int J Cancer. 2020;146(11):3065–76.
36. Testa U, Pelosi E, Castelli G. Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. Med Sci (Basel). 2018;6(2). pii:E31.
37. Kawakami H, Zaanan A, Sinicrope FA. Microsatellite instability testing and its role in the management of colorectal cancer. Curr Treat Options Oncol. 2015;16(7):30.

38. Zeinalian M, Hashemzadeh-Chaleshtori M, Salehi R, Emami MH. Clinical Aspects of Microsatellite Instability Testing in Colorectal Cancer. Adv Biomed Res. 2018;7:28.
39. Battaglin F, Naseem M, Lenz HJ, Salem ME. Microsatellite instability in colorectal cancer: overview of its clinical significance and novel perspectives. Clin Adv Hematol Oncol. 2018;16(11):735–45.
40. Hu W, Yang Y, Qi L, Chen J, Ge W, Zheng S. Subtyping of microsatellite instability-high colorectal cancer. Cell Commun Signal. 2019;17:79.
41. Wang B, Li F, Zhou X, Ma Y, Fu W. Is microsatellite instability-high really a favorable prognostic factor for advanced colorectal cancer? A meta-analysis. World J Surg Onc. 2019;17:169.
42. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. Nat Rev Gastroenterol Hepatol. 2019;16(6):361–75.
43. Xiao Y, Freeman GJ. The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. Cancer Discov. 2015;5(1):16–8.
44. Tian S, Roepman P, Popovici V, Michaut M, Majewski I, Salazar R, et al. A robust genomic signature for the detection of colorectal cancer patients with microsatellite instability phenotype and high mutation frequency. J Pathol. 2012;228(4):586–95.
45. Visone R, Bacalini MG, Di Franco S, Ferracin M, Colorito ML, Pagotto S, et al. DNA methylation of shelf, shore and open sea CpG positions distinguish high microsatellite instability from low or stable microsatellite status colon cancer stem cells. Epigenomics. 2019;11(6):587–604.
46. Huisman C, Wisman GB, Kazemier HG, van Vugt MA, van der Zee AG, Schuuring E, et al. Functional validation of putative tumor suppressor gene C13ORF18 in cervical cancer by Artificial Transcription Factors. Mol Oncol. 2013;7(3):669–79.
47. Liu H, Brannon AR, Reddy AR, Alexe G, Seiler MW, Arreola A, et al. Identifying mRNA targets of microRNA dysregulated in cancer: with application to clear cell Renal Cell Carcinoma. BMC Syst Biol. 2010;4:51.
48. Hermsen M, Postma C, Baak J, Weiss M, Rapallo A, Sciutto A, et al. Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. Gastroenterology. 2002;123(4):1109–19.

Figures

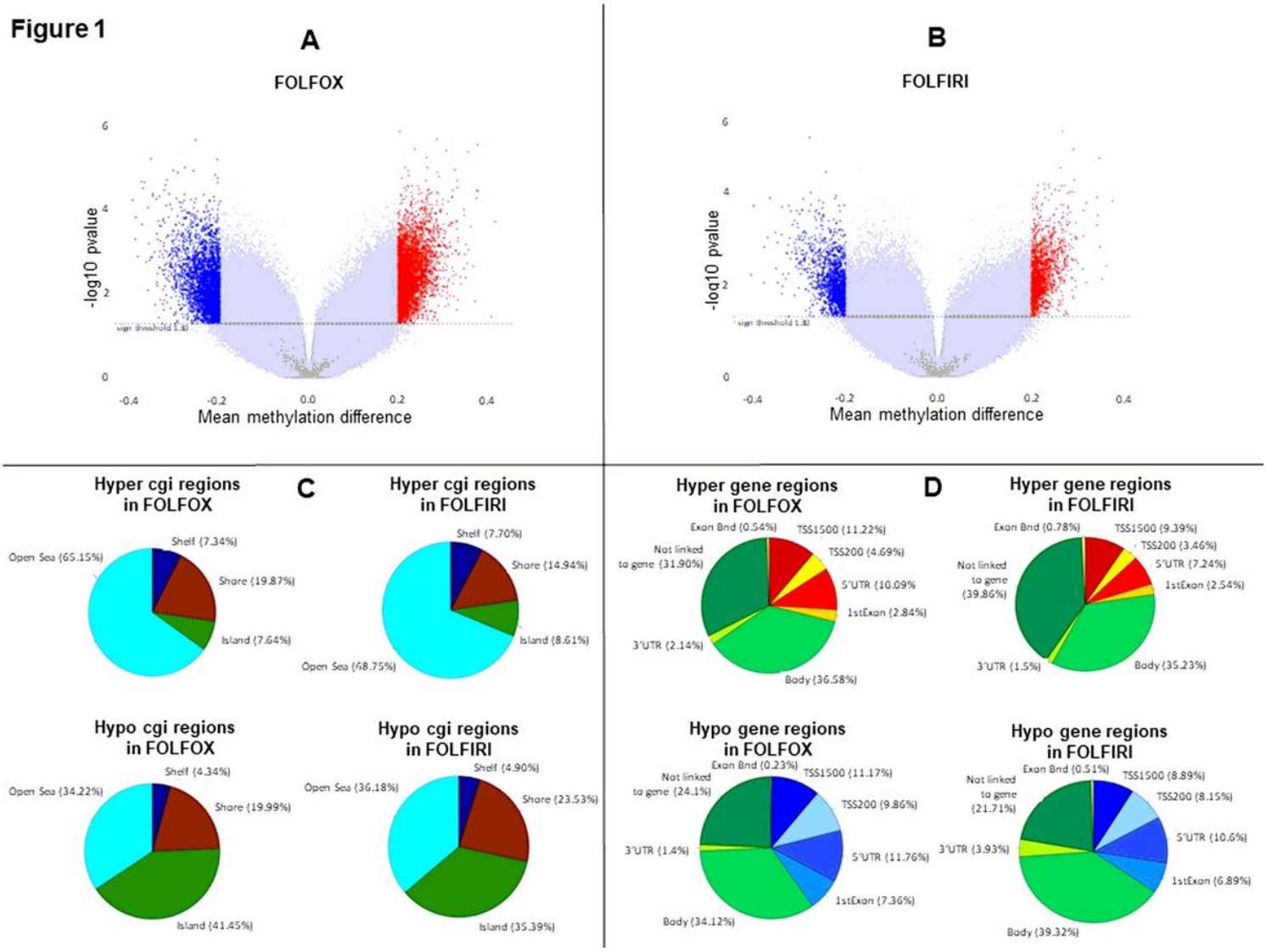

Figure 1

Methylation profile is remodeled in primary-resistant versus drug-sensitive mCRCs. A-B. Volcano Plots representing differentially methylated probes between primary-resistant and drug-sensitive mCRCs. Overall, statistically significant differentially methylated probes were 74843 and 36876 in, respectively, FOLFOX and FOLFIRI datasets (p-value < 0.05). In particular, statistically significant probes with an absolute difference of Beta value > 0.2 are highlighted as blue dots, corresponding to hypomethylated probes (3227 in FOLFOX and 1475 in FOLFIRI datasets), or as red dots, corresponding to hypermethylated probes (3899 in FOLFOX and 1393 in FOLFIRI datasets). C-D. Differentially methylated probes distribution according to genomic regions.

Figure 2

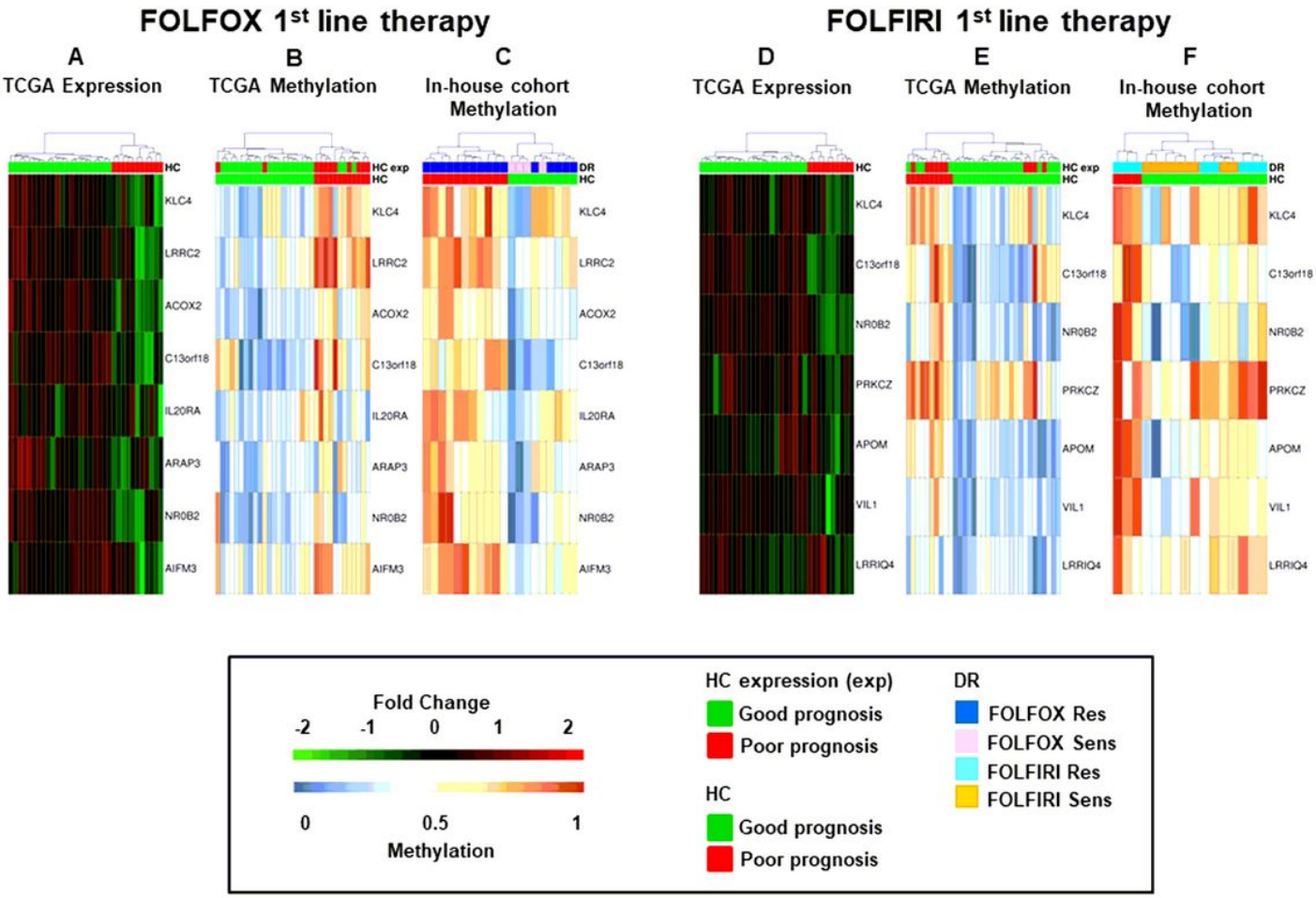

Figure 2

Hierarchical clustering of mCRCs according to the 8 and 7-genes signatures. A-F. Heatmaps of functionally methylated genes in TCGA COAD dataset (A-B and D-E) and in in-house first-line FOLFOX (C) or FOLFIRI (F) datasets. A and D. Differential gene expression profiles in TCGA COAD. B and E. Differential methylation profiles in TCGA COAD dataset. C and F. Differential methylation profiles in in-house FOLFOX or FOLFIRI cohorts. HC, hierarchical clustering; DR, drug response.

Figure 3

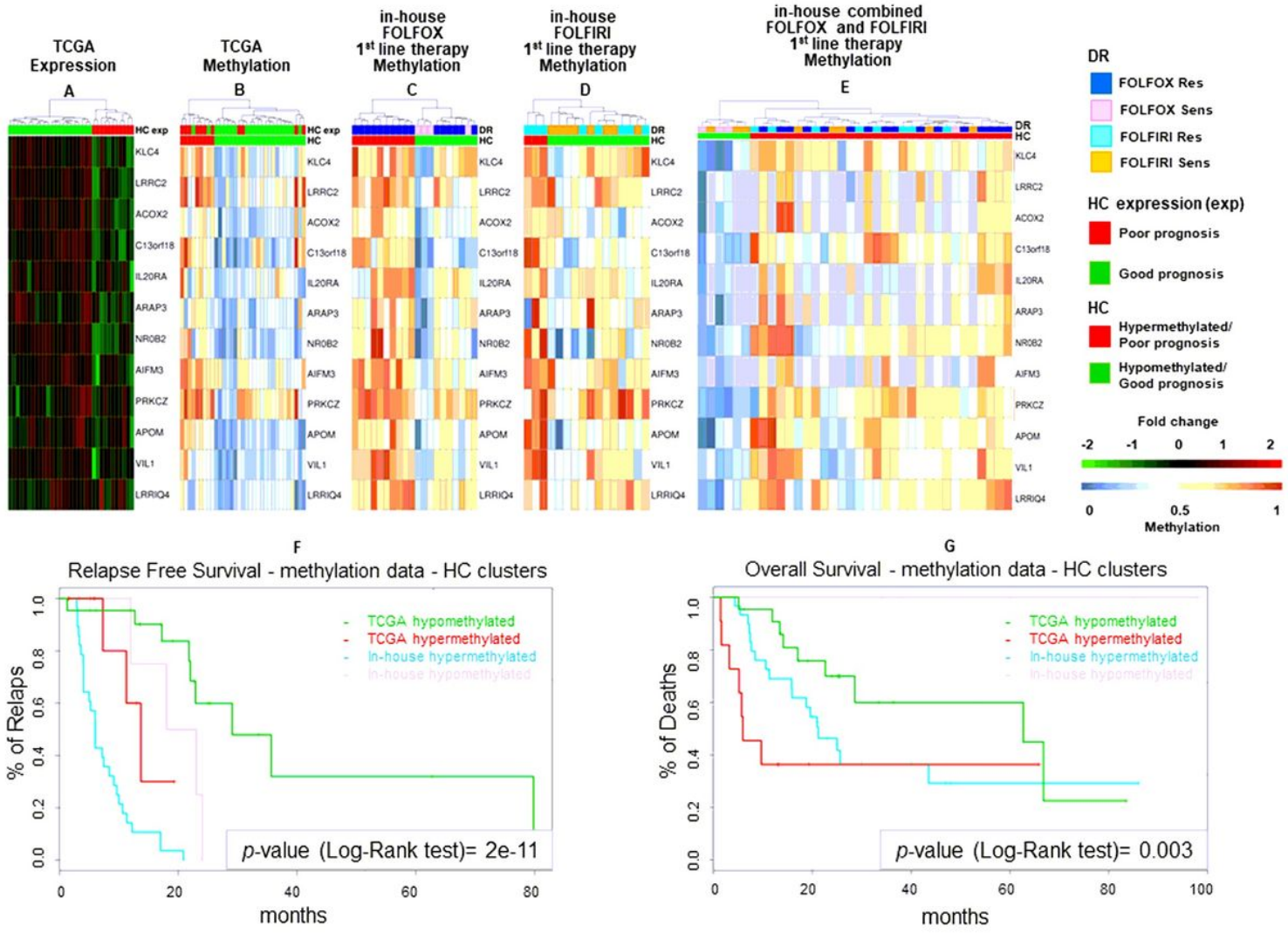

Figure 3

Hierarchical clustering and Kaplan-Meyer survival curves of mCRCs according to the 12-genes signature. A-E. Heatmaps of fMET genes in COAD TCGA dataset (A-B) and in 1st-line FOLFOX (C), FOLFIRI (D) or combined FOLFOX/FOLFIRI (E) in-house datasets. A-B. Differential gene expression (A) and methylation (B) profiles in TCGA COAD. C-E. Methylation profiles in in-house cohorts. F-G. Relapse free (F) and overall (G) survival curves according to TCGA COAD or in-house clusters, as reported in A, B and E. HC, hierarchical clustering; DR, drug response.

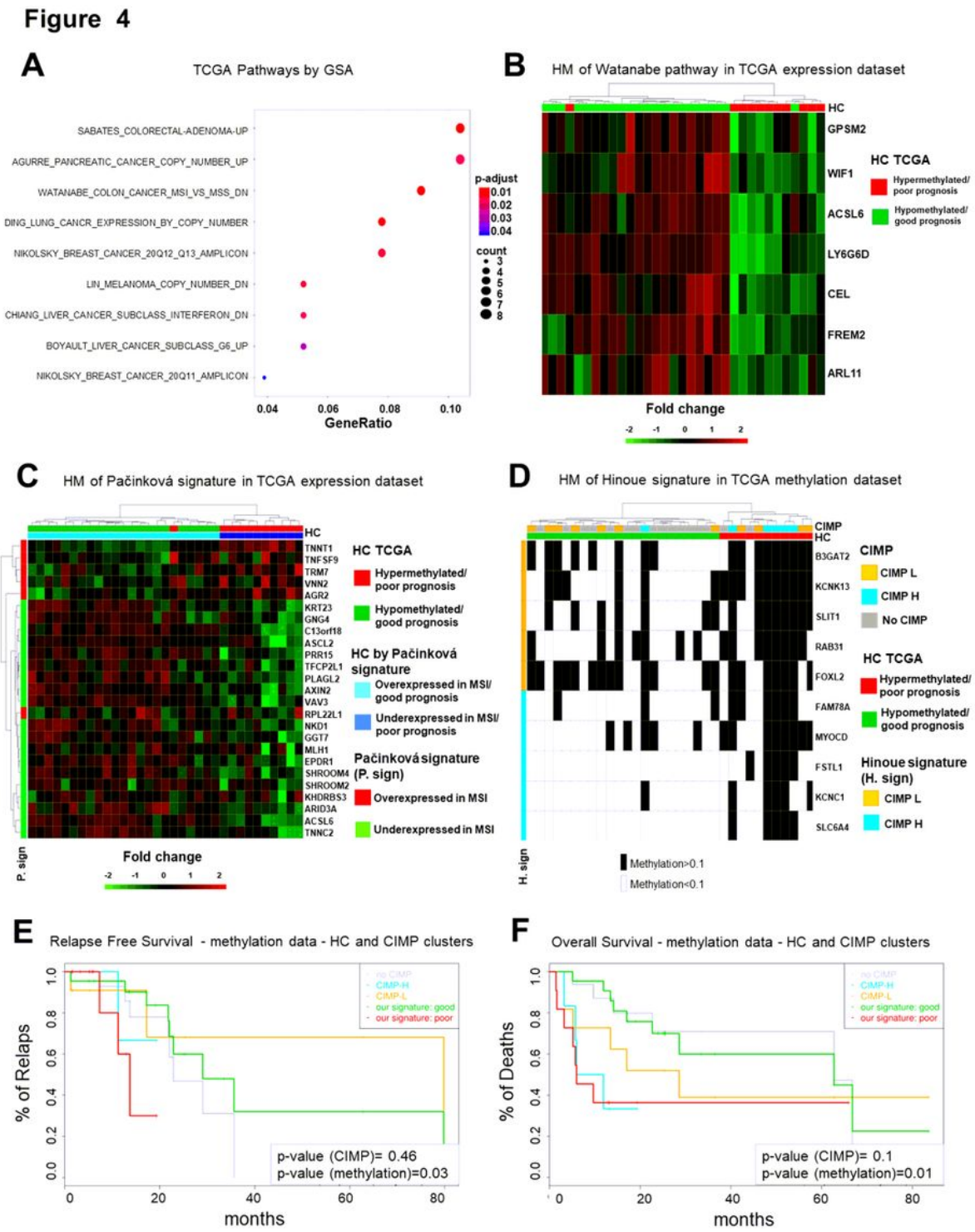

Figure 4

The characterization of the poor prognosis hypermethylated cluster highlights a MSI-like/CIMP-H phenotype. A. Significant enrichments for signaling pathways upon Gene Set Analysis (GSA). B. Heatmap (HM) of differentially expressed genes enriching Watanabe gene set in 33 mCRCs from TCGA COAD dataset. C. Heatmap of differentially expressed genes from the MSI-like gene expression Pačínková signature in 33 mCRCs from COAD TCGA database. D. Heatmap of CIMP status in 33 mCRCs from TCGA

database according to the 12-genes signature. CIMP status is labeled in in black. E-F. Relapse free and overall survival curves of TCGA COAD patients according to the 12-genes signature or CIMP status. HC, hierarchical clustering; DR, drug response.

Figure 5

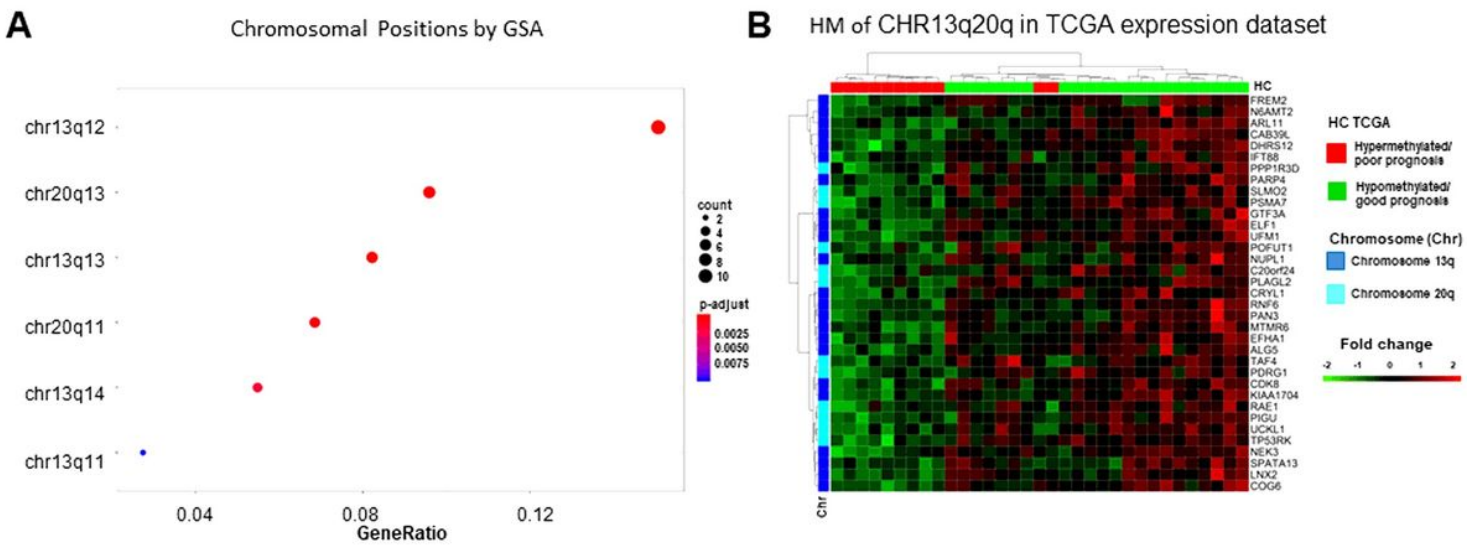

Figure 5

The poor prognosis hypermethylated cluster is enriched of genes located on arm q of chromosomes 13 and 20. A. Significant enrichments for the genomic positional collections upon GSA analysis. B. Heatmap (HM) of differentially expressed genes enriching chromosome 13 arm q and chromosome 20 arm q gene sets in 33 mCRCs from TCGA COAD dataset.

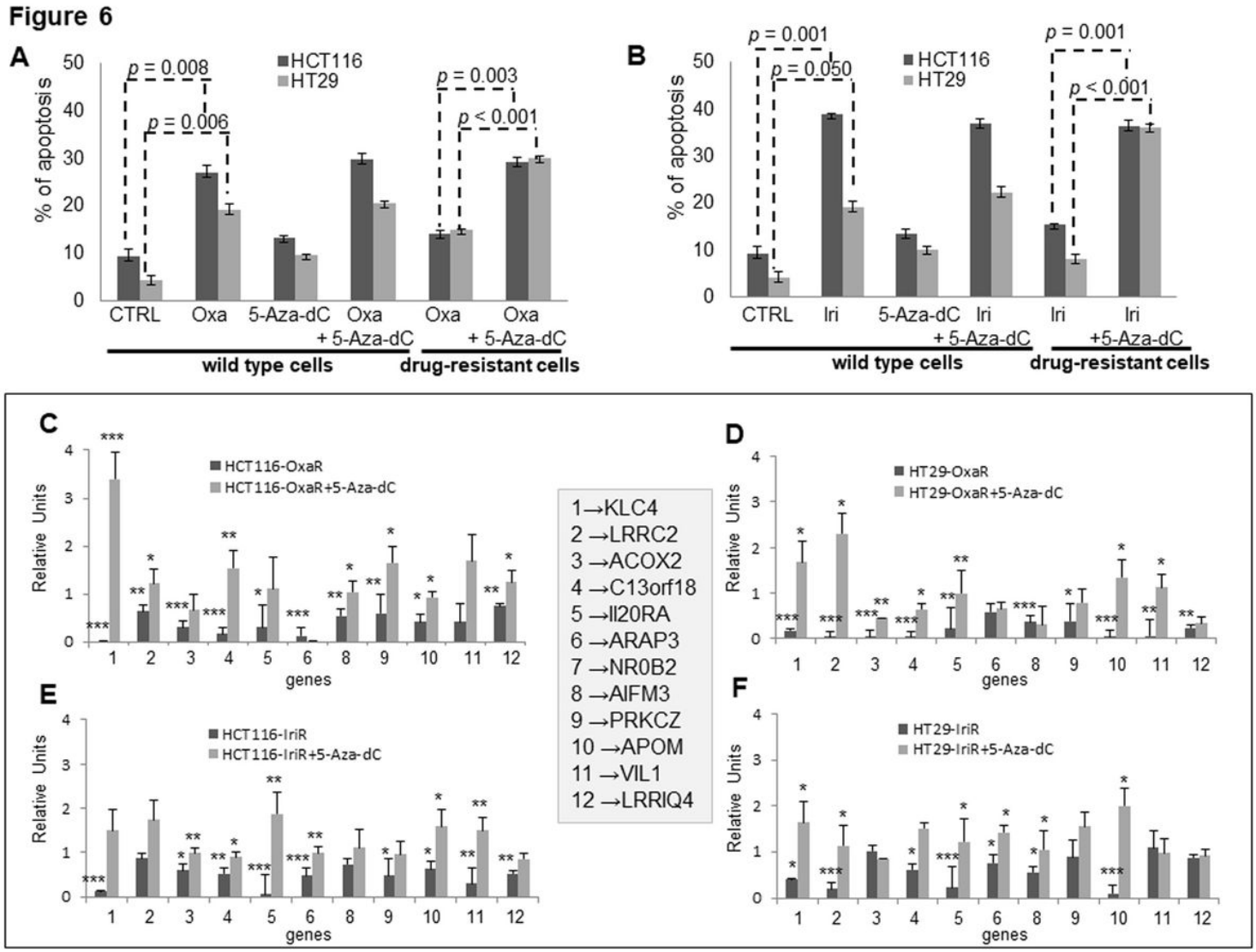

Figure 6

Validation of the 12-genes signature in drug-resistant CRC cell lines. A-B. Apoptotic cell death in HCT116 and HT29 drug-sensitive and drug-resistant CRC cell lines exposed to 10 uM 5-Aza-dC for 48h and/or 3 uM oxaliplatin (Oxa) (A) or 2uM irinotecan (Iri) (B) for 24h. C-F. Real time differential expression analysis of 11 genes belonging to the prognostic signature between drug-resistant and drug-sensitive CRC cell lines before and after exposure to 10 uM 5-Aza-dC for 48h. C. HCT116-OxaR: D. HT29-OxaR; E. HCT115-IriR; F, HT29-IriR. Significantly modulated genes are indicated by asterisks: *= $p < 0.05$; **= $p < 0.01$; ***= $p < 0.001$. Apoptosis and PCR analyses were performed in triplicate.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalDataset1.xlsx
- AdditionalTables.pdf

- AdditionalFigures.pdf
- AdditionalData.docx