Fadda A, Gentilini D, Moi L, Barault L, Leoni VP, Sulas P, Zorcolo L, Restivo A, Cabras F, Fortunato F, Zavattari C, Varesco L, Gismondi V, De Miglio MR, Scanu AM, Colombi F, Lombardi P, Sarotto I, Loi E, Leone F, Giordano S, Di Nicolantonio F, Columbano A, Zavattari P.

Colorectal cancer early methylation alterations affect the crosstalk between cell and surrounding environment, tracing a biomarker signature specific for this tumor.

Int J Cancer. 2018 Mar 15. doi: 10.1002/ijc.31380.

Colorectal cancer early methylation alterations affect the crosstalk between cell and surrounding environment, tracing a biomarker signature specific for this tumor

Antonio Fadda[1]*, Davide Gentilini[2,3]*, Loredana Moi[1]*, Ludovic Barault[4,5], Vera Piera Leoni[6], Pia Sulas[6], Luigi Zorcolo[7], Angelo Restivo[7], Francesco Cabras[7], Federica Fortunato[7], Cesare Zavattari[8], Liliana Varesco[9], Viviana Gismondi[9], Maria Rosaria De Miglio[10], Antonio Mario Scanu[10], Federica Colombi[4], Pasquale Lombardi[4], Ivana Sarotto[5], Eleonora Loi[1], Francesco Leone[4,5], Silvia Giordano[4,5], Federica Di Nicolantonio[4,5], Amedeo Columbano[6], Patrizia Zavattari[1#]

[1] Unit of Biology and Genetics, Department of Biomedical Sciences, University of Cagliari, Italy
[2] Department of Brain and Behavioral Sciences, University of Pavia, Italy
[3] Bioinformatics and Statistical Genomics Unit, Istituto Auxologico Italiano IRCCS, Cusano Milanino, Milan, Italy
[4] Department of Oncology, University of Torino, Italy
[5] Candiolo Cancer Institute-FPO, IRCCS, Italy
[6] Unit of Oncology and Molecular Pathology, Department of Biomedical Sciences, University of Cagliari, Italy
[7] Department of Surgery, Colorectal Surgery Center, University of Cagliari, Italy
[8] Independent Researcher, Machine Learning, Lucca, Italy
[9] Unit of Hereditary Cancer, IRCCS AOU San Martino-IST, Genoa, Italy
[10] Department of Clinical and Experimental Medicine, University of Sassari, Italy

* These authors equally contributed to this work
[#] Corresponding author

Running title: CRC early methylation biomarkers impair signalling crosstalk

Correspondence:
Patrizia Zavattari, PhD
Department of Biomedical Sciences, Unit of Biology and Genetics, University of Cagliari
Cittadella Universitaria di Monserrato
SP 8, Km 0.700 - 09042, Monserrato,
Cagliari, Italy
Phone: +39-070-675-4101 / Fax: +39-070-675-4100
email: pzavattari@unica.it

The authors declare no potential conflicts of interest.

PZ, FDN, AC have a patent 102017000072650 pending.

Arrays raw data are stored in our Dept. server and will be available on request

Novelty and Impact: We identified and validated a panel of 74 altered CpG islands, able to discriminate CRCs and adenomas from peritumoral and normal mucosa, with very high specificity (100%) and sensitivity (99.9%); three selected markers were tested and detected through non-invasive techniques, both in cfDNA and in stool DNA. Our results also demonstrated that the earliest methylation alterations affect genes coding for proteins involved in the crosstalk between cell and surrounding environment

Word count: 4088

Abstract

Colorectal cancer (CRC) develops through the accumulation of both genetic and epigenetic alterations. However, while the former are already used as prognostic and predictive biomarkers, the latter are less well characterized. Here, performing global methylation analysis on both CRCs and adenomas by Illumina Infinium HumanMethylation450 Bead Chips, we identified a panel of 74 altered *CpG islands*, demonstrating that the earliest methylation alterations affect genes coding for proteins involved in the crosstalk between cell and surrounding environment. The panel discriminates CRCs and adenomas from peritumoral and normal mucosa with very high specificity (100%) and sensitivity (99.9%). Interestingly, over 70% of the hypermethylated islands resulted in downregulation of gene expression. To establish the possible usefulness of these non-invasive markers for detection of colon cancer, we selected three biomarkers and identified the presence of altered methylation in stool DNA and plasma cell-free circulating DNA from CRC patients.

Introduction

Colorectal cancer (CRC), the third most prevalent cancer in the world [1], develops over years or decades, as a result of the accumulation of genetic and epigenetic alterations. At least three distinct types have been described: 1) chromosomal instability (CIN), the most frequent (80-85%), is characterized by aneuploidy, chromosomal gains/losses and accumulation of mutations in specific oncogenes (e.g. *KRAS*, *PIK3CA*, *BRAF*, etc) and tumor suppressors genes (e.g. *APC*, *SMAD4*, *TP53*); 2) microsatellite instability (MSI), caused by dysfunction of DNA mismatch repair genes, 3) CpG island methylation phenotype (CIMP), characterized by a widespread CpG island methylation, accounting for about 20-40% of CRCs. These types are not mutually exclusive, with tumors exhibiting features of more than one [2]. Although CIN tumors are characterized by mutations in oncogenes and tumor suppressor genes, epigenetic changes may also play a relevant role in neoplastic transformation even at the earliest stages [3-5]. Better knowledge of these epigenetic alterations could also help clarify which cellular processes are affected by such early changes and drive future functional studies.

Genes altered by methylation not only might represent potential biomarkers for early detection of CRC [6], but could also be important prognostic and predictive markers to improve therapeutic interventions [7,8]. Indeed, therapies based on epigenetic reprogramming have been recently tested in several types of tumors [9,10]. Moreover, since the screening of epigenetic alterations can also be carried out in circulating tumor DNA (ctDNA) and in DNA isolated from faeces of patients at risk of CRC, the detection of these biomarkers through less invasive procedures makes their identification of great value [8].

The aims of the present study were: i) to identify signature alterations in the CRC methylome through over 485k CpG loci of CRCs and peritumoral tissues; ii) to test whether these alterations represent early events in CRC development; iii) to explore the use of non-invasive techniques (stool and ctDNA) to reveal altered methylation. We identified the most altered pathways since early stages of CRC tumorigenesis and defined a panel of CpG islands differentially methylated as

biomarkers capable to detect CRC even at its earliest stages. This panel was successfully cross-validated *in silico* using methylation data from more than 500 CRC samples. Selected biomarkers were then confirmed in ctDNAs and in DNAs from stool samples of CRC patients.

Materials and methods

Samples collection and processing

Samples for whole genome methylation analysis

The methylome analysis was first performed in 18 pairs of primary CRCs among which four had matched peritumoral samples, taken at a distance of more than 10 cm from the neoplastic tissue (selected to represent the four anatomic region affected by cancer: left, right, sigmoid colon, rectum). Specimens came from patients diagnosed with CRC and operated at the Department of General Surgery, University of Cagliari (Italy) (**Figure 1** and **Supporting Information Table 1**). For this study, cases with familial adenomatous polyposis or human nonpolyposis colorectal cancer were excluded.

In a second step, methylome analysis was conducted in 21 adenomas and three matched intestinal mucosa controls, from 21 patients bearing an adenoma (**Figure 1** and **Supporting Information Table 2**). Lesions were removed during endoscopy [11]. DNA samples were collected at the National Institute for Cancer Research of Genoa (Italy).

Stool samples for methyl-BEAMing analyses

Stool samples were collected from 24 patients with colorectal cancer and taken intraoperatively from the bowel resection specimen. All stools samples were immediately frozen after collection and stored at −80°C until being processed. DNA extraction was performed using the QIAamp DNA Stool Mini Kit according to the manufacturer's instructions. All samples were collected at Department of Surgical Sciences, University of Cagliari and Department of Clinical and experimental medicine, University of Sassari. (**Figure 1**)

Plasma samples for methylBEAMing analyses

Blood draws were collected from 45 cases of CRC enrolled at the medical oncology department of the Candiolo Cancer Institute-FPO, IRCCS (Torino, Italy) between November 2015 and April 2016. Twelve cases were under adjuvant therapy after surgical resection of their lesion and were considered with no evidence of disease (NED). Remaining cases (N=33) were metastatic CRC with different levels of tumor burden (**Figure 1**). Whole blood was processed within samples three hours after collection. Samples were centrifuged at 1600g for 10 minutes for phase separation. Plasma was collected and submitted to a second centrifugation step at 3000g for 10 minutes to remove platelets and cell debris. Upper phase was collected, aliquoted and stored at -80°C until further processing. One milliliter of plasma was processed for DNA extraction using the Maxwell® RSC ccfDNA Plasma Kit (Promega) using 100µl for the elution volume.

Samples for Real-Time qRT-PCR validation

A total of 26 RNA samples (eight CRC with matched peritumoral tissues and ten individual CRCs), extracted as described in the previous paragraph, were tested by qRT-PCR (**Figure 1**).

DNA extraction and bisulfite conversion from tissue samples

Genomic DNA was extracted from tumoral and peritumoral tissue using the DNeasy Blood & Tissue Kit (Qiagen) and bisulfite converted using EZ DNA Methylation Gold Kit ™ (Zymo Research) according to the manufacturer's instructions. Quality control and quantification of DNA were performed before and after bisulfite conversion. Further details are available on supplementary materials.

DNA methylation assay

Four microliters of bisulfite-converted DNA were used for hybridization on Infinium HumanMethylation 450 BeadChip (450K), following the Illumina Infinium HD Methylation protocol. Data were acquired on an Illumina HiScan SQ scanner. Image intensities were extracted using GenomeStudio (2010.3).

MethylBEAMing analysis

The primers used for pyrosequencing were coupled with Tag sequence as previously described [9]. Two microliters of bisulfite converted cfDNA was amplified in replicate and processed following the same protocol previously described [10]. (**Supporting Information Table 3**).

Purified beads were run on a BD Accuri C6 (Becton-Dickinson), methylation percentage was expressed as the number of events in the methylated gate divided by the sum of events in methylated and unmethylated gates multiply by 100.

Markers selection for validation of locus specific methylation alterations

We selected three CpG islands, based on both a large methylation Δ between tumor and non-tumor tissue and the feasibility of the assay. Thirteen regions belonging to the biomarker panel identified were selected for locus specific investigation. Since cell free circulating and stool DNA is known to be composed of short fragments [12], significant probes which were distant from a maximum of 150bp (defining 23 loci) were selected to assess their potential to discriminate tumors from normal tissue in The Cancer Genome Atlas (TCGA) COREAD cohort using ROC analyses. Ten loci demonstrated a positive predictive value of 1 and a negative predictive value of at least 0.5. Among those, three regions allowed methylation independent amplification assay with highly specific probe design (*GRIA4*, *SLC8A1*, *SYN3*) (**Supporting Information Table 4**) and were firstly validated in a second data set of 78 tumoral and matched peritumoral tissue samples by pyrosequencing analysis (see Supplementary Materials and methods).

Statistical Analysis

Power calculation

The power was estimated on the calculation of the parameters of interest using the GSE48684 dataset. Based on this preliminary data the number of samples was calculated assuming an effect size that consider a differential methylation level of at least 10%, using a type I error of 10e-8 (which takes into account the need to correct for multiple test) and a statistical power of 0.8. Using these criteria, the number of samples to be analyzed should be 30.

Methylome data management

Illumina Methylation 450K raw data were analysed using the RnBeads analysis software package, as previously described [13,14]. In addition to CpG sites, four sets of genomic regions were covered in the analysis (tiling, genes, promoters and CpG Islands). Corrected p values (Benjamini & Hochberg) were computed as previously described [13,14]. The selected differently methylated

CpG loci were annotated, by referring to the 450k manifest, in order to obtain a gene list based on HGNC database. The beta values obtained for each sample at each CpG locus analyzed, were used in an analysis of Unsupervised Hierarchical Clustering (UHC), that, comparing locus to locus, allows to trace a tree (called dendrogram) whose closest branches indicate the most similarity in the methylation pattern, and vice versa. Finally, beta values were also visualized by building a HeatMap.

Biomarker specificity and sensitivity evaluation

We evaluated in three datasets (discovery set, GSE48684, TCGA COREAD) the tumoral and non-tumoral beta value distribution relative to the 74 biomarkers. The combined p-value of each biomarker was obtained using the Fisher's method, by combining the p-value of each CpG site belonging to a particular CpG island, calculated using the Wilcoxon test.

To evaluate the association of each biomarker with a binary outcome a logistic regression model was fitted to the individual biomarker data and the new probabilities used to calculate the area under the curve (AUC), specificity and sensitivity values using the "*OptimalCutpoints*" package [15].

To evaluate the ability of the entire biomarkers panel to correctly classify between two possible conditions (i.e tumoral *vs*. non-tumoral), we built a support vector machine learning model (SVM), using the GSE48684 dataset as a training set and the TCGA COREAD as test set and evaluate the prediction performance ability in the test set calculating the confusion matrix using the "*caret*" package.

Bioinformatic analysis

Functional annotations of the differently methylated loci, were conducted by the ToppGene package [16], from which are derived the information on the pathways potentially involved. The significance value was corrected for multiple comparisons (Bonferroni or Benjamini & Hochberg).

Three cross-validation datasets were retrieved from the database NCBI Gene Expression Omnibus (GEO) portal (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE48684, GSE52270, GSE53051. Processed data were used for all datasets above mentioned. TCGA COREAD cohort is available online at RnBeads website under Methylome Resources (http://rnbeads.mpi-inf.mpg.de/methylomes.php). For each of these datasets the mean methylation value for each CpG

island of interest has been calculated and visualized using Bioconductor package "*ComplexHeatmap*" [17].

Ethical approval

All the biological samples analysed were obtained with written informed consent signed from patients and ethical approval granted by the relative Ethics Committee: CRC tissue samples by the Independent Ethics Committee of the University Hospital of Cagliari; stool samples by the Independent Ethics Committee of the University Hospital of Cagliari and by Comitato Bioetico della Azienda Sanitaria Locale di Sassari; Plasma samples by the Ethical Committee of the IRCCS Cancer Center of Candiolo.

Results

CRC methylome alterations are early events in colon carcinogenesis: identification of an early epigenetic biomarkers panel

The differential methylation analysis conducted on 18 CRCs and 4 peritumoral samples (**Figure 1** and **Supporting Information Table 1**), powerfully discriminated tumor samples from peritumoral tissues. We identified 22307 CpG loci differentially methylated with an adjusted p-value threshold of 0.05 (Benjamini & Hochberg multiple test correction). To test whether these alterations represent an early event in carcinogenesis, we conducted a differential methylation analysis on 21 adenomas *vs.* 3 control mucosae (**Figure 1** and **Supporting Information Table 2**). As expected, the results of the analysis in adenomas were less robust than in CRCs. Using the nominal threshold (p-value <0.05), 43999 CpG loci resulted differently methylated. In both CRCs and adenomas, a wide hypomethylation of the tumoral samples was found by analyzing the entire genome, divided into portions (*tiling*), however, by restricting the analysis to the regulatory regions, a switch towards hypermethylation, especially in the CpG islands, was evident (**Figure 2A and 2B**).

The list of genes whose associated CpG islands, were significantly altered (875 in CRCs and 2393 in adenomas), was subjected to a gene enrichment and candidate gene prioritization analysis by Toppgene, allowing the identification of the pathways most affected by aberrant methylation: Wnt signaling, Neuronal System, Cadherin signaling, Transmission across Chemical Synapses, Neuroactive ligand-receptor interaction, Neurotransmitter Release Cycle, GABAergic synapse,

Core extracellular matrix, Calcium signaling, Cholinergic synapse (**Supporting Information Table 5**). For the list of the 171 genes found modified in CRCs and 432 in adenomas belonging to 10 significantly involved pathways, see **Supporting Information Table 6**. Notably, Toppgene results show that the most affected pathways were largely comparable between CRCs and adenomas (**Figure 2C**).

Remarkably, most of the genes whose associated CpG islands were altered in CRCs, were "already" aberrant in adenomas. In **Figure 3** and **Supporting Information Table 7** are listed 74 CpG islands resulting from the comparison, with the respective average beta values and Δs calculated between the average value in the tested samples (CRCs or adenomas) and in the respective controls. Interestingly, only two islands, at *GNG7* and *GRIN3B* genes, underwent hypomethylation in both adenomas and CRCs. Of note, the Δ average beta value was always higher in CRCs than in adenomas (except for *EDNRB*), probably due to a greater methylation heterogeneity among adenomas. Surprisingly, none of the genes listed in **Figure 3** resulted dysregulated by the transcriptome analysis. However, all those genes displayed an extremely low level of expression, close to the background intensity level. Therefore, an investigation by qRT-PCR was undertaken (**Supporting Information Figure 1**). UHC was conducted on the CRCs and adenomas (**Figure 3**) discovery set, based on the 74 biomarkers panel. A multi-cluster gene functional enrichment analysis conducted only for the genes associated to the 74 CpG islands, showed a significant enrichment of proteins involved in the crosstalk between cell and surrounding environment (see **Figure 3** and **Figure 4**).

In silico validation of the CRC early biomarkers panel: specificity for primary tumors and metastases

To validate and increase the statistical power of our discovery dataset results, first, we cross-validated *in silico* our set of markers, as predictors of cancer and adenoma, using methylation data from two databases: I) TCGA COREAD cohort [18] (**Figure 5A**); II) GSE48684 [3](**Figure 5B**). The 74 biomarker panel was able to separate most carcinomas and adenomas from normal mucosa. ROC analysis was performed for each marker using TCGA COREAD cohort. The specificity of many markers was equal to 1, i.e. 100% (ranging from 0.89 and 1), the sensitivity was ≥ 0.9 in over 70% of the islands (ranging from 0.7 and 0.97) (**Supporting Information Table 7** and **Supporting Information Table 8**). The robustness of the identified biomarkers was supported by p-values lower than the threshold for claiming genome-wide significance ($P < 10^{-8}$). We also evaluated the

specificity, sensitivity, AUC and accuracy for the entire panel, obtaining respectively: SP = 1; SE = 0.9942; AUC = 1; ACC = 0.9971.

To verify if the panel is specific for CRC, we examined the GSE52270 data set [19]. The panel resulted specific for CRC, but not for the other analyzed cancer types (**Figure 5C**).

Interestingly, the UHC of our CpG islands panel in a fourth data set (GSE53051) [20], assigned colon cancer metastases on the same branches of CRC, with the exception of two cases (**Figure 5D**). These results suggest that these alterations are not counter-selected during carcinogenesis and are maintained in metastases.

Methylation analysis in stool DNA samples

Based on the consistency of the methylation alterations identified in the tumor tissues, we wondered whether these methylation patterns could be detected in DNA extracted from stools. The goal was to set up in CRC patient stools the experimental test as a possible non-invasive technique to detect early/asymptomatic CRCs. We selected three CpG islands, associated to *GRIA4*, *SLC8A1* and *SYN3* genes, based on both a large methylation Δ between tumor and non-tumor tissue and the feasibility of the assay experimental design (**Supporting Information Table 4**). To assess the reproducibility of the three selected markers, we performed a methylation analysis by pyrosequencing in a second data set of 78 tumoral and 78 matched peritumoral samples (**Figure 1**). All three selected islands, associated to *GRIA4*, *SLC8A1* and *SYN3* genes, were significantly hypermethylated in tumor *vs*. peritumoral samples (**Supporting Information Figure 2**). Methylation of the same CpG islands, was thus assessed, by digital PCR analyses, in DNA isolated from CRC patients' stool samples, taken at the time of surgical resection. As shown in **Figure 6A**, all except three tested samples (87.5%) showed more than 1% of methylation for at least one of the three markers. In particular, 79.2% of samples showed more than 1% of methylation at *GRIA4* (average 21%); 70.8% at *SLC8A1* (average 10%); 62.5% at *SYN3* (average 13%).

Methylation analysis in ctDNA

Tumor fragments of DNA in the blood stream can be used as a surrogate sample to tissue biopsy. Cell free circulating DNA is a mixture of nucleic acids from normal (mainly from leukocytes) and from tumor tissues (ctDNA) [21] Liquid biopsy test using genetic or epigenetic alterations in plasma DNA has been proposed for early diagnosis [6] and early detection of relapse [22-24]. Methylation of *GRIA4*, *SLC8A1* and *SYN3* was assessed in circulating DNA of CRC patients. Assays were successful in at least one replicate for 43, 45 and 41 cases for *GRIA4*, *SLC8A1* and *SYN3*, respectively. Some patients were receiving adjuvant therapy post-surgical resection of their

tumor with no radiologic evidence of disease (NED) and were expected to display very low or no ctDNA. The aim of these exploratory analyses was to compare methylation levels at the three loci between NED and patients still bearing a lesion. Median methylation at *GRIA4* and *SLC8A1* was significantly different (u-test, *GRIA4*: p-value =0.029; *SLC8A1*: p-value =0.024). *SYN3* did not show any difference, possibly due to low number of cases which displayed methylation. The positivity exclusively seen in patients with lesions demonstrates the specificity of the selected markers.

Among clinical features (**Supporting Information Table 9**), CEA levels, a surrogate of tumor burden, could discriminate two subgroups among the patients still bearing a lesion. According to the literature, we divided the samples into three groups: NED, CEA-low and CEA-high, using a threshold of 5 ng/ml; indeed, CEA elevation over 5 ng/ml was found to be a very accurate marker of recurrence [25]. As shown in **Figure 6B**, dividing the samples in the three groups, did not result in significant differences between NED and CEA-low, relatively to the observed methylation levels, while significantly higher methylation levels (p-value = 0.00223 and p-value = 0.01972, respectively) were evident for *GRIA4* and *SLC8A1* in the CEA-high group. Of note, patients with low CEA whose primary tumor was mutated for *RAS* or *BRAF*, did not display the genetic alteration in ctDNA (**Figure 6C**). Interestingly, the three NED subjects with high levels of CEA, did not show hypermethylation in ctDNA.

Discussion

The genetic and epigenetic landscape of CRC has been extensively studied so far. However, current technological advances in DNA methylation analysis have enabled the identification and validation of new biomarkers in a more unbiased way. Nevertheless, the concept of CIMP phenotype remains debated [26], mostly due to lack of consensus in its evaluation warranting unbiased genome-wide approaches to characterize epigenetic changes involved in CRC onset and development [4]. We observed limited correlation between CIMP phenotype and hypermethylation of the CpG islands included in our panel, perhaps because we selected loci as being the most informative on the early

alterations, or perhaps because CIMP was classically defined on methylation alterations discovered before the advent of genomic approaches to assess global methylome.

Although several studies have been conducted analysing methylome alterations in CRCs, there is still urgent need for biomarker discovery. In our work, we identified a panel of 74 CpG islands aberrantly methylated not only in advanced stages of disease, but already detectable in adenomas, when the tumor is often asymptomatic. The performance ability of the panel was cross-validated *in silico* by analysing it in hundreds of samples, including CRCs, adenomas, normal counterparts and other tumor types. The identified panel appears very robust and informative (sensitivity 99.99%), specific for CRC (specificity 100%) from early to metastatic stages. Multiple studies have investigated the use of single or combined DNA methylation-based biomarkers for diagnostic purposes [8]. The performance ability of *SEPT9* and *VIM*, the only two commercially available methylation biomarkers, to identify CRC/adenomas greatly varies depending on the experimental design. However, combining both the markers into a panel could improve the diagnostic accuracy and achieve higher clinical sensitivity. Not only our results are consistent with this observation but the performance ability of most of our biomarkers outperforms that of commercially available ones.

Importantly, this set of markers could be detected in CRC patients even through non-invasive techniques, as we identified hypermethylated tumoral DNA both in stool and plasma samples of CRC patients. Overall, the panel shows a good diagnostic and prognostic value even in the non-invasive assessment, strengthening its potential value in screening and follow-up of CRC patients. Previously, three selected markers (*AGTR1*, *WNT2*, *SLIT2*) were validated in stool DNA samples from CRC patients, showing a sensitivity of 21%, 40% and 52% respectively [27]. A panel of these genes reached a sensitivity of 78%, based on the criteria that at least one of the genes was methylated. In comparison, our three selected biomarkers performed better in terms of sensitivity, with a percentage of samples showing more than 1% of methylation, ranging from 62.5% to 79.2%. Based on a similar criterion, the panel of our selected markers guaranteed an overall sensitivity of 87.5%. We conclude that, even in these samples, the tested CpG islands resulted excellent tumor markers despite the technical difficulties related to a biological challenging matrix, such as stool. Nevertheless, we acknowledge the exploratory nature of our analyses in a limited stool dataset from patients who had already received a CRC diagnosis. Future studies are warranted to assess the sensitivity and specificity of the loci identified in this work in stool DNA from asymptomatic individuals and early stage disease.

Concerning cell free DNA samples, two markers (*GRIA4* and *SLC8A1*), were able to distinguish NED patients from metastatic and CEA-low from CEA-high. Indeed, methylation alterations displayed a negative predictive value higher than CEA (as demonstrated by the three NED patients with elevated CEA levels but not hypermethylated DNA). The absence of alterations observed in the low-CEA subgroup might be due to specific biologic features of these tumors, impairing the release of circulating markers (e.g.: well differentiated status of neoplastic cells, absence of vascular network, etc.). On the other hand, the limited sensitivity of the techniques employed to assess methylation could have yielded false negative results. Indeed, we acknowledge that the current digital PCR technique may have limited sensitivity for cases with low tumor burden, since the assays was designed for monitoring purposes (not early detection) privileging highly linear quantification, instead of high sensitivity [28]. Further assay improvements are therefore warranted for the early detection setting. Moreover, studies with larger number of patient plasma samples and controls from healthy individuals are required to establish and exploit the potential of our markers in liquid biopsies.

In addition to identifying potentially new powerful biomarkers (only partially already correlated to cancer, see **Supporting Information Table 10**), in CRC early diagnosis and traceability of minimal residual disease, this study has highlighted the biological processes mainly affected by early methylation alterations in colon carcinogenesis. In fact, as shown in **Figure 4**, the crosstalk between tumor cells and surrounding environment resulted particularly involved, in terms of membrane receptors, solute transporters and cell adhesion molecules. Functional annotation analysis has highlighted the enrichment of protocadherins [29-33], integrins [34], members of the solute carrier family [35-40], and, extensively, G-protein coupled receptors involved in the transduction of neuroactive signals [41,42].

Silencing of genes coding for protocadherins (PCDHs) by means of promoter hypermethylation, such as PCDH10 in different carcinomas [29,30] and PCDH8 in breast cancer [31], has been shown in several cancer types. In particular, it has been demonstrated that a cluster of PCDHs located on chromosome 5q31 undergoes a mechanism of long range epigenetic silencing (LRES) by hypermethylation [32]. PCDH LRES was described in Wilms tumor [32], breast [43] prostate [44] and colorectal cancer [45], suggesting a tumor suppressor role of this protocadherin cluster in several types of cancer. As other LRES loci have been identified in breast [43], head and neck [46], lung [46], prostate [33] and colon cancer [47], silencing of tumor suppressor genes by this mechanism to appears a emerges as a common feature in human cancer [33].

The list of genes reported in **Supporting Information Table 7** also highlights the involvement of several carriers, including many members of the solute carrier family (SLC). These genes, unlike PCDHs, are located on different chromosomes, but their common alteration suggests a combinatorial control, a mechanism well known in the regulation of gene expression. Indeed, methylation-mediated silencing of solute carriers was already reported in CRC, lung, prostate (SLC5A8) [35-37], breast and gastric cancers (SLC19A3) [38,39] and gliomas (SLC22A18) [40].

Functional studies will clarify the role of the epigenetic alterations identified in the present study, especially considering that over 70% of hypermethylated islands resulted in downstream downregulation **(Supporting Information Figure 3)**. Remarkably, almost all these molecules are located at membrane or extracellular matrix level (see **Figure 3**), therefore also potential optimal therapeutic targets.

Acknowledgements

We wish to thank Francesco and Manuela Zavattari for their skillful contribution to the drawing and graphic design of the graphical abstract.

Grants: This work was supported by grants from Fondazione Banco di Sardegna (2012) and Fondo per la Ricerca Locale (ex 60%), Università di Cagliari to PZ, Regione Autonoma della Sardegna (CRP-79303) to PZ and LZ. Partly supported also by grants AIRC IG n. 17707 (FDN); Fondo per la Ricerca Locale (ex 60%), Università di Torino, 2014 (FDN) and AIRC 2010 Special Program Molecular Clinical Oncology 5 per mille, project no. 9970 (FL).

Author contributions

AF: DNA extraction from tissue samples and bisulfite treatment, acquisition and analysis of qRT-PCR data, pyrosequencing, methylation data management and analyses, CIMP phenotype definition, transcriptome analysis, statistical analysis, bioinformatic analysis, analysis and interpretation of data, drafting of the manuscript.
DG: methylation data management and analyses, supervision on statistical analysis and bioinformatic analysis.
LM: DNA extraction from tissue and stool samples, bisulfite treatment, DNA whole genome methylation assay, pyrosequencing, RNA extraction, transcriptome assay, genetic mutations screening.
LB: DNA extraction from plasma samples, bisulfite treatment, markers selection, methylBEAMing assay and analysis on ctDNA from plasma samples and stool DNA, analysis and interpretation of ctDNA data, drafting of the manuscript.
VPL: DNA extraction from tissue samples and pyrosequencing analysis.

PS: support on transcriptome assay and analysis.
LZ, AR, FC, FF, DS: CRC/peritumoral tissues and stool samples collection, clinical data collection.
CZ: preliminary methylome data analysis.
LV, VG: adenomas/normal mucosa DNA samples and clinical data collection.
MRDM: DNA extraction from stool samples, interpretation of clinical data.
AMS: stool samples and clinical data collection.
FL, FC, PL: plasma samples and clinical data collection, interpretation of ctDNA data.
IS: immunohistochemistry
EL: Final revision of the manuscript and participation in the implementation of the requested changes.
SG, FDN, AC, PZ: study concept and design, analysis and interpretation of data, drafting of the manuscript.
FDN: supervision on methylBEAMing assay and analysis on ctDNA from plasma samples and stool DNA.
PZ: supervision on experiments conducted on tissues and faeces samples and related analyzes.
All Authors discussed the results and commented the manuscript.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66:7–30.
2. Goel A, Arnold CN, Niedzwiecki D, Chang DK, Ricciardiello L, Carethers JM, Dowell JM, Wasserman L, Compton C, Mayer RJ, Bertagnolli MM, Boland CR. Characterization of sporadic colon cancer by patterns of genomic instability. *Cancer Res* 2003;63:1608–14.
3. Luo Y, Wong C-J, Kaz AM, Dzieciatkowski S, Carter KT, Morris SM, Wang J, Willis JE, Makar KW, Ulrich CM, Lutterbaugh JD, Shrubsole MJ, et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. *Gastroenterology* 2014;147:418–29.e8.
4. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, Houlihan PS, Krouse RS, Prasad AR, Einspahr JG, Buckmeier J, Alberts DS, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. *J Natl Cancer Inst* 2005;97:1330–8.
5. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP. Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Res* 1998;58:5489–94.
6. Church TR, Wandell M, Lofton-Day C, Mongin SJ, Burger M, Payne SR, Castaños-Vélez E, Blumenstein BA, Rösch T, Osborn N, Snover D, Day RW, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut* 2014;63:317–25.
7. Phipps AI, Limburg PJ, Baron JA, Burnett-Hartman AN, Weisenberger DJ, Laird PW, Sinicrope FA, Rosty C, Buchanan DD, Potter JD, Newcomb PA. Association between molecular subtypes of colorectal cancer and patient survival. *Gastroenterology* 2015;148:77–87.e2.

8. Lam K, Pan K, Linnekamp JF, Medema JP, Kandimalla R. DNA methylation based biomarkers in colorectal cancer: A systematic review. *Biochim Biophys Acta* 2016;1866:106–20.
9. Diehl F, Schmidt K, Durkee KH, Moore KJ, Goodman SN, Shuber AP, Kinzler KW, Vogelstein B. Analysis of mutations in DNA isolated from plasma and stool of colorectal cancer patients. *Gastroenterology* 2008;135:489–98.
10. Barault L, Amatu A, Bleeker FE, Moutinho C, Falcomatà C, Fiano V, Cassingena A, Siravegna G, Milione M, Cassoni P, de Braud F, Rudà R, et al. Digital PCR quantification of MGMT methylation refines prediction of clinical benefit from alkylating agents in glioblastoma and metastatic colorectal cancer. *Ann Oncol* 2015;26:1994–9.
11. De Benedetti L, Sciallero S, Gismondi V, James R, Bafico A, Biticchi R, Masetti E, Bonelli L, Heouaine A, Picasso M. Association of APC gene mutations and histological characteristics of colorectal adenomas. *Cancer Res* 1994;54:3553–6.
12. Roth C, Kasimir-Bauer S, Pantel K, Schwarzenbach H. Screening for circulating nucleic acids and caspase activity in the peripheral blood as potential diagnostic tools in lung cancer. *Mol Oncol* 2011;5:281–91.
13. Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods* 2014;11:1138–40.
14. Gentilini D, Scala S, Gaudenzi G, Garagnani P, Capri M, Cescon M, Grazi GL, Bacalini MG, Pisoni S, Dicitore A, Circelli L, Santagata S, et al. Epigenome-wide association study in hepatocellular carcinoma: Identification of stochastic epigenetic mutations through an innovative statistical approach. *Oncotarget* 2017;8:41890–902.
15. López-Ratón M, Rodríguez-Álvarez MX, Suárez CC, Sampedro FG. OptimalCutpoints : An R Package for Selecting Optimal Cutpoints in Diagnostic Tests. *J Stat Softw* 2014;61:1–36.
16. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;37:W305-11.
17. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32:2847–9.
18. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487:330–7.
19. Heyn H, Vidal E, Ferreira HJ, Vizoso M, Sayols S, Gomez A, Moran S, Boque-Sastre R, Guil S, Martinez-Cardus A, Lin CY, Royo R, et al. Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. *Genome Biol* 2016;17:11.
20. Timp W, Bravo HC, McDonald OG, Goggins M, Umbricht C, Zeiger M, Feinberg AP, Irizarry R a. Large hypomethylated blocks as a universal defining epigenetic alteration in human solid tumors. *Genome Med* 2014;6:61.
21. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol* 2013;10:472–84.
22. Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, Silliman N, Tacey M, Wong H-L, Christie M, Kosmider S, Skinner I, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci Transl Med* 2016;8:346ra92.

23. Siravegna G, Mussolin B, Buscarino M, Corti G, Cassingena A, Crisafulli G, Ponzetti A, Cremolini C, Amatu A, Lauricella C, Lamba S, Hobor S, et al. Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. *Nat Med* 2015;21:795–801.
24. Garrigou S, Perkins G, Garlan F, Normand C, Didelot A, Le Corre D, Peyvandi S, Mulot C, Niarra R, Aucouturier P, Chatellier G, Nizard P, et al. A Study of Hypermethylated Circulating Tumor DNA as a Universal Colorectal Cancer Biomarker. *Clin Chem* 2016;62:1129–39.
25. Flamini E, Mercatali L, Nanni O, Calistri D, Nunziatini R, Zoli W, Rosetti P, Gardini N, Lattuneddu A, Verdecchia GM, Amadori D. Free DNA and carcinoembryonic antigen serum levels: an important combination for diagnosis of colorectal cancer. *Clin Cancer Res* 2006;12:6985–8.
26. Gallois C, Laurent-Puig P, Taieb J. Methylator phenotype in colorectal cancer: A prognostic factor or not? *Crit Rev Oncol Hematol* 2016;99:74–80.
27. Carmona FJ, Azuara D, Berenguer-Llergo A, Fernández AF, Biondo S, de Oca J, Rodriguez-Moranta F, Salazar R, Villanueva A, Fraga MF, Guardiola J, Capellá G, et al. DNA methylation biomarkers for noninvasive diagnosis of colorectal cancer. *Cancer Prev Res (Phila)* 2013;6:656–65.
28. Barault L, Amatu A, Siravegna G, Ponzetti A, Moran S, Cassingena A, Mussolin B, Falcomatà C, Binder AM, Cristiano C, Oddo D, Guarrera S, et al. Discovery of methylated circulating DNA biomarkers for comprehensive non-invasive monitoring of treatment response in metastatic colorectal cancer. *Gut* 2017 Oct 5.
29. Ying J, Li H, Seng TJ, Langford C, Srivastava G, Tsao SW, Putti T, Murray P, Chan a TC, Tao Q. Functional epigenetics identifies a protocadherin PCDH10 as a candidate tumor suppressor for nasopharyngeal, esophageal and multiple other carcinomas with frequent methylation. *Oncogene* 2006;25:1070–80.
30. Yu J, Cheng YY, Tao Q, Cheung KF, Lam CNY, Geng H, Tian L-W, Wong YP, Tong JHM, Ying J-M, Jin H, To KF, et al. Methylation of protocadherin 10, a novel tumor suppressor, is associated with poor prognosis in patients with gastric cancer. *Gastroenterology* 2009;136:640–51.e1.
31. Yu JS, Koujak S, Nagase S, Li C-M, Su T, Wang X, Keniry M, Memeo L, Rojtman A, Mansukhani M, Hibshoosh H, Tycko B, et al. PCDH8, the human homolog of PAPC, is a candidate tumor suppressor of breast cancer. *Oncogene* 2008;27:4657–65.
32. Dallosso AR, Hancock AL, Szemes M, Moorwood K, Chilukamarri L, Tsai H-H, Sarkar A, Barasch J, Vuononvirta R, Jones C, Pritchard-Jones K, Royer-Pokora B, et al. Frequent long-range epigenetic silencing of protocadherin gene clusters on chromosome 5q31 in Wilms' tumor. *PLoS Genet* 2009;5:e1000745.
33. Coolen MW, Stirzaker C, Song JZ, Statham AL, Kassir Z, Moreno CS, Young AN, Varma V, Speed TP, Cowley M, Lacaze P, Kaplan W, et al. Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. *Nat Cell Biol* 2010;12:235–46.
34. Kok-Sin T, Mokhtar NM, Ali Hassan NZ, Sagap I, Mohamed Rose I, Harun R, Jamal R. Identification of diagnostic markers in colorectal cancer via integrative epigenomics and genomics data. *Oncol Rep* 2015;34:22–32.
35. Valo S, Kaur S, Ristimäki A, Renkonen-Sinisalo L, Järvinen H, Mecklin J-P, Nyström M, Peltomäki P. DNA methylation changes in Lynch syndrome associated colorectal adenomas and carcinomas. *Cancer Res* 2015;75.

36. Ikeda K, Shiraishi K, Koga T, Motooka Y, Fujino K, Shibata H, Mori T, Suzuki M. Prognostic Significance of Aberrant Methylation of Solute Carrier Gene Family 5A8 in Lung Adenocarcinoma. *Ann Thorac Surg* 2015;99:1755–9.
37. Park JY, Zheng W, Kim D, Cheng JQ, Kumar N, Ahmad N, Pow-Sang J. Candidate tumor suppressor gene SLC5A8 is frequently down-regulated by promoter hypermethylation in prostate tumor. *Cancer Detect Prev* 2007;31:359–65.
38. Ng EKO, Leung CPH, Shin VY, Wong CLP, Ma ESK, Jin HC, Chu K-M, Kwong A. Quantitative analysis and diagnostic significance of methylated SLC19A3 DNA in the plasma of breast and gastric cancer patients. *PLoS One* 2011;6:e22233.
39. Liu X, Lam EKY, Wang X, Zhang J, Cheng YY, Lam YW, Ng EKO, Yu J, Chan FKL, Jin H, Sung JJY. Promoter hypermethylation mediates downregulation of thiamine receptor SLC19A3 in gastric cancer. *Tumour Biol* 2009;30:242–8.
40. Chu S-H, Ma Y-B, Feng D-F, Zhang H, Zhu Z-A, Li Z-Q, Jiang P-C. Correlation of low SLC22A18 expression with poor prognosis in patients with glioma. *J Clin Neurosci* 2012;19:95–8.
41. Naumov VA, Generozov E V., Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov S V., Alekseev M V., Shelygin YA, Govorun VM. Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. *Epigenetics* 2013;8:921–34.
42. Feigin ME. Harnessing the genome for characterization of G-protein coupled receptors in cancer pathogenesis. *FEBS J* 2013;280:4729–38.
43. Novak P, Jensen T, Oshiro MM, Watts GS, Kim CJ, Futscher BW. Agglomerative epigenetic aberrations are a common event in human breast cancer. *Cancer Res* 2008;68:8616–25.
44. Kobayashi Y, Absher DM, Gulzar ZG, Young SR, McKenney JK, Peehl DM, Brooks JD, Myers RM, Sherlock G. DNA methylation profiling reveals novel biomarkers and important roles for DNA methyltransferases in prostate cancer. *Genome Res* 2011;21:1017–27.
45. Dallosso a R, Øster B, Greenhough A, Thorsen K, Curry TJ, Owen C, Hancock a L, Szemes M, Paraskeva C, Frank M, Andersen CL, Malik K. Long-range epigenetic silencing of chromosome 5q31 protocadherins is involved in early and late stages of colorectal tumorigenesis through modulation of oncogenic pathways. *Oncogene* 2012;31:4409–19.
46. Smith LT, Lin M, Brena RM, Lang JC, Schuller DE, Otterson G a, Morrison CD, Smiraglia DJ, Plass C. Epigenetic regulation of the tumor suppressor gene TCF21 on 6q23-q24 in lung and head and neck cancer. *Proc Natl Acad Sci U S A* 2006;103:982–7.
47. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nat Genet* 2006;38:540–9.
48. De Benedetti L, Varesco L, Pellegata NS, Losi L, Gismondi V, Casarino L, Sciallero S, Bonelli L, Biticchi R, Bafico A. Genetic events in sporadic colorectal adenomas: K-ras and p53 heterozygous mutations are not sufficient for malignant progression. *Anticancer Res* 1993;13:667–70.
49. Berg M, Hagland HR, Søreide K. Comparison of CpG island methylator phenotype (CIMP) frequency in colon cancer using different probe- and gene-specific scoring alternatives on recommended multi-gene panels. *PLoS One* 2014;9:e86657.

50. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 2016;44:e71.

Figure Legends

Figure 1: Project workflow and description of sample sets used for the different working stages

Discovery sets: CRC and adenoma DNA samples collected for the study of methylome (450K), transcriptome (HT-12) and gene expression by qRT-PCR. In silico validation sets: data used to validate the identified biomarkers panel and for the evaluation of RNA-Seq (TCGA data set). Screening / wet validation sets: samples used for locus-specific methylation alterations in a further collection of tissues (tumor and peritumoral), stool samples and in the plasma of patients suffering from CRC.

Figure 2: Differential methylation and functional annotation analyses results obtained from the discovery data set methylome study

A) RnBeads differential methylation analysis performed on the data obtained using Infinium HumanMethylation 450 BeadChip, for single CpGs and for sets of pre-defined genomic regions such as genome-wide 5kb tiling regions, genes, promoters and CpG islands, in CRC *vs*. peritumoral tissue samples. Each dot represents the average beta value (methylation value) for each CpG locus in the region, resulting from the average of the samples belonging to that group; in red are those significantly differentially methylated. The pathways significantly altered are shown (in descending order of significance), according to the functional annotation analysis performed by ToppGene package. This analysis was conducted on the genes corresponding to the CpG islands significantly altered, according to the 450K manifest. The outputs generated by ToppGene show: the number of genes belonging to the pathway, the number of altered genes belonging to the pathway, and significance levels, raw and corrected for multiple testing according to Bonferroni and to Benjamini & Hochberg.

B) RnBeads differential methylation analysis performed in adenomas *vs*. normal mucosa samples.

C) The most significantly altered pathways are also shown, according to the functional annotation analysis performed by ToppGene. Circos plot resulting by the comparison of pathways significantly altered in CRC (left) and adenomas (right). For each pathway, identified by a number (whose correspondence is given in the supplementary material), are given the number of genes involved and the relative frequency. The red arrow indicates the pathways which show significant enrichment

of altered loci (decreasing p value) in CRC; the blue arrow in adenomas. The name of the five pathways most significantly altered in CRC are given and highlighted in color. Colored beams link the shared genes. The pathways significantly altered, obtained by functionally annotating the entire list of genes altered in CRC and adenomas, are also shown.

Figure 3: Discovery set unsupervised hierarchical clustering analysis based on the average methylation beta value for each of the 74 CpG islands

Heat map obtained by UHC of CRCs, adenomas, peritumoral tissue and normal mucosa samples. To the right of the heat map the genes associated to the 74 CpG islands aberrantly methylated and in common between CRC and adenomas are listed. All CRCs grouped on a branch separated from peritumoral and normal mucosa samples, with the sole exception of samples 352T and 279T.While seven adenomas branched along with peritumoral and normal mucosa samples, the remaining ones closely resembled the methylation pattern of carcinomas. Peritumoral and normal mucosa samples show an extremely similar pattern. No correlation trend was observed between methylation patterns and staging, localization or mutational pattern in both CRCs and adenomas. Other information reported: MS, microsatellites stability or instability; APC and KRAS mutational status; CIMP like, CIMP status as defined in the Supplementary methods; histology, Dukes tumoral staging, adenomas grade; localization.

Figure 4: Network visualization of multi-cluster gene functional enrichment analysis performed by Toppcluster tool on the genes associated to the 74 biomarkes

In green squares: pathway annotation. In blue squares: GO: molecular function annotation. In red hexagons: gene set.

Figure 5: *In silico* validation, by unsupervised hierarchical clustering analysis, of the early biomarkers panel in different data sets

A) Heat map of TCGA COREAD cohort: 248 colon and 94 rectum adenocarcinomas, 37 colon and 7 rectum normal tissues.

B) Heat map of GSE48684 data set: 42 adenomas, 64 carcinomas and 41 normal mucosa from colon. It is evident that the biomarkers panel was able to separate the majority of carcinomas and adenomas from normal mucosa, while no correlation between the methylation degree and tumor staging or location could be observed.

C) Heat map of GSE52270 data set: 103 CRC samples, 18 colon peritumoral tissues, 66 breast

cancers and 19 no-tumoral breast tissues, 48 glioblastomas and 10 white matter samples. The biomarkers panel is specific for colorectal cancer, but not for the other analyzed cancer types, with rare exceptions only for few markers.

D) Heat map of GSE53051 data set: 9 CRC samples, 18 colon peritumoral tissues, 10 adenomas, 16 colon cancer metastases, 14 breast cancers and 10 no-tumoral breast tissues, 9 lung cancer and 11 no-tumoral lung tissues, 29 pancreas cancer and 12 no-tumoral pancreas tissues, 70 thyroid cancer and 13 no-tumoral thyroid. Colon cancer metastases localized in the lung and in the liver are on the same branches of CRC, with the exception of two cases.

Figure 6: Validation of *GRIA4*, *SLC8A1* and *SYN3* specific methylation alterations in tissues, stool DNA samples and cfDNA

A) MethylBEAMing analysis results obtained for *GRIA4*, *SLC8A1* and *SYN3* in DNA isolated from stool samples of CRC patients, taken at the time of surgical resection. Colored bars show the methylation percentage at the three islands, cumulated for each sample.

B) MethylBEAMing methylation values in cfDNA, isolated from plasma, for *GRIA4*, *SLC8A1* and *SYN3*. Samples were divided into three groups: NED, CEA-low and CEA-high, using a threshold of 5 ng/ml. **c**) Graph showing the percentages of genetic and methylation alterations detected in plasma (left Y axis) and CEA levels (right Y axis) for all samples, divided into NED, CEA-low and CEA-high.

All shown comparison were performed using Mann-withney U-test. Asterisks indicate the significance level: * = p value <0.05; ** = p value <0.01; *** = p value <0.001.

Figure 1

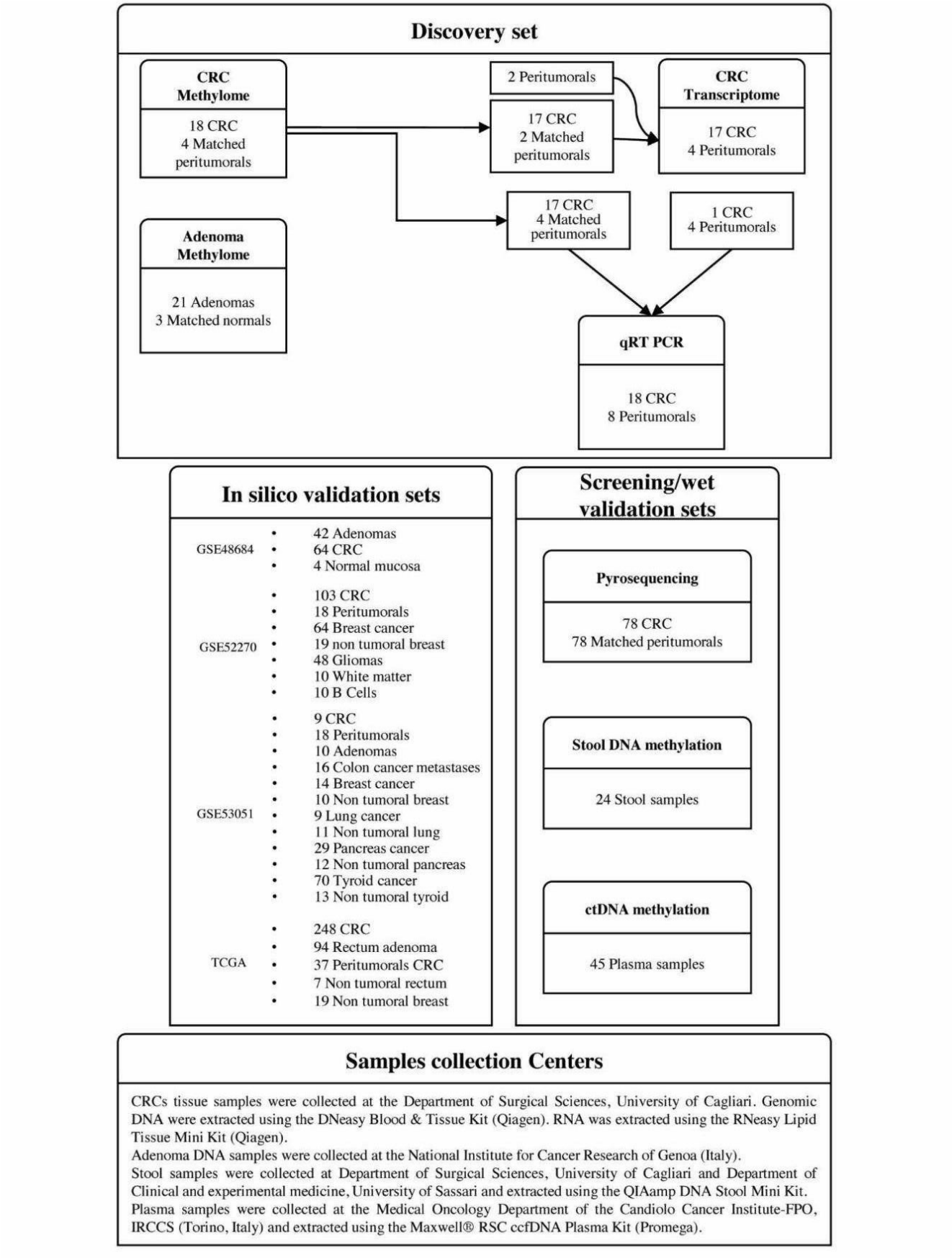

Discovery set
CRC Methylome
18 CRC
4 Matched peritumorals
2 Peritumorals
17 CRC
2 Matched peritumorals
CRC Transcriptome
17 CRC
4 Peritumorals
Adenoma Methylome
21 Adenomas
3 Matched normals
17 CRC
4 Matched peritumorals
1 CRC
4 Peritumorals
qRT PCR
18 CRC
8 Peritumorals
In silico validation sets
GSE48684
42 Adenomas
64 CRC
4 Normal mucosa
GSE52270
103 CRC
18 Peritumorals
64 Breast cancer
19 non tumoral breast
48 Gliomas
10 White matter
10 B Cells
GSE53051
9 CRC
18 Peritumorals
10 Adenomas
16 Colon cancer metastases
14 Breast cancer
10 Non tumoral breast
9 Lung cancer
11 Non tumoral lung
29 Pancreas cancer
12 Non tumoral pancreas
70 Tyroid cancer
13 Non tumoral tyroid
TCGA
248 CRC
94 Rectum adenoma
37 Peritumorals CRC
7 Non tumoral rectum
19 Non tumoral breast
Screening/wet validation sets
Pyrosequencing
78 CRC
78 Matched peritumorals
Stool DNA methylation
24 Stool samples
ctDNA methylation
45 Plasma samples
Samples collection Centers
CRCs tissue samples were collected at the Department of Surgical Sciences, University of Cagliari. Genomic DNA were extracted using the DNeasy Blood & Tissue Kit (Qiagen). RNA was extracted using the RNeasy Lipid Tissue Mini Kit (Qiagen).
Adenoma DNA samples were collected at the National Institute for Cancer Research of Genoa (Italy).
Stool samples were collected at Department of Surgical Sciences, University of Cagliari and Department of Clinical and experimental medicine, University of Sassari and extracted using the QIAamp DNA Stool Mini Kit.
Plasma samples were collected at the Medical Oncology Department of the Candiolo Cancer Institute-FPO, IRCCS (Torino, Italy) and extracted using the Maxwell® RSC ccfDNA Plasma Kit (Promega).

Figure 2

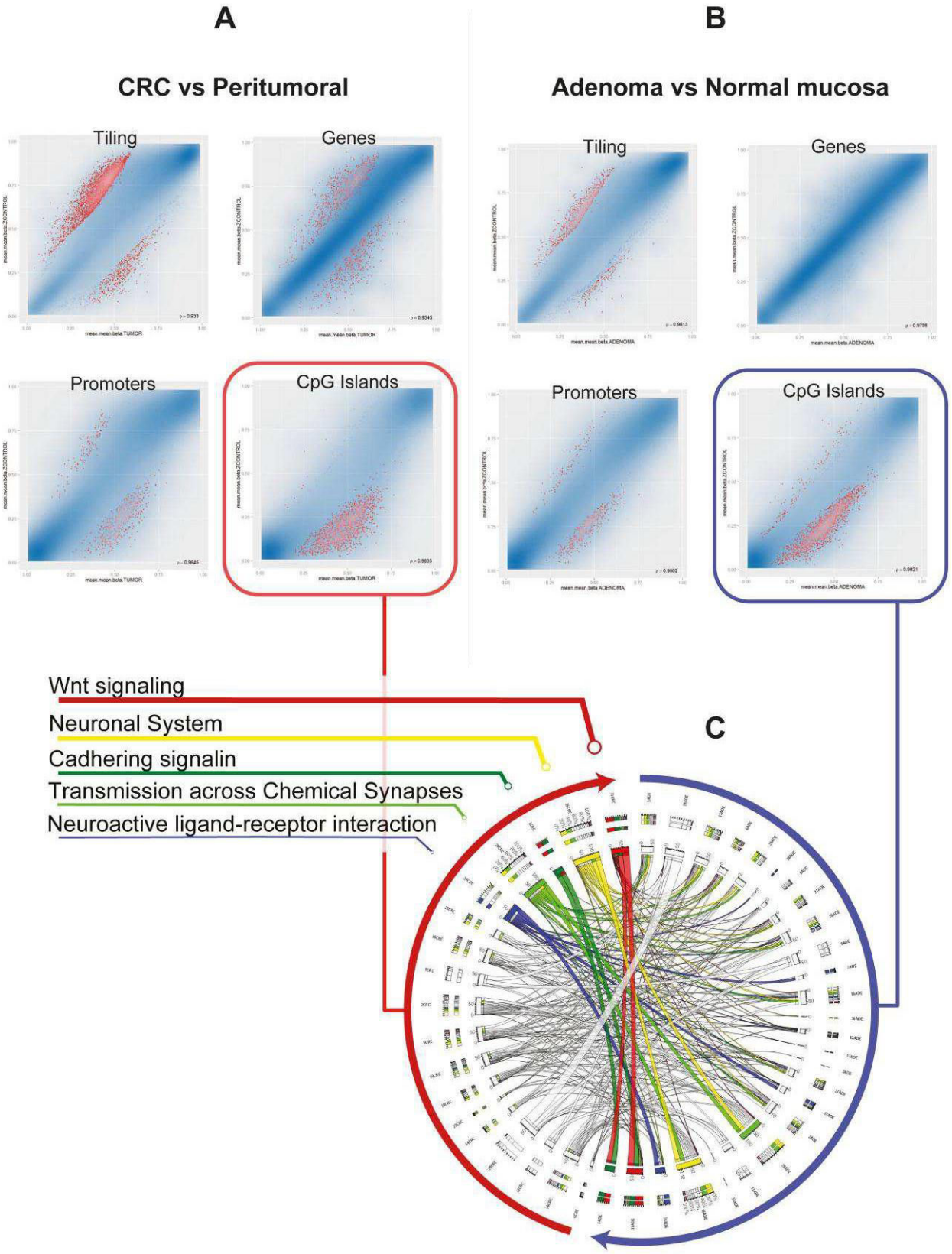

A
CRC vs Peritumoral
Tiling
Genes
Promoters
CpG Islands
B
Adenoma vs Normal mucosa
Tiling
Genes
Promoters
CpG Islands
Wnt signaling
Neuronal System
Cadhering signalin
Transmission across Chemical Synapses
Neuroactive ligand-receptor interaction
C

Figure 3

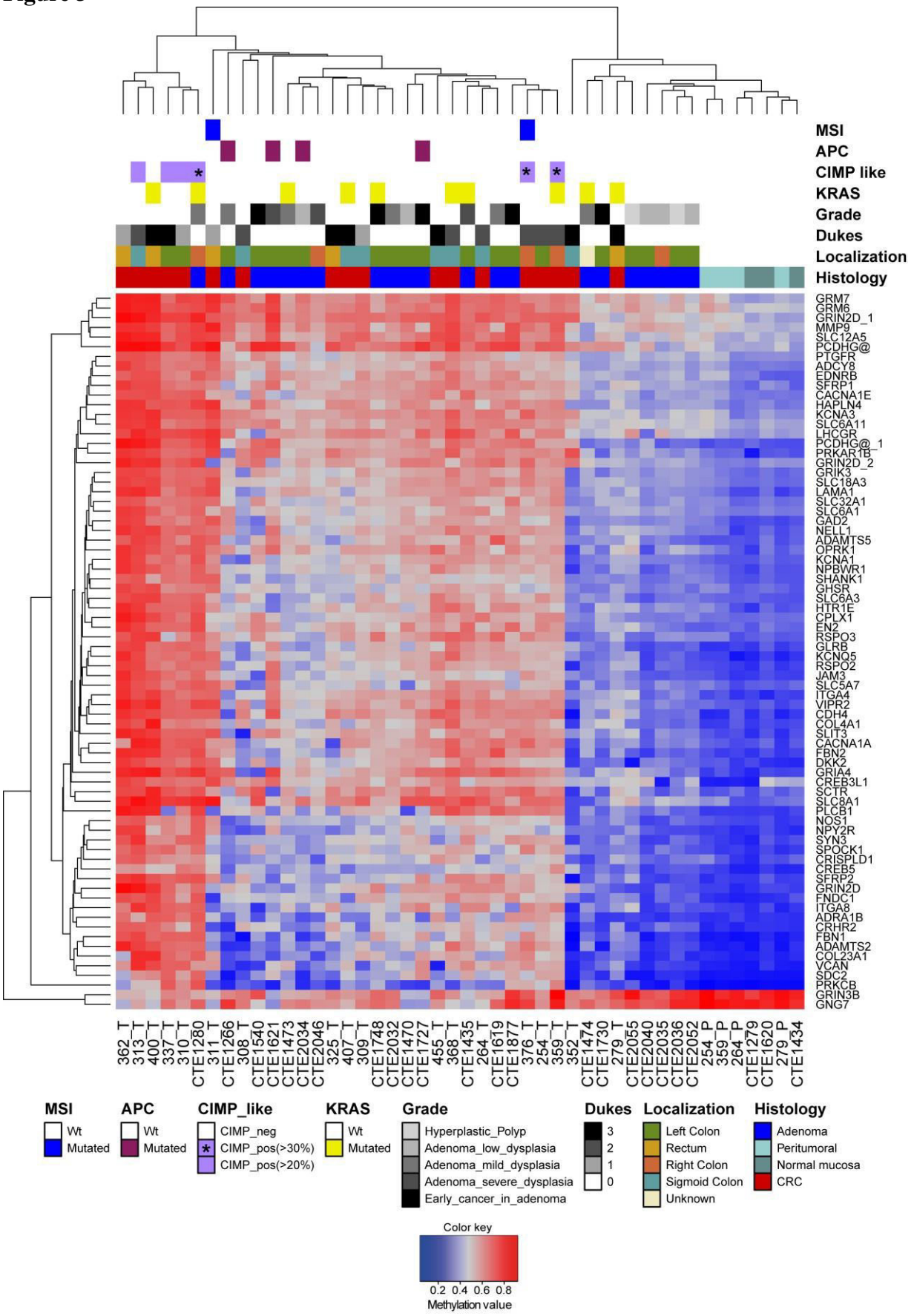

MSI
APC
CIMP like
KRAS
Grade
Dukes
Localization
Histology
GRM7
GRM6
GRIN2D_1
MMP9
SLC12A5
PCDHG@
PTGFR
ADCY8
EDNRB
SFRP1
CACNA1E
HAPLN4
KCNA3
SLC6A11
LHCGR
PCDHG@_1
PRKAR1B
GRIN2D_2
GRIK3
SLC18A3
LAMA1
SLC32A1
SLC6A1
GAD2
NELL1
ADAMTS5
OPRK1
KCNA1
NPBWR1
SHANK1
GHSR
SLC6A3
HTR1E
CPLX1
EN2
RSPO3
GLRB
KCNQ5
RSPO2
JAM3
SLC5A7
ITGA4
VIPR2
CDH4
COL4A1
SLIT3
CACNA1A
FBN2
DKK2
GRIA4
CREB3L1
SCTR
SLC8A1
PLCB1
NOS1
NPY2R
SYN3
SPOCK1
CRISPLD1
CREB5
SFRP2
GRIN2D
FNDC1
ITGA8
ADRA1B
CRHR2
FBN1
ADAMTS2
COL23A1
VCAN
SDC2
PRKCB
GRIN3B
GNG7
362_T
313_T
400_T
337_T
310_T
CTE1280
311_T
CTE1266
308_T
CTE1540
CTE1621
CTE1473
CTE2034
CTE2046
325_T
407_T
309_T
CTE1748
CTE2032
CTE1470
CTE1727
455_T
368_T
CTE1435
264_T
CTE1619
CTE1877
376_T
254_T
359_T
352_T
CTE1474
CTE1730
279_T
CTE2055
CTE2040
CTE2035
CTE2036
CTE2052
254_P
359_P
264_P
CTE1279
CTE1620
279_P
CTE1434
MSI
Wt
Mutated
APC
Wt
Mutated
CIMP_like
CIMP_neg
CIMP_pos(>30%)
CIMP_pos(>20%)
KRAS
Wt
Mutated
Grade
Hyperplastic_Polyp
Adenoma_low_dysplasia
Adenoma_mild_dysplasia
Adenoma_severe_dysplasia
Early_cancer_in_adenoma
Dukes
3
2
1
0
Localization
Left Colon
Rectum
Right Colon
Sigmoid Colon
Unknown
Histology
Adenoma
Peritumoral
Normal mucosa
CRC
Color key
0.2 0.4 0.6 0.8
Methylation value

Figure 4

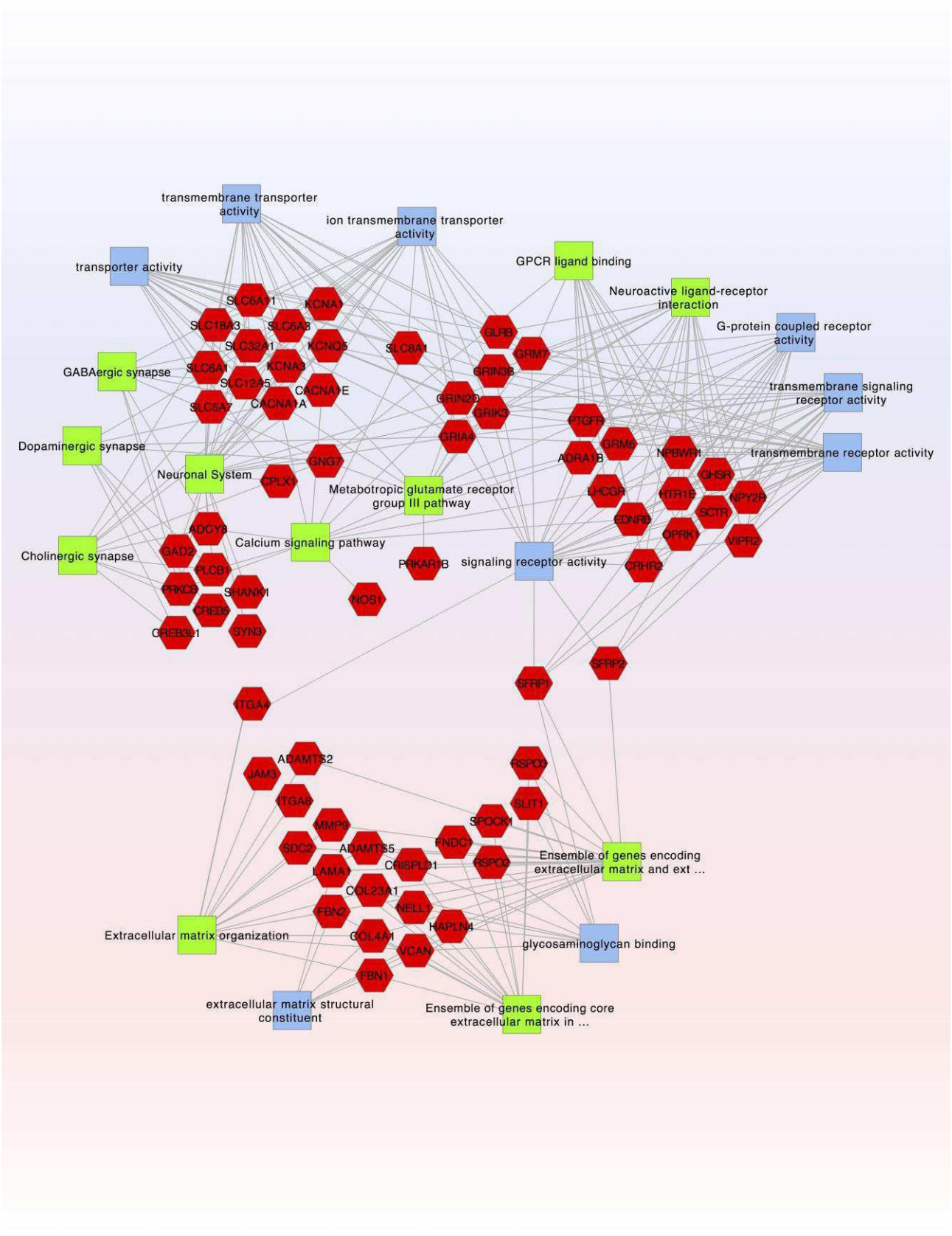

transmembrane transporter activity
ion transmembrane transporter activity
transporter activity
GPCR ligand binding
Neuroactive ligand-receptor interaction
G-protein coupled receptor activity
GABAergic synapse
transmembrane signaling receptor activity
Dopaminergic synapse
transmembrane receptor activity
Neuronal System
Metabotropic glutamate receptor group III pathway
Calcium signaling pathway
Cholinergic synapse
signaling receptor activity
SLC6A1
KCNA1
SLC18A3
SLC6A3
SLC32A1
KCNQ5
SLC8A1
GLRB
GRM7
SLC6A1
KCNA3
GRIN3B
SLC12A5
CACNA1E
GRIN2D
SLC5A7
CACNA1A
GRIK3
GRIA4
PTGFR
GRM6
ADRA1B
NPBWR1
GHSR
LHCGR
HTR1E
NPY2R
SCTR
EDNRB
OPRK1
VIPR2
CRHR2
GNG7
CPLX1
ADCY3
GAD2
PLCB1
PRKCZ
SHANK1
CREB5
CREB3L1
SYN3
PRKAR1B
NOS1
SFRP2
SFRP1
ITGA4
ADAMTS2
JAM3
ITGA8
RSPO3
SLIT1
SPOCK1
MMP9
SDC2
ADAMTS5
FNDC1
RSPO2
CRISPLD1
LAMA1
COL23A1
FBN2
NELL1
HAPLN4
COL4A1
VCAN
FBN1
Ensemble of genes encoding extracellular matrix and ext ...
Extracellular matrix organization
glycosaminoglycan binding
extracellular matrix structural constituent
Ensemble of genes encoding core extracellular matrix in ...

Figure 5

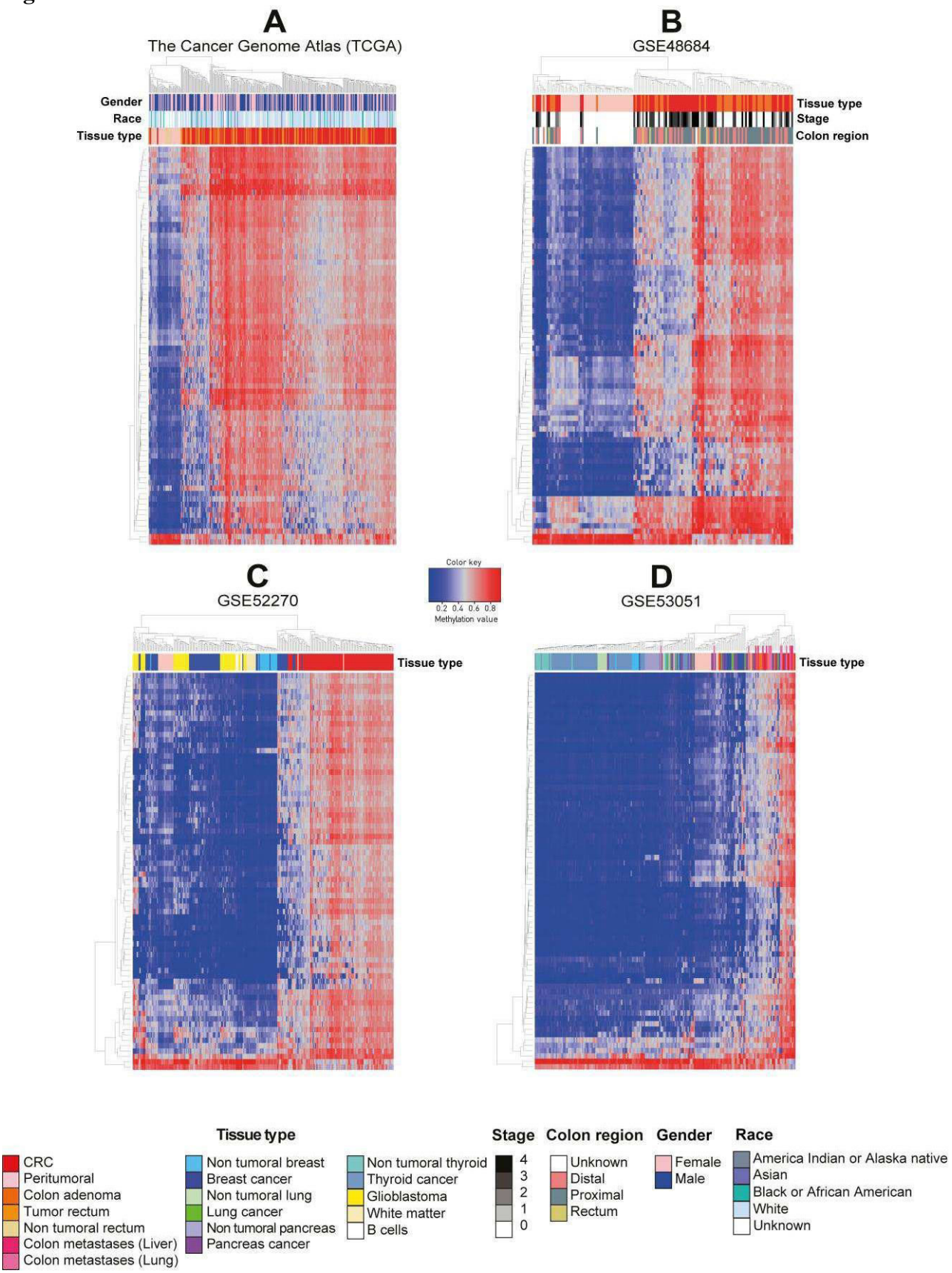
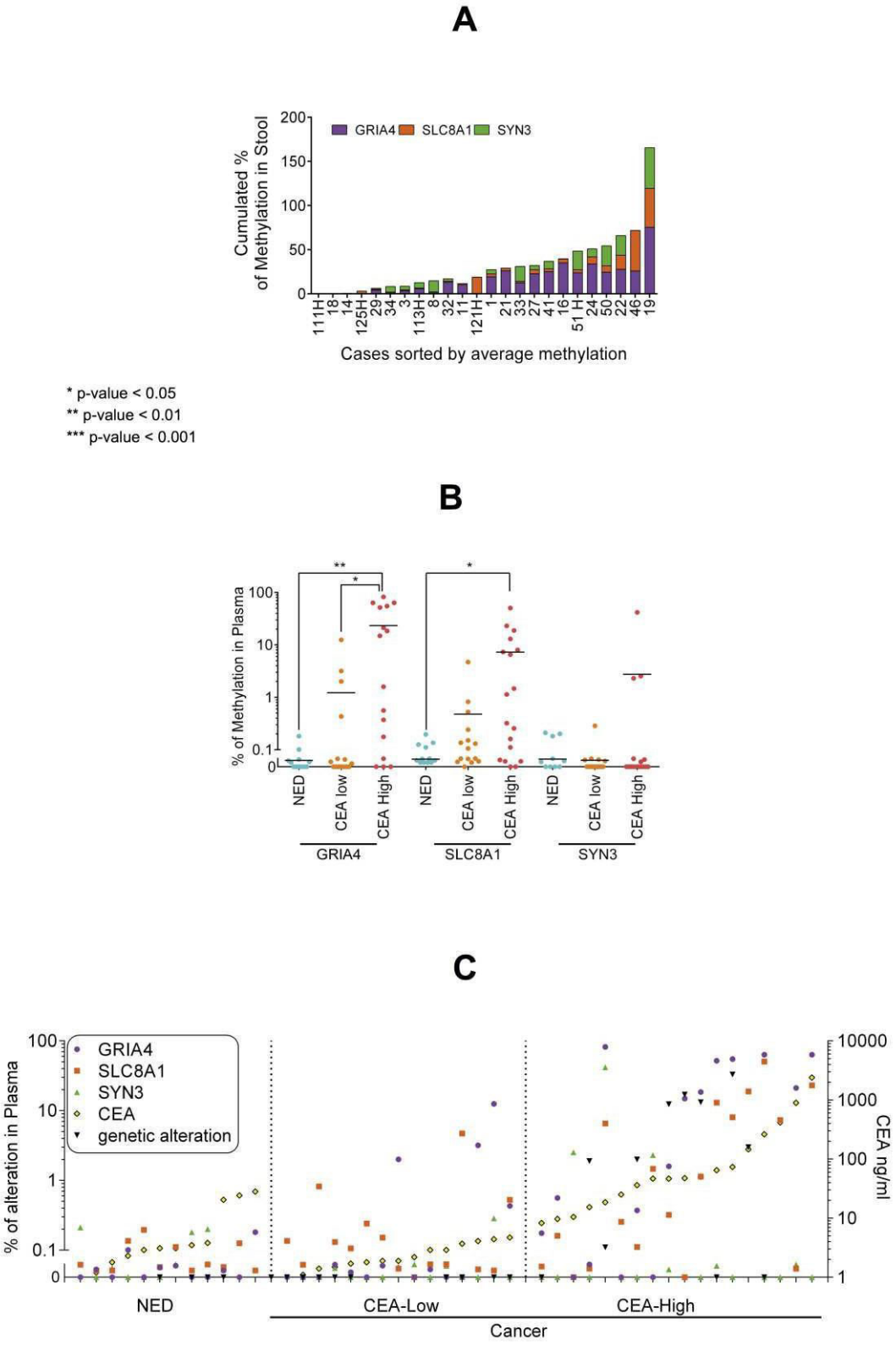
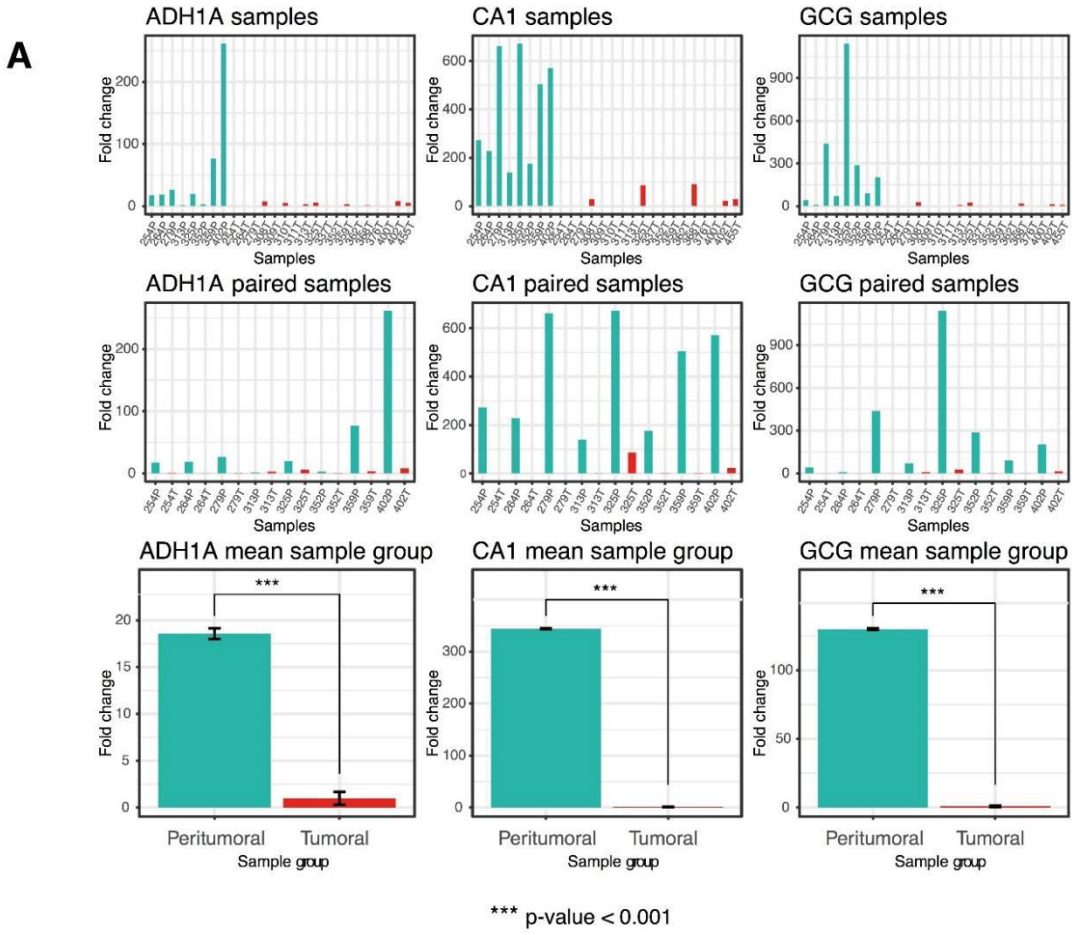
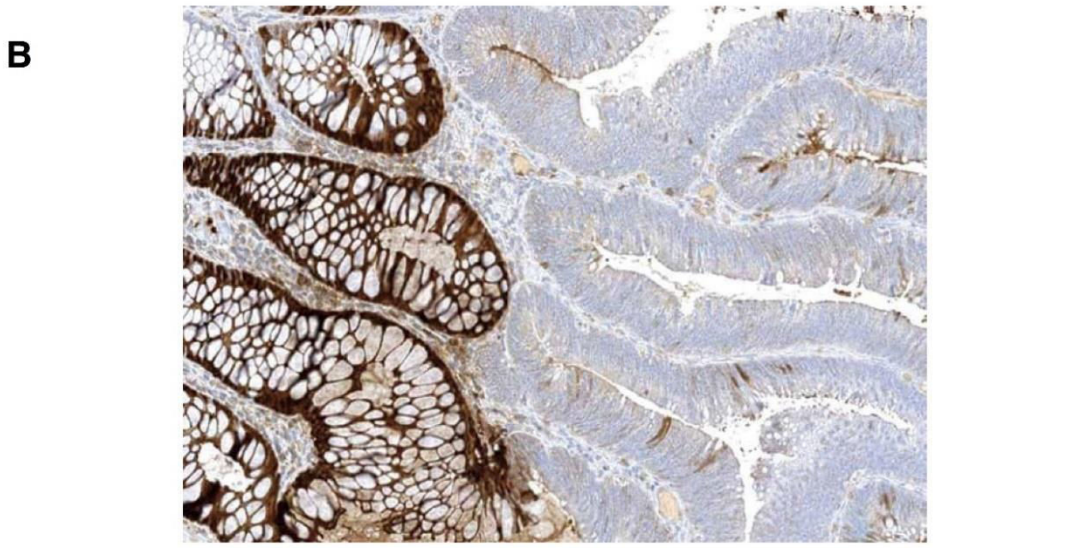

A
The Cancer Genome Atlas (TCGA)
Gender
Race
Tissue type
B
GSE48684
Tissue type
Stage
Colon region
Color key
0.2 0.4 0.6 0.8
Methylation value
C
GSE52270
Tissue type
D
GSE53051
Tissue type
Tissue type
CRC
Peritumoral
Colon adenoma
Tumor rectum
Non tumoral rectum
Colon metastases (Liver)
Colon metastases (Lung)
Non tumoral breast
Breast cancer
Non tumoral lung
Lung cancer
Non tumoral pancreas
Pancreas cancer
Non tumoral thyroid
Thyroid cancer
Glioblastoma
White matter
B cells
Stage
4
3
2
1
0
Colon region
Unknown
Distal
Proximal
Rectum
Gender
Female
Male
Race
America Indian or Alaska native
Asian
Black or African American
White
Unknown

Figure 6


A
GRIA4
SLC8A1
SYN3
Cumulated % of Methylation in Stool
200
150
100
50
0
111H
18
14
125H
29
34
3
113H
8
32
11
121H
1
21
33
27
41
16
51 H
24
50
22
46
19
Cases sorted by average methylation
* p-value < 0.05
** p-value < 0.01
*** p-value < 0.001
B
% of Methylation in Plasma
100
10
1
0.1
0
**
*
*
NED
CEA low
CEA High
GRIA4
SLC8A1
SYN3
C
% of alteration in Plasma
100
10
1
0.1
0
GRIA4
SLC8A1
SYN3
CEA
genetic alteration
10000
1000
100
10
1
CEA ng/ml
NED
CEA-Low
CEA-High
Cancer

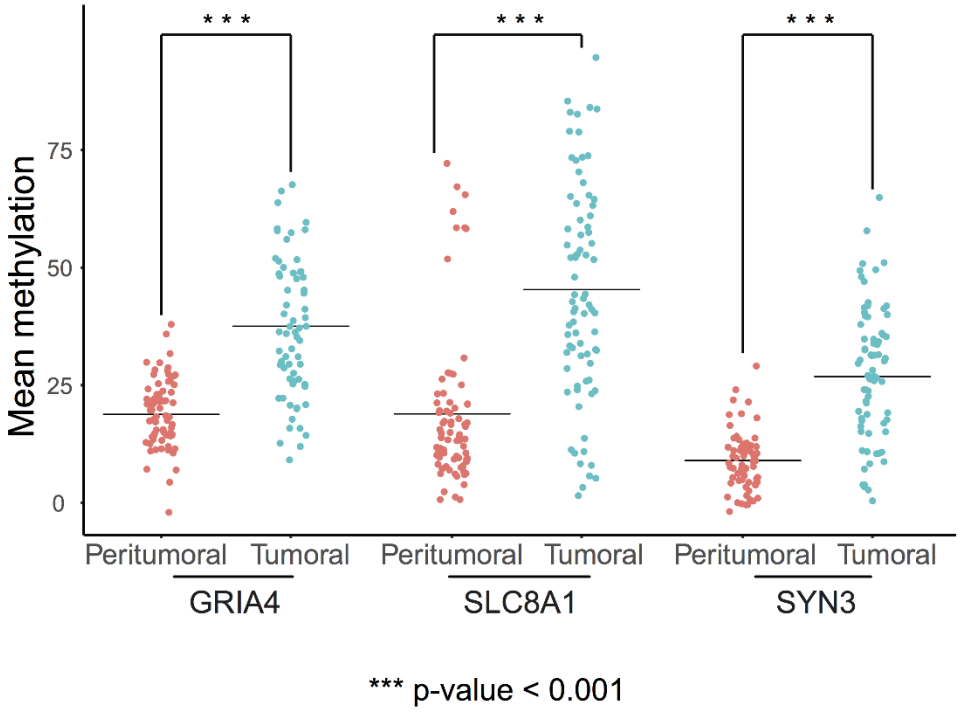




Supporting Information Figure 1: Validation of changes in gene expression observed by WGGE analysis

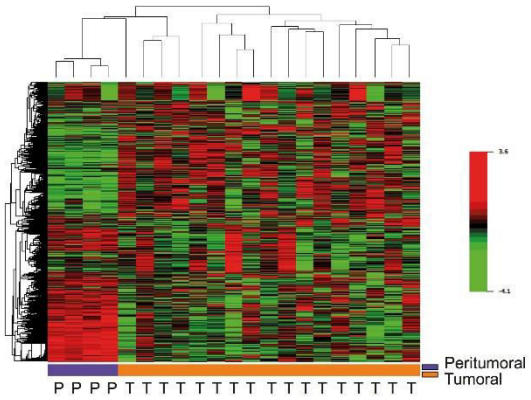
A) Validation of changes in gene expression observed by WGGE analysis, evaluated by qRT-PCR on three selected genes: ADH1A (Alcohol Dehydrogenase 1A), CA1 (Carbonic Anhydrase I), and GCG (Glucagon). For each gene is shown the fold change of 8 peritumoral (P) and 18 tumor (T) samples, of 8 matched pairs (P / T), and the average of P versus T.

B) Colorectal adenoma (left) and carcinoma (right) tissues, immunostained for CA1.

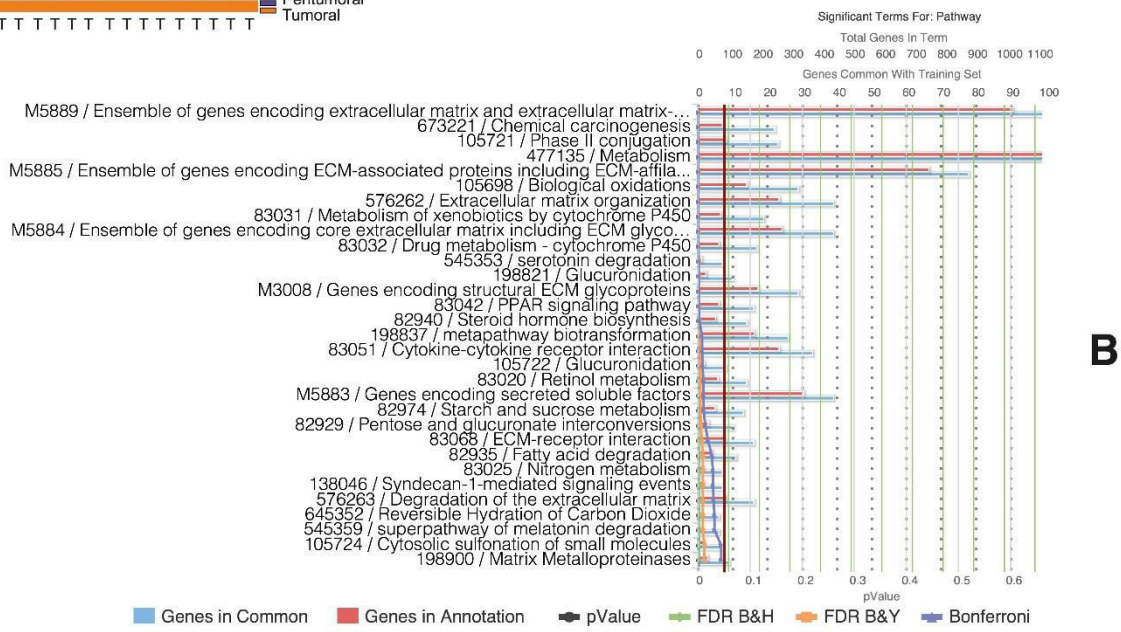


Supporting Information Figure 2: Pyrosequencing validation of *GRIA4*, *SLC8A1* and *SYN3* specific methylation alterations in a second data set of tumoral and peritumoral tissues

Pyrosequencing methylation analysis of three selected islands, related to the *GRIA4*, *SLC8A1* and *SYN3* genes, performed in a second data set of 78 tumoral and 78 peritumoral samples. On the Y axis is reported the methylation average value of the CpG loci analyzed for each island.

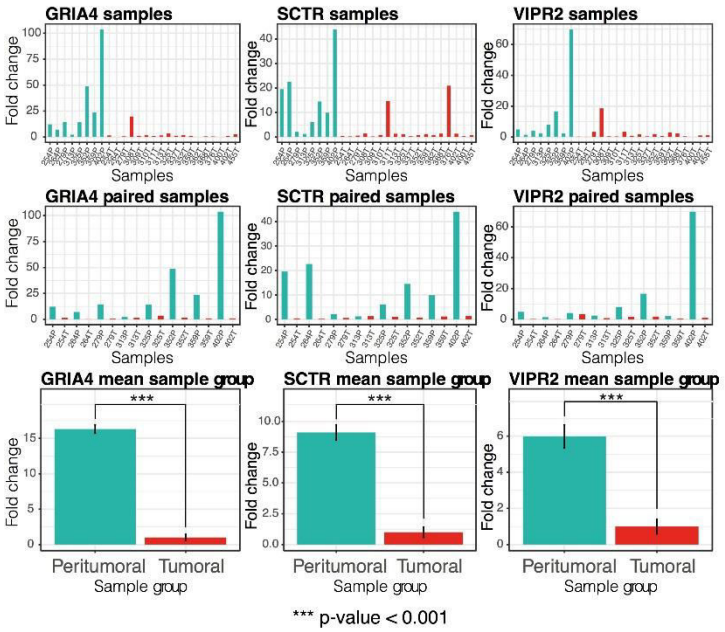

P P P P T T T T T T T T T T T T T T T T
Peritumoral
Tumoral

A

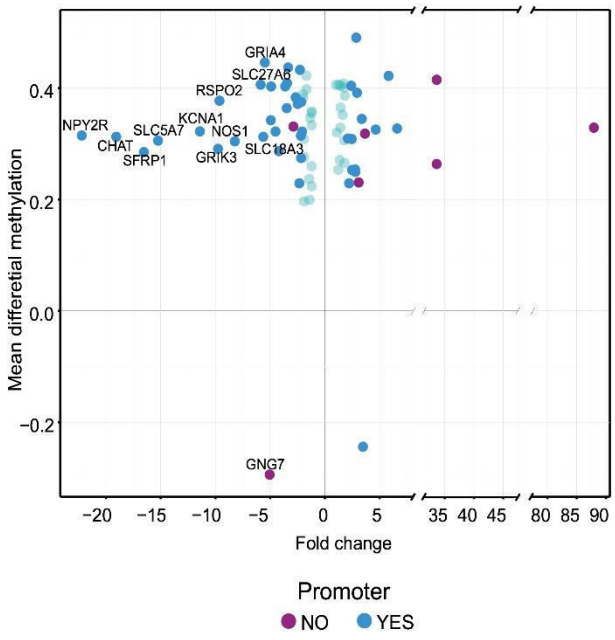

Significant Terms For: Pathway
Total Genes In Term
Genes Common With Training Set
M5889 / Ensemble of genes encoding extracellular matrix and extracellular matrix-...
673221 / Chemical carcinogenesis
105721 / Phase II conjugation
477135 / Metabolism
M5885 / Ensemble of genes encoding ECM-associated proteins including ECM-affila...
105698 / Biological oxidations
576262 / Extracellular matrix organization
83031 / Metabolism of xenobiotics by cytochrome P450
M5884 / Ensemble of genes encoding core extracellular matrix including ECM glyco...
83032 / Drug metabolism - cytochrome P450
545353 / serotonin degradation
198821 / Glucuronidation
M3008 / Genes encoding structural ECM glycoproteins
83042 / PPAR signaling pathway
82940 / Steroid hormone biosynthesis
198837 / metapathway biotransformation
83051 / Cytokine-cytokine receptor interaction
105722 / Glucuronidation
83020 / Retinol metabolism
M5883 / Genes encoding secreted soluble factors
82974 / Starch and sucrose metabolism
82929 / Pentose and glucuronate interconversions
83068 / ECM-receptor interaction
82935 / Fatty acid degradation
83025 / Nitrogen metabolism
138046 / Syndecan-1-mediated signaling events
576263 / Degradation of the extracellular matrix
645352 / Reversible Hydration of Carbon Dioxide
545359 / superpathway of melatonin degradation
105724 / Cytosolic sulfonation of small molecules
198900 / Matrix Metalloproteinases
pValue
Genes in Common
Genes in Annotation
pValue
FDR B&H
FDR B&Y
Bonferroni

B

C


GRIA4 samples
SCTR samples
VIPR2 samples
GRIA4 paired samples
SCTR paired samples
VIPR2 paired samples
GRIA4 mean sample group
SCTR mean sample group
VIPR2 mean sample group
Fold change
Samples
Peritumoral
Tumoral
Sample group
***
*** p-value < 0.001

D


GRIA4
SLC27A6
RSPO2
NPY2R
KCNA1
SLC5A7
NOS1
CHAT
SLC18A3
SFRP1
GRIK3
GNG7
Mean differetial methylation
Fold change
Promoter
NO
YES

Supporting Information Figure 3: Differential gene expression and functional annotation analyses

A) Heat map showing the results of an unsupervised hierarchical clustering analysis performed on the average intensity values obtained from the WGGE analysis on the discovery CRC and peritumoral tissue samples sets. WGGE analysis was conducted using the HumanHT-12 v4 Expression BeadChip. As expected, all CRCs grouped on a branch separated from peritumoral samples. The results of the differential expression analysis, showed the presence of 963 significantly dysregulated transcripts (329 downregulated, 171 upregulated, using a cutoff value of $|FC| > 2$), some of which were also validated by qRT-PCR and at the protein level by immunostaining (**Supporting Information Figure 1**).

B) Pathways resulting dysregulated by functional annotation of genes differentially expressed in the WGGE analysis, performed by ToppGene package. The analysis showed that the most altered pathways were involved in biological oxidation processes. The highlighted pathways did not overlap with those of the methylome (compare **Supporting Information Table 5**), apart from those relating to the extracellular matrix. A number of down-regulated pathways (such as those involved in amines degradation) are likely downstream of those modified as a consequence of hypermethylation.

C) Functional validation of CpG islands hypermethylation as target genes downregulation, evaluated by qRT-PCR on three selected genes: GRIA4 (glutamate receptor, ionotropic, AMPA 4), SCTR (secretin receptor), and VIPR2 (vasoactive intestinal peptide receptor 2). For each gene is shown the fold change of 8 peritumoral (P) and 18 tumor (T) samples, of 8 matched pairs (P / T), and the average of P versus T. All the three tested genes were significantly down-regulated in cancer.

D) TCGA CRC RNAseq data corresponding to the gene neighbors the 74 selected altered CpG islands. The X axis shows the expression fold change value; along the Y axis the average value of differential methylation for each CpG island. In blue, CpG islands included in promoter regions (according to the definition given in Methods), in purple those not falling into promoters. More faded the symbols corresponding to the CpG islands showing a not significant fold change (between -2 and +2). Almost all the selected altered CpG islands were located in the promoter region of genes (defined as the sequence between 2kb upstream and 1kb downstream the TSS), and the majority of them (> 70%) were associated to down-regulation of the corresponding genes in the tumor tissues. The only islands associated to the three genes strongly upregulated, are actually not located into their promoter regions but in the gene body or downstream.

Supporting Information Table 1: CRC Discovery Set

Samples	KRAS	MSI	Dukes	Histology	CIMP_like	Anatomic site	Methylome analysis	Transcriptome analysis	qRT-PCR
254_P	Wt	Wt		Peritumoral	CIMP_neg	Left Colon	Yes	Yes	Yes
264_P	Wt	Wt		Peritumoral	CIMP_neg	Sigmoid colon	Yes	No	Yes
279_P	Wt	Wt		Peritumoral	CIMP_neg	Rectum	Yes	Yes	Yes
313_P	Wt	Wt		Peritumoral	NA	NA	No	Yes	Yes
325_P	Wt	Wt		Peritumoral	NA	NA	No	Yes	Yes
352_P	Wt	Wt		Peritumoral	NA	NA	No	No	Yes
359_P	Wt	Wt		Peritumoral	CIMP_neg	Right Colon	Yes	No	Yes
402_P	Wt	wt		Peritumoral	NA	NA	No	no	Yes
254_T	Wt	Wt	b	CRC	CIMP_neg	Left Colon	Yes	Yes	Yes
264_T	Wt	Wt	b	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	Yes
279_T	Mutated	Wt	d	CRC	CIMP_neg	Rectum	Yes	Yes	Yes
308_T	Wt	Wt	b	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	Yes
309_T	Wt	Wt	a	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	Yes
310_T	Wt	Wt	a	CRC	CIMP_pos(>20%)	Left Colon	Yes	Yes	Yes
311_T	Wt	Mutated	a	CRC	CIMP_neg	Rectum	Yes	Yes	Yes
313_T	Wt	Wt	b	CRC	CIMP_pos(>20%)	Sigmoid colon	Yes	Yes	Yes
325_T	Wt	Wt	d	CRC	CIMP_neg	Rectum	Yes	Yes	Yes
337_T	Wt	Wt	d	CRC	CIMP_pos(>20%)	Left Colon	Yes	Yes	Yes
352_T	Wt	Wt	d	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	Yes
359_T	Mutated	Wt	b	CRC	CIMP_pos(>30%)	Right Colon	Yes	Yes	Yes
362_T	Wt	Wt	a	CRC	CIMP_neg	Rectum	Yes	Yes	Yes
368_T	Mutated	Wt	b	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	Yes
376_T	Wt	Mutated	b	CRC	CIMP_pos(>30%)	Right Colon	Yes	No	Yes
400_T	Mutated	Wt	d	CRC	CIMP_neg	Rectum	Yes	Yes	Yes
402_T	NA	NA	NA	CRC	NA	NA	No	No	Yes
407_T	Mutated	Wt	d	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	No
455_T	Wt	Wt	d	CRC	CIMP_neg	Sigmoid colon	Yes	Yes	Yes

Supporting Information Table 2: Adenomas Discovery Set

Sample	KRAS	APC	Grade	Histology	CIMP_like	Anatomic site
CTE1279	Wt	Wt		Normal	CIMP_neg	Right Colon
CTE1434	Wt	Wt		Normal	CIMP_neg	Left Colon
CTE1620	Wt	Wt		Normal	CIMP_neg	Left Colon
CTE1266	Wt	Mutated	Adenoma_mild_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE1280	Mutated	Wt	Adenoma_mild_dysplasia	Adenoma	CIMP_pos(>30%)	Right Colon
CTE1435	Mutated	Wt	Adenoma_severe_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE1470	Wt	Wt	Adenoma_low_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE1473	Mutated	Wt	Adenoma_mild_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE1474	Mutated	Wt	Adenoma_mild_dysplasia	Adenoma	CIMP_neg	NA
CTE1540	Wt	Wt	Early cancer in adenoma	Adenoma	CIMP_neg	Left Colon
CTE1619	Wt	Wt	Adenoma_mild_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE1621	Wt	Mutated	Adenoma_severe_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE1727	Wt	Mutated	Early cancer in adenoma	Adenoma	CIMP_neg	Left Colon
CTE1730	Wt	Wt	Early cancer in adenoma	Adenoma	CIMP_neg	Left Colon
CTE1748	Mutated	Wt	Early cancer in adenoma	Adenoma	CIMP_neg	Left Colon
CTE1877	Wt	Wt	Early cancer in adenoma	Adenoma	CIMP_neg	Left Colon
CTE2032	Wt	Wt	Adenoma_mild_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE2034	Wt	Mutated	Adenoma_low_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE2035	Wt	Wt	Adenoma_low_dysplasia	Adenoma	CIMP_neg	Right Colon
CTE2036	Wt	Wt	Hyperplastic polyp	Adenoma	CIMP_neg	Left Colon
CTE2040	Wt	Wt	Adenoma_low_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE2046	Wt	Wt	Adenoma_severe_dysplasia	Adenoma	CIMP_neg	Right Colon
CTE2052	Wt	Wt	Adenoma_low_dysplasia	Adenoma	CIMP_neg	Left Colon
CTE2055	Wt	Wt	Hyperplastic polyp	Adenoma	CIMP_neg	Left Colon

Supporting Information Table 3: List of primers and probes used in the present study

Assay	Gene	Primer/Probe	Sequence (5'-3')
Genetic mutations screening	*KRAS* (exon 2)	Forward	ACTGGTGGAGTATTTGATAGTGTAT
		Reverse	AGAATGGTCCTGCACCAGTAA
	KRAS (exon 3)	Forward	TCCAGACTGTGTTTCTCCCT
		Reverse	AACCCACCTATAATGGTGAATATCT
Pyrosequencing analysis	*SYN3*	Forward	[btn]TTGGGTAGGTTTTTGGGATAGATAG
		Reverse	ATAAAAACAATCTTAAAATCCACAAT
		Sequence	ACAATCTTAAAATCCACAATC
	SLC8A1	Forward	TTTGGGGAAAGATTTTAGGGATTA
		Reverse	[btn]AAAATCCAAACCTCCCCACCCACT
		Sequence	GGGAAAGATTTTAGGGATTAT
	GRIA4	Forward	GGGTTGGTGTAGGTTTGTT
		Reverse	[btn]CTCCCCCCTTACTTTCTCACATACACACAA
		Sequence	GTGTAGGTTTGTTGGG
MethylBEAMing analysis	*SYN3*	Forward	TCCCGCGAAATTAATACGACTTGGGTAGGTTTTTGGGATAGATAG
		Reverse	GCTGGAGCTCTGCAGCTAATAAAAACAATCTTAAAATCCACAAT
		met-probe	/5Alex647N/GACAAACGACCCCCGCACG
		unm-probe	/5Alex488N/CAACAAACAACCCCCACACA
	SLC8A1	Forward	TCCCGCGAAATTAATACGACTTTGGGGAAAGATTTTAGGGATTA
		Reverse	GCTGGAGCTCTGCAGCTAAAAATCCAAACCTCCCCACCCACT
		met-probe	/5Alex647N/GCGAACATCCCTCCTTCCG
		unm-probe	/5Alex488N/ACAAACATCCCTCCTTCCA
	GRIA4	Forward	TCCCGCGAAATTAATACGACGGGTTGGTGTAGGTTTGTT
		Reverse	GCTGGAGCTCTGCAGCTACTCCCCCCTTACTTTCTCACATACACACAA
		met-probe	/5Alex647N/AACGCCGCGACCGCCACAC
		unm-probe	/5Alex488N/CACCACAACCACCACACACA
qRT-PCR analysis	*SCTR*	Forward	GGATGGTGGAGGTGGAATG
		Reverse	AAGGTTTCTGACCAGCCATC
	VIPR2	Forward	GTCTCTTGCAACAGGAAGCA
		Reverse	TCTCAGGATGAAGGACAGGAA
	GRIA4	Forward	TCATGTGGACAACATTGAGACA
		Reverse	ATCATAGAGTCCAAAAATGGCAAA
	ADH1A	Forward	TGGAGGTGTGGATTTTTCATT
		Reverse	CCCTACGATGACACTTGTGC
	CA1	Forward	CAGTACAAATGAGCATGGTTCAG
		Reverse	GCAAGGCTGGAGTACTTTGC
	GCG	Forward	CCAAGATTTTGTGCAGTGGTT
		Reverse	GGTAAAGGTCCCTTCAGCAT
	TFRC (endogenus)	Forward	GGCACAGCTCTCCTATTGAAAC
		Reverse	CAAAGTCTCCAGCACTCCAACT

Supporting Information Table 4: Markers selection for validation of locus specific methylation alterations

	Chromosome	Start	End	Threshold	Specificity	Sensitivity	npv	ppv
ADAMTS2-A	chr5	178771329	178771394	0.313	0.956	0.923	0.573	0.995
ADAMTS2-B	chr5	178771776	178771782	0.229	1.000	0.908	0.542	1.000
ADAMTS2-C	chr5	178772372	178772391	0.147	1.000	0.851	0.421	1.000
CREB5-D	chr7	28449847	28450002	0.305	1.000	0.918	0.570	1.000
CRHR2-E	chr7	30722114	30722178	0.155	1.000	0.884	0.484	1.000
CRHR2-F	chr7	30722320	30722362	0.335	1.000	0.884	0.484	1.000
GRIN2D-G	chr19	48947560	48947632	0.333	0.933	0.969	0.764	0.993
HTR1E-H	chr6	87647370	87647528	0.393	1.000	0.865	0.446	1.000
KCNA1-I	chr12	5018715	5018806	0.269	1.000	0.904	0.529	1.000
KCNA1-J	chr12	5018984	5019092	0.233	1.000	0.906	0.536	1.000
OPRK1-K	chr8	54164081	54164162	0.355	1.000	0.894	0.506	1.000
OPRK1-L	chr8	54164296	54164392	0.358	0.933	0.930	0.592	0.992
PRKAR1B-M	chr7	751830	751963	0.479	1.000	0.964	0.750	1.000
SLC8A1-N	chr2	40678618	40678692	0.475	1.000	0.896	0.511	1.000
SYN3-O	chr22	33453893	33453995	0.372	1.000	0.923	0.584	1.000
SYN3-P	chr22	33454209	33454325	0.342	0.978	0.925	0.587	0.997
SYN3-Q	chr22	33454444	33454505	0.185	0.956	0.834	0.384	0.994
SCTR-R	chr2	120281719	120281813	0.202	0.978	0.923	0.579	0.997
GRIA4-S	chr11	105481283	105481406	0.347	1.000	0.957	0.714	1.000
VIPR2-T	chr7	158936508	158936632	0.320	1.000	0.959	0.726	1.000
VIPR2-U	chr7	158936632	158936739	0.148	0.978	0.978	0.830	0.998
VIPR2-V	chr7	158937494	158937610	0.343	1.000	0.947	0.672	1.000
VIPR2-W	chr7	158937969	158938060	0.440	1.000	0.961	0.738	1.000

Supporting Information Table 5:

A) CRC pathways most affected by aberrant methylations

Name	Source	p-value	q-value Bonferroni	q-value FDR B&H	q-value FDR B&Y	Hit Count in Query List	Hit Count in Genome
Wnt signaling pathway	PantherDB	1,67E-10	2,13E-07	1,46E-07	1,13E-06	35	305
Neuronal System	BioSystems: REACTOME	2,28E-10	2,92E-07	1,46E-07	1,13E-06	34	293
Cadherin signaling pathway	PantherDB	5,74E-10	7,33E-07	2,44E-07	1,89E-06	24	159
Transmission across Chemical Synapses	BioSystems: REACTOME	2,97E-09	3,79E-06	9,48E-07	7,33E-06	26	200
Neuroactive ligand-receptor interaction	BioSystems: KEGG	4,46E-08	5,69E-05	1,14E-05	8,80E-05	29	275
Neurotransmitter Release Cycle	BioSystems: REACTOME	2,53E-07	3,23E-04	5,38E-05	4,16E-04	10	37
GABAergic synapse	BioSystems: KEGG	1,63E-06	2,09E-03	2,98E-04	2,30E-03	14	90
Ensemble of genes encoding core extracellular matrix including ECM glycoproteins, collagens and proteoglycans	MSigDB C2 BIOCARTA (v5.1)	1,87E-06	2,39E-03	2,99E-04	2,31E-03	26	275
Calcium signaling pathway	BioSystems: KEGG	2,82E-06	3,61E-03	4,01E-04	3,10E-03	20	181
Cholinergic synapse	BioSystems: KEGG	5,37E-06	6,86E-03	6,86E-04	5,30E-03	15	113
Metabotropic glutamate receptor group III pathway	PantherDB	8,01E-06	1,02E-02	9,30E-04	7,19E-03	11	64
Ionotropic glutamate receptor pathway	PantherDB	1,82E-05	2,33E-02	1,94E-03	1,50E-02	9	46
Na_2-dependent neurotransmitter transporters	BioSystems: REACTOME	2,59E-05	3,31E-02	2,24E-03	1,73E-02	6	19
GABA synthesis, release, reuptake and degradation	BioSystems: REACTOME	2,59E-05	3,31E-02	2,24E-03	1,73E-02	6	19
Ensemble of genes encoding extracellular matrix and extracellular matrix-associated proteins	MSigDB C2 BIOCARTA (v5.1)	2,75E-05	3,51E-02	2,24E-03	1,73E-02	59	1028
Extracellular matrix organization	BioSystems: REACTOME	2,80E-05	3,58E-02	2,24E-03	1,73E-02	23	264
Glutamatergic synapse	BioSystems: KEGG	3,31E-05	4,23E-02	2,40E-03	1,85E-02	14	116
Cell adhesion molecules (CAMs)	BioSystems: KEGG	3,38E-05	4,32E-02	2,40E-03	1,85E-02	16	147
Cardiac Progenitor Differentiation	BioSystems: WikiPathways	5,95E-05	7,60E-02	4,00E-03	3,09E-02	9	53
Heart Development	BioSystems: WikiPathways	1,51E-04	1,93E-01	9,19E-03	7,11E-02	8	47
Insulin secretion	BioSystems: KEGG	1,51E-04	1,93E-01	9,19E-03	7,11E-02	11	87
GABA-B receptor II signaling	PantherDB	1,65E-04	2,10E-01	9,56E-03	7,39E-02	7	36

NCAM1 interactions	BioSystems: REACTOME	1,97E-04	2,52E-01	1,10E-02	8,47E-02	7	37
Morphine addiction	BioSystems: KEGG	2,75E-04	3,51E-01	1,46E-02	1,13E-01	11	93
Nicotine addiction	BioSystems: KEGG	3,28E-04	4,19E-01	1,67E-02	1,29E-01	7	40
Acetylcholine Neurotransmitter Release Cycle	BioSystems: REACTOME	3,44E-04	4,39E-01	1,69E-02	1,31E-01	4	11
Circadian entrainment	BioSystems: KEGG	3,98E-04	5,08E-01	1,88E-02	1,46E-01	11	97
Dopaminergic synapse	BioSystems: KEGG	4,56E-04	5,83E-01	2,08E-02	1,61E-01	13	131
Monoamine Transport	BioSystems: WikiPathways	5,98E-04	7,64E-01	2,56E-02	1,98E-01	6	32
Genes encoding collagen proteins	MSigDB C2 BIOCARTA (v5.1)	6,00E-04	7,67E-01	2,56E-02	1,98E-01	7	44
Genes encoding structural ECM glycoproteins	MSigDB C2 BIOCARTA (v5.1)	9,41E-04	1,00E+00	3,88E-02	3,00E-01	16	196
GABA synthesis	BioSystems: REACTOME	1,17E-03	1,00E+00	4,48E-02	3,46E-01	2	2
Transmembrane transport of small molecules	BioSystems: REACTOME	1,18E-03	1,00E+00	4,48E-02	3,46E-01	32	537
taurine and hypotaurine metabolic	Pathway Ontology	1,19E-03	1,00E+00	4,48E-02	3,46E-01	3	7
Maturity onset diabetes of the young	BioSystems: KEGG	1,27E-03	1,00E+00	4,53E-02	3,50E-01	5	25
GPCRs, Class C Metabotropic glutamate, pheromone	BioSystems: WikiPathways	1,28E-03	1,00E+00	4,53E-02	3,50E-01	4	15
Axon guidance	BioSystems: REACTOME	1,36E-03	1,00E+00	4,70E-02	3,63E-01	19	262
Collagen biosynthesis and modifying enzymes	BioSystems: REACTOME	1,43E-03	1,00E+00	4,81E-02	3,72E-01	8	65

B) Adenoma pathways most affected by aberrant methylations

Name	Source	p-value	q-value Bonferroni	q-value FDR B&H	q-value FDR B&Y	Hit Count in Query List	Hit Count in Genome
Cadherin signaling pathway	PantherDB	3,96E-26	7,88E-23	7,88E-23	6,45E-22	60	159
Wnt signaling pathway	PantherDB	1,44E-20	2,86E-17	1,43E-17	1,17E-16	77	305
Neuroactive ligand-receptor interaction	BioSystems: KEGG	3,73E-17	7,43E-14	2,48E-14	2,03E-13	67	275
Neuronal System	BioSystems: REACTOME	1,94E-12	3,87E-09	9,68E-10	7,91E-09	61	293
Neural Crest Differentiation	BioSystems: WikiPathways	6,20E-10	1,23E-06	2,47E-07	2,02E-06	29	101
Extracellular matrix organization	BioSystems: REACTOME	2,14E-09	4,26E-06	7,09E-07	5,80E-06	51	264
Transmission across Chemical Synapses	BioSystems: REACTOME	4,48E-09	8,91E-06	1,27E-06	1,04E-05	42	200
Calcium signaling pathway	BioSystems: KEGG	2,47E-08	4,91E-05	6,14E-06	5,02E-05	38	181
GPCR ligand binding	BioSystems: REACTOME	5,41E-07	1,08E-03	1,20E-04	9,78E-04	66	445
Nicotine addiction	BioSystems: KEGG	1,26E-06	2,51E-03	2,51E-04	2,05E-03	14	40
Cardiac Progenitor Differentiation	BioSystems: WikiPathways	2,12E-06	4,22E-03	3,84E-04	3,14E-03	16	53
G protein signaling via Galphaq family	Pathway Ontology	3,12E-06	6,22E-03	4,82E-04	3,94E-03	8	14
G alpha (q) signalling events	BioSystems: REACTOME	3,36E-06	6,68E-03	4,82E-04	3,94E-03	35	193
Voltage gated Potassium channels	BioSystems: REACTOME	3,39E-06	6,75E-03	4,82E-04	3,94E-03	14	43
Glutamatergic synapse	BioSystems: KEGG	3,70E-06	7,36E-03	4,91E-04	4,01E-03	25	116
Class A/1 (Rhodopsin-like receptors)	BioSystems: REACTOME	4,65E-06	9,25E-03	5,78E-04	4,73E-03	49	316
Ensemble of genes encoding core extracellular matrix including ECM glycoproteins, collagens and proteoglycans	MSigDB C2 BIOCARTA (v5.1)	6,27E-06	1,25E-02	7,34E-04	6,00E-03	44	275
Metabotropic glutamate receptor group III pathway	PantherDB	7,16E-06	1,43E-02	7,92E-04	6,47E-03	17	64
Morphine addiction	BioSystems: KEGG	1,02E-05	2,03E-02	1,07E-03	8,75E-03	21	93
Maturity onset diabetes of the young	BioSystems: KEGG	1,08E-05	2,16E-02	1,08E-03	8,81E-03	10	25
Developmental Biology	BioSystems: REACTOME	1,15E-05	2,29E-02	1,09E-03	8,90E-03	59	419
Peptide ligand-binding receptors	BioSystems: REACTOME	1,46E-05	2,90E-02	1,32E-03	1,08E-02	33	189
Neurotransmitter Release Cycle	BioSystems: REACTOME	1,78E-05	3,54E-02	1,54E-03	1,26E-02	12	37
Circadian entrainment	BioSystems: KEGG	2,03E-05	4,04E-02	1,68E-03	1,38E-02	21	97
GABAergic synapse	BioSystems: KEGG	2,12E-05	4,23E-02	1,69E-03	1,38E-02	20	90
Ensemble of genes encoding extracellular matrix and extracellular	MSigDB C2 BIOCARTA (v5.1)	2,28E-05	4,55E-02	1,75E-03	1,43E-02	118	1028

matrix-associated proteins							
Cholinergic synapse	BioSystems: KEGG	2,41E-05	4,79E-02	1,77E-03	1,45E-02	23	113
Gastrin-CREB signalling pathway via PKC and MAPK	BioSystems: REACTOME	3,44E-05	6,84E-02	2,44E-03	2,00E-02	35	214
Ionotropic glutamate receptor pathway	PantherDB	4,18E-05	8,33E-02	2,87E-03	2,35E-02	13	46
Collagen formation	BioSystems: REACTOME	4,39E-05	8,73E-02	2,91E-03	2,38E-02	19	87
Heart Development	BioSystems: WikiPathways	5,36E-05	1,07E-01	3,44E-03	2,82E-02	13	47
Endogenous cannabinoid signaling	PantherDB	5,54E-05	1,10E-01	3,45E-03	2,82E-02	9	24
GABA-B receptor II signaling	PantherDB	7,41E-05	1,48E-01	4,47E-03	3,66E-02	11	36
NCAM1 interactions	BioSystems: REACTOME	9,79E-05	1,95E-01	5,74E-03	4,69E-02	11	37
Amoebiasis	BioSystems: KEGG	1,24E-04	2,46E-01	7,03E-03	5,75E-02	21	109
Peptide GPCRs	BioSystems: WikiPathways	1,39E-04	2,77E-01	7,68E-03	6,28E-02	16	72
Collagen biosynthesis and modifying enzymes	BioSystems: REACTOME	1,43E-04	2,84E-01	7,68E-03	6,28E-02	15	65
Neurotransmitter Receptor Binding And Downstream Transmission In The Postsynaptic Cell	BioSystems: REACTOME	1,70E-04	3,38E-01	8,74E-03	7,15E-02	25	144
Integrin cell surface interactions	BioSystems: REACTOME	1,71E-04	3,41E-01	8,74E-03	7,15E-02	15	66
G alpha (i) signalling events	BioSystems: REACTOME	3,02E-04	6,02E-01	1,50E-02	1,23E-01	35	238
Potassium Channels	BioSystems: REACTOME	3,52E-04	7,01E-01	1,71E-02	1,40E-01	19	101
GABA synthesis, release, reuptake and degradation	BioSystems: REACTOME	4,34E-04	8,64E-01	2,06E-02	1,68E-01	7	19
Insulin secretion	BioSystems: KEGG	4,48E-04	8,92E-01	2,06E-02	1,69E-01	17	87
Retrograde endocannabinoid signaling	BioSystems: KEGG	4,56E-04	9,08E-01	2,06E-02	1,69E-01	19	103
Genes encoding collagen proteins	MSigDB C2 BIOCARTA (v5.1)	5,24E-04	1,00E+00	2,31E-02	1,89E-01	11	44
Synaptic Vesicle Pathway	BioSystems: WikiPathways	5,38E-04	1,00E+00	2,31E-02	1,89E-01	12	51
Endothelin signaling pathway	PantherDB	5,48E-04	1,00E+00	2,31E-02	1,89E-01	15	73
FOXA transcription factor networks	BioSystems: Pathway Interaction Database	5,56E-04	1,00E+00	2,31E-02	1,89E-01	3	3
GPCRs, Class C Metabotropic glutamate, pheromone	BioSystems: WikiPathways	6,82E-04	1,00E+00	2,77E-02	2,26E-01	6	15
Regulation of beta-cell development	BioSystems: REACTOME	8,49E-04	1,00E+00	3,38E-02	2,76E-01	9	33
Regulation of gene expression in beta cells	BioSystems: REACTOME	8,70E-04	1,00E+00	3,39E-02	2,78E-01	7	21
Acetylcholine Neurotransmitter Release Cycle	BioSystems: REACTOME	9,88E-04	1,00E+00	3,78E-02	3,09E-01	5	11

Elastic fibre formation	BioSystems: REACTOME	1,15E-03	1,00E+00	4,19E-02	3,43E-01	10	41
Type II diabetes mellitus	BioSystems: KEGG	1,15E-03	1,00E+00	4,19E-02	3,43E-01	11	48
Transmembrane transport of small molecules	BioSystems: REACTOME	1,18E-03	1,00E+00	4,19E-02	3,43E-01	63	537
Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway	PantherDB	1,18E-03	1,00E+00	4,19E-02	3,43E-01	19	111
Cell adhesion molecules (CAMs)	BioSystems: KEGG	1,36E-03	1,00E+00	4,74E-02	3,88E-01	23	147
Unblocking of NMDA receptor, glutamate binding and activation	BioSystems: REACTOME	1,47E-03	1,00E+00	4,95E-02	4,05E-01	6	17
Focal Adhesion	BioSystems: WikiPathways	1,47E-03	1,00E+00	4,95E-02	4,05E-01	27	184

Supporting Information Table 6:

A) genes belonging to the pathways most affected by methylation alterations in CRC

SYMBOL	ENSEMBL	GENENAME
ADAMTS2	ENSG00000087116	ADAM metallopeptidase with thrombospondin type 1 motif 2
ADAMTS5	ENSG00000154736	ADAM metallopeptidase with thrombospondin type 1 motif 5
ADCY1	ENSG00000164742	adenylate cyclase 1 (brain)
ADCY8	ENSG00000155897	adenylate cyclase 8 (brain)
ADRA1B	ENSG00000170214	adrenoceptor alpha 1B
AGRN	ENSG00000188157	agrin
APC2	ENSG00000115266	adenomatosis polyposis coli 2
BMP2	ENSG00000125845	bone morphogenetic protein 2
BRINP2	ENSG00000198797	BMP/retinoic acid inducible neural specific 2
C1QL2	ENSG00000144119	complement component 1, q subcomponent-like 2
C1QL3	ENSG00000165985	complement component 1, q subcomponent-like 3
CACNA1A	ENSG00000141837	calcium voltage-gated channel subunit alpha1 A
CACNA1B	ENSG00000148408	calcium voltage-gated channel subunit alpha1 B
CACNA1E	ENSG00000198216	calcium voltage-gated channel subunit alpha1 E
CAMK2B	ENSG00000058404	calcium/calmodulin-dependent protein kinase II beta
CBLN1	ENSG00000102924	cerebellin 1 precursor
CBLN2	ENSG00000141668	cerebellin 2 precursor
CBLN4	ENSG00000054803	cerebellin 4 precursor
CD34	ENSG00000174059	CD34 molecule
CD8A	ENSG00000153563	CD8a molecule
CDH3	ENSG00000062038	cadherin 3
CDH4	ENSG00000179242	cadherin 4
CDH4	ENSG00000280641	cadherin 4
CHAT	ENSG00000070748	choline O-acetyltransferase
CLEC14A	ENSG00000176435	C-type lectin domain family 14 member A
CNTN1	ENSG00000018236	contactin 1
CNTNAP1	ENSG00000108797	contactin associated protein 1
COL14A1	ENSG00000187955	collagen type XIV alpha 1
COL15A1	ENSG00000204291	collagen type XV alpha 1
COL23A1	ENSG00000050767	collagen type XXIII alpha 1
COL24A1	ENSG00000171502	collagen type XXIV alpha 1
COL26A1	ENSG00000160963	collagen type XXVI alpha 1
COL4A1	ENSG00000187498	collagen type IV alpha 1
COL4A2	ENSG00000134871	collagen type IV alpha 2
CPLX1	ENSG00000168993	complexin 1
CREB3L1	ENSG00000157613	cAMP responsive element binding protein 3-like 1
CREB5	ENSG00000146592	cAMP responsive element binding protein 5
CRHBP	ENSG00000145708	corticotropin releasing hormone binding protein
CRHR2	ENSG00000106113	corticotropin releasing hormone receptor 2
CRISPLD1	ENSG00000121005	cysteine rich secretory protein LCCL domain containing 1
CTBP2	ENSG00000175029	C-terminal binding protein 2
DKK2	ENSG00000155011	dickkopf WNT signaling pathway inhibitor 2

DRD5	ENSG00000169676	dopamine receptor D5
EDNRB	ENSG00000136160	endothelin receptor type B
EMILIN3	ENSG00000183798	elastin microfibril interfacer 3
EN2	ENSG00000164778	engrailed homeobox 2
FBN1	ENSG00000166147	fibrillin 1
FBN2	ENSG00000138829	fibrillin 2
FGF12	ENSG00000114279	fibroblast growth factor 12
FGF14	ENSG00000102466	fibroblast growth factor 14
FGF5	ENSG00000138675	fibroblast growth factor 5
FGF8	ENSG00000107831	fibroblast growth factor 8
FNDC1	ENSG00000164694	fibronectin type III domain containing 1
FREM3	ENSG00000183090	FRAS1 related extracellular matrix 3
GABBR2	ENSG00000136928	gamma-aminobutyric acid type B receptor subunit 2
GABRB1	ENSG00000163288	gamma-aminobutyric acid type A receptor beta1 subunit
GAD1	ENSG00000128683	glutamate decarboxylase 1
GAD2	ENSG00000136750	glutamate decarboxylase 2
GDF6	ENSG00000156466	growth differentiation factor 6
GDNF	ENSG00000168621	glial cell derived neurotrophic factor
GHSR	ENSG00000121853	growth hormone secretagogue receptor
GLRB	ENSG00000109738	glycine receptor beta
GNG7	ENSG00000176533	G protein subunit gamma 7
GPC6	ENSG00000183098	glypican 6
GPR83	ENSG00000123901	G protein-coupled receptor 83
GRIA4	ENSG00000152578	glutamate ionotropic receptor AMPA type subunit 4
GRIK3	ENSG00000163873	glutamate ionotropic receptor kainate type subunit 3
GRIK5	ENSG00000105737	glutamate ionotropic receptor kainate type subunit 5
GRIN2D	ENSG00000105464	glutamate ionotropic receptor NMDA type subunit 2D
GRIN3B	ENSG00000116032	glutamate ionotropic receptor NMDA type subunit 3B
GRM6	ENSG00000113262	glutamate receptor, metabotropic 6
GRM7	ENSG00000196277	glutamate receptor, metabotropic 7
HAPLN4	ENSG00000187664	hyaluronan and proteoglycan link protein 4
HCN1	ENSG00000164588	hyperpolarization activated cyclic nucleotide gated potassium channel 1
HPSE2	ENSG00000172987	heparanase 2 (inactive)
HRH2	ENSG00000113749	histamine receptor H2
HRNR	ENSG00000197915	hornerin
HTR1E	ENSG00000168830	5-hydroxytryptamine receptor 1E
ITGA4	ENSG00000115232	integrin subunit alpha 4
ITGA8	ENSG00000077943	integrin subunit alpha 8
JAM3	ENSG00000166086	junctional adhesion molecule 3
KCNA1	ENSG00000111262	potassium voltage-gated channel subfamily A member 1
KCNA3	ENSG00000177272	potassium voltage-gated channel subfamily A member 3
KCNH1	ENSG00000143473	potassium voltage-gated channel subfamily H member 1
KCNK7	ENSG00000173338	potassium two pore domain channel subfamily K member 7
KCNQ2	ENSG00000075043	potassium voltage-gated channel subfamily Q member 2
KCNQ2	ENSG00000281151	potassium voltage-gated channel subfamily Q member 2
KCNQ5	ENSG00000185760	potassium voltage-gated channel subfamily Q member 5
KY	ENSG00000174611	kyphoscoliosis peptidase

LAMA1	ENSG00000101680	laminin subunit alpha 1
LHCGR	ENSG00000138039	luteinizing hormone/choriogonadotropin receptor
LRRC4	ENSG00000128594	leucine rich repeat containing 4
LTBP4	ENSG00000090006	latent transforming growth factor beta binding protein 4
MMP9	ENSG00000100985	matrix metallopeptidase 9
NCAM1	ENSG00000149294	neural cell adhesion molecule 1
NCAM2	ENSG00000154654	neural cell adhesion molecule 2
NELL1	ENSG00000165973	neural EGFL like 1
NGF	ENSG00000134259	nerve growth factor
NOS1	ENSG00000089250	nitric oxide synthase 1
NPBWR1	ENSG00000183729	neuropeptides B/W receptor 1
NPBWR2	ENSG00000277339	neuropeptides B/W receptor 2
NPBWR2	ENSG00000125522	neuropeptides B/W receptor 2
NPY2R	ENSG00000185149	neuropeptide Y receptor Y2
NPY4R	ENSG00000204174	neuropeptide Y receptor Y4
NRG1	ENSG00000157168	neuregulin 1
NRG3	ENSG00000185737	neuregulin 3
NRXN2	ENSG00000110076	neurexin 2
NTSR1	ENSG00000101188	neurotensin receptor 1 (high affinity)
OPRK1	ENSG00000082556	opioid receptor, kappa 1
PCDHGA1	ENSG00000204956	protocadherin gamma subfamily A, 1
PCDHGA10	ENSG00000253846	protocadherin gamma subfamily A, 10
PCDHGA11	ENSG00000253873	protocadherin gamma subfamily A, 11
PCDHGA12	ENSG00000253159	protocadherin gamma subfamily A, 12
PCDHGA2	ENSG00000081853	protocadherin gamma subfamily A, 2
PCDHGA3	ENSG00000254245	protocadherin gamma subfamily A, 3
PCDHGA4	ENSG00000262576	protocadherin gamma subfamily A, 4
PCDHGA5	ENSG00000253485	protocadherin gamma subfamily A, 5
PCDHGA6	ENSG00000253731	protocadherin gamma subfamily A, 6
PCDHGA7	ENSG00000253537	protocadherin gamma subfamily A, 7
PCDHGA8	ENSG00000253767	protocadherin gamma subfamily A, 8
PCDHGA9	ENSG00000261934	protocadherin gamma subfamily A, 9
PCDHGB1	ENSG00000254221	protocadherin gamma subfamily B, 1
PCDHGB2	ENSG00000253910	protocadherin gamma subfamily B, 2
PCDHGB3	ENSG00000262209	protocadherin gamma subfamily B, 3
PCDHGB4	ENSG00000253953	protocadherin gamma subfamily B, 4
PCDHGB5	ENSG00000276547	protocadherin gamma subfamily B, 5
PCDHGB6	ENSG00000253305	protocadherin gamma subfamily B, 6
PCDHGB7	ENSG00000254122	protocadherin gamma subfamily B, 7
PCDHGC3	ENSG00000240184	protocadherin gamma subfamily C, 3
PCDHGC4	ENSG00000242419	protocadherin gamma subfamily C, 4
PCDHGC5	ENSG00000240764	protocadherin gamma subfamily C, 5
PDCD1	ENSG00000188389	programmed cell death 1
PDCD1	ENSG00000276977	programmed cell death 1
PDE1C	ENSG00000154678	phosphodiesterase 1C
PDGFD	ENSG00000170962	platelet derived growth factor D
PLCB1	ENSG00000182621	phospholipase C beta 1

PPP2R5C	ENSG00000078304	protein phosphatase 2 regulatory subunit B', gamma
PRKAR1B	ENSG00000188191	protein kinase cAMP-dependent type I regulatory subunit beta
PRKCB	ENSG00000166501	protein kinase C beta
PTGFR	ENSG00000122420	prostaglandin F receptor
PTPRS	ENSG00000105426	protein tyrosine phosphatase, receptor type S
PXDN	ENSG00000130508	peroxidasin
RELN	ENSG00000189056	reelin
RSPO2	ENSG00000147655	R-spondin 2
RSPO3	ENSG00000146374	R-spondin 3
RYR2	ENSG00000198626	ryanodine receptor 2
S1PR5	ENSG00000180739	sphingosine-1-phosphate receptor 5
SCTR	ENSG00000080293	secretin receptor
SDC2	ENSG00000169439	syndecan 2
SEMA6B	ENSG00000167680	semaphorin 6B
SFRP1	ENSG00000104332	secreted frizzled-related protein 1
SFRP2	ENSG00000145423	secreted frizzled-related protein 2
SHANK1	ENSG00000161681	SH3 and multiple ankyrin repeat domains 1
SLC12A5	ENSG00000124140	solute carrier family 12 member 5
SLC18A3	ENSG00000187714	solute carrier family 18 member A3
SLC32A1	ENSG00000101438	solute carrier family 32 member 1
SLC5A7	ENSG00000115665	solute carrier family 5 member 7
SLC6A1	ENSG00000157103	solute carrier family 6 member 1
SLC6A11	ENSG00000132164	solute carrier family 6 member 11
SLC6A15	ENSG00000072041	solute carrier family 6 member 15
SLC6A2	ENSG00000103546	solute carrier family 6 member 2
SLC6A3	ENSG00000142319	solute carrier family 6 member 3
SLC6A3	ENSG00000276996	solute carrier family 6 member 3
SLC6A5	ENSG00000165970	solute carrier family 6 member 5
SLC8A1	ENSG00000183023	solute carrier family 8 member A1
SLIT2	ENSG00000145147	slit guidance ligand 2
SLIT3	ENSG00000184347	slit guidance ligand 3
SPOCK1	ENSG00000152377	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1
STX4	ENSG00000103496	syntaxin 4
SYN3	ENSG00000185666	synapsin III
T	ENSG00000164458	T brachyury transcription factor
TSPEAR	ENSG00000175894	thrombospondin-type laminin G domain and EAR repeats
VCAN	ENSG00000038427	versican
VEGFC	ENSG00000150630	vascular endothelial growth factor C
VIPR2	ENSG00000106018	vasoactive intestinal peptide receptor 2
VWC2	ENSG00000188730	von Willebrand factor C domain containing 2

B) genes belonging to the pathways most affected by methylation alterations in adenoma

SYMBOL	ENSEMBL	GENENAME
ABLIM2	ENSG00000163995	actin binding LIM protein family member 2
ACAN	ENSG00000157766	aggrecan
ACTA1	ENSG00000143632	actin, alpha 1, skeletal muscle
ACTN2	ENSG00000077522	actinin alpha 2
ACVR1C	ENSG00000123612	activin A receptor type 1C
ADAM30	ENSG00000134249	ADAM metallopeptidase domain 30
ADAMTS10	ENSG00000142303	ADAM metallopeptidase with thrombospondin type 1 motif 10
ADAMTS12	ENSG00000151388	ADAM metallopeptidase with thrombospondin type 1 motif 12
ADAMTS12	ENSG00000281690	ADAM metallopeptidase with thrombospondin type 1 motif 12
ADAMTS16	ENSG00000145536	ADAM metallopeptidase with thrombospondin type 1 motif 16
ADAMTS2	ENSG00000087116	ADAM metallopeptidase with thrombospondin type 1 motif 2
ADAMTS20	ENSG00000173157	ADAM metallopeptidase with thrombospondin type 1 motif 20
ADAMTS5	ENSG00000154736	ADAM metallopeptidase with thrombospondin type 1 motif 5
ADAMTSL3	ENSG00000156218	ADAMTS like 3
ADCY8	ENSG00000155897	adenylate cyclase 8 (brain)
ADCY9	ENSG00000162104	adenylate cyclase 9
ADCYAP1	ENSG00000141433	adenylate cyclase activating polypeptide 1 (pituitary)
ADCYAP1R1	ENSG00000078549	adenylate cyclase activating polypeptide 1 (pituitary) receptor type I
ADRA1A	ENSG00000120907	adrenoceptor alpha 1A
ADRA1B	ENSG00000170214	adrenoceptor alpha 1B
AGTR1	ENSG00000144891	angiotensin II receptor type 1
ANK1	ENSG00000029534	ankyrin 1
APC2	ENSG00000115266	adenomatosis polyposis coli 2
ARHGEF25	ENSG00000240771	Rho guanine nucleotide exchange factor 25
ARTN	ENSG00000117407	artemin
ATP2A3	ENSG00000074370	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3
AVPR1A	ENSG00000166148	arginine vasopressin receptor 1A
AXIN2	ENSG00000168646	axin 2
BCAN	ENSG00000132692	brevican
BMP2	ENSG00000125845	bone morphogenetic protein 2
BMP4	ENSG00000125378	bone morphogenetic protein 4
BMP7	ENSG00000101144	bone morphogenetic protein 7
BRINP2	ENSG00000198797	BMP/retinoic acid inducible neural specific 2
C1QL1	ENSG00000131094	complement component 1, q subcomponent-like 1
C1QL2	ENSG00000144119	complement component 1, q subcomponent-like 2
C1QL3	ENSG00000165985	complement component 1, q subcomponent-like 3
C1QTNF8	ENSG00000184471	C1q and tumor necrosis factor related protein 8
CACNA1A	ENSG00000141837	calcium voltage-gated channel subunit alpha1 A
CACNA1B	ENSG00000148408	calcium voltage-gated channel subunit alpha1 B
CACNA1E	ENSG00000198216	calcium voltage-gated channel subunit alpha1 E
CACNA1G	ENSG00000006283	calcium voltage-gated channel subunit alpha1 G
CACNA1H	ENSG00000196557	calcium voltage-gated channel subunit alpha1 H
CACNA1I	ENSG00000100346	calcium voltage-gated channel subunit alpha1 I
CACNB2	ENSG00000165995	calcium voltage-gated channel auxiliary subunit beta 2

CACNG8	ENSG00000142408	calcium voltage-gated channel auxiliary subunit gamma 8
CAPNS1	ENSG00000126247	calpain small subunit 1
CASR	ENSG00000036828	calcium sensing receptor
CBLN1	ENSG00000102924	cerebellin 1 precursor
CBLN2	ENSG00000141668	cerebellin 2 precursor
CBLN4	ENSG00000054803	cerebellin 4 precursor
CD38	ENSG00000004468	CD38 molecule
CDH1	ENSG00000039068	cadherin 1
CDH2	ENSG00000170558	cadherin 2
CDH4	ENSG00000179242	cadherin 4
CDH4	ENSG00000280641	cadherin 4
CDH8	ENSG00000150394	cadherin 8
CELSR1	ENSG00000075275	cadherin EGF LAG seven-pass G-type receptor 1
CELSR3	ENSG00000008300	cadherin EGF LAG seven-pass G-type receptor 3
CHAD	ENSG00000136457	chondroadherin
CHAT	ENSG00000070748	choline O-acetyltransferase
CHL1	ENSG00000134121	cell adhesion molecule L1 like
CHRDL2	ENSG00000054938	chordin-like 2
CHRM2	ENSG00000181072	cholinergic receptor muscarinic 2
CHRNA4	ENSG00000101204	cholinergic receptor nicotinic alpha 4 subunit
CLEC14A	ENSG00000176435	C-type lectin domain family 14 member A
CNTN1	ENSG00000018236	contactin 1
CNTNAP1	ENSG00000108797	contactin associated protein 1
COL20A1	ENSG00000101203	collagen type XX alpha 1
COL21A1	ENSG00000124749	collagen type XXI alpha 1
COL23A1	ENSG00000050767	collagen type XXIII alpha 1
COL25A1	ENSG00000188517	collagen type XXV alpha 1
COL26A1	ENSG00000160963	collagen type XXVI alpha 1
COL2A1	ENSG00000139219	collagen type II alpha 1
COL4A1	ENSG00000187498	collagen type IV alpha 1
COL4A2	ENSG00000134871	collagen type IV alpha 2
COL5A1	ENSG00000130635	collagen type V alpha 1
COL5A3	ENSG00000080573	collagen type V alpha 3
COL9A1	ENSG00000112280	collagen type IX alpha 1
CPLX1	ENSG00000168993	complexin 1
CREB3L1	ENSG00000157613	cAMP responsive element binding protein 3-like 1
CREB5	ENSG00000146592	cAMP responsive element binding protein 5
CRHBP	ENSG00000145708	corticotropin releasing hormone binding protein
CRHR2	ENSG00000106113	corticotropin releasing hormone receptor 2
CRISPLD1	ENSG00000121005	cysteine rich secretory protein LCCL domain containing 1
CRMP1	ENSG00000072832	collapsin response mediator protein 1
CSNK1G2	ENSG00000133275	casein kinase 1 gamma 2
CST3	ENSG00000101439	cystatin C
CTNNA2	ENSG00000066032	catenin alpha 2
CTSL	ENSG00000135047	cathepsin L
CXCL2	ENSG00000081041	C-X-C motif chemokine ligand 2
DGKG	ENSG00000058866	diacylglycerol kinase gamma

DGKI	ENSG00000157680	diacylglycerol kinase iota
DGKQ	ENSG00000145214	diacylglycerol kinase theta
DKK2	ENSG00000155011	dickkopf WNT signaling pathway inhibitor 2
DLX5	ENSG00000105880	distal-less homeobox 5
DNM1	ENSG00000106976	dynamin 1
DOCK1	ENSG00000150760	dedicator of cytokinesis 1
EBF1	ENSG00000164330	early B-cell factor 1
EDN3	ENSG00000124205	endothelin 3
EDNRB	ENSG00000136160	endothelin receptor type B
EFEMP1	ENSG00000115380	EGF containing fibulin-like extracellular matrix protein 1
EGFLAM	ENSG00000164318	EGF like, fibronectin type III and laminin G domains
EGLN2	ENSG00000269858	egl-9 family hypoxia-inducible factor 2
EGLN3	ENSG00000129521	egl-9 family hypoxia-inducible factor 3
ELFN1	ENSG00000225968	extracellular leucine-rich repeat and fibronectin type III domain containing 1
EMILIN1	ENSG00000138080	elastin microfibril interfacer 1
EMILIN3	ENSG00000183798	elastin microfibril interfacer 3
EN1	ENSG00000163064	engrailed homeobox 1
EN2	ENSG00000164778	engrailed homeobox 2
ERBB2	ENSG00000141736	erb-b2 receptor tyrosine kinase 2
F2RL1	ENSG00000164251	F2R like trypsin receptor 1
F7	ENSG00000057593	coagulation factor VII
FAT1	ENSG00000083857	FAT atypical cadherin 1
FBLN7	ENSG00000144152	fibulin 7
FBN1	ENSG00000166147	fibrillin 1
FBN2	ENSG00000138829	fibrillin 2
FFAR1	ENSG00000126266	free fatty acid receptor 1
FGF12	ENSG00000114279	fibroblast growth factor 12
FGF14	ENSG00000102466	fibroblast growth factor 14
FGF17	ENSG00000158815	fibroblast growth factor 17
FGF3	ENSG00000186895	fibroblast growth factor 3
FGF5	ENSG00000138675	fibroblast growth factor 5
FGF8	ENSG00000107831	fibroblast growth factor 8
FNDC1	ENSG00000164694	fibronectin type III domain containing 1
FOXA2	ENSG00000125798	forkhead box A2
FOXA3	ENSG00000170608	forkhead box A3
FREM3	ENSG00000183090	FRAS1 related extracellular matrix 3
FZD10	ENSG00000111432	frizzled class receptor 10
FZD2	ENSG00000180340	frizzled class receptor 2
GABBR1	ENSG00000204681	gamma-aminobutyric acid type B receptor subunit 1
GABBR1	ENSG00000206466	gamma-aminobutyric acid type B receptor subunit 1
GABBR1	ENSG00000232632	gamma-aminobutyric acid type B receptor subunit 1
GABBR1	ENSG00000237051	gamma-aminobutyric acid type B receptor subunit 1
GABBR1	ENSG00000232569	gamma-aminobutyric acid type B receptor subunit 1
GABBR1	ENSG00000237112	gamma-aminobutyric acid type B receptor subunit 1
GABBR1	ENSG00000206511	gamma-aminobutyric acid type B receptor subunit 1
GABRA2	ENSG00000151834	gamma-aminobutyric acid type A receptor alpha2 subunit
GABRA5	ENSG00000186297	gamma-aminobutyric acid type A receptor alpha5 subunit

GABRB3	ENSG00000166206	gamma-aminobutyric acid type A receptor beta3 subunit
GABRG3	ENSG00000182256	gamma-aminobutyric acid type A receptor gamma3 subunit
GAD1	ENSG00000128683	glutamate decarboxylase 1
GAD2	ENSG00000136750	glutamate decarboxylase 2
GALR1	ENSG00000166573	galanin receptor 1
GATA4	ENSG00000136574	GATA binding protein 4
GDF1	ENSG00000130283	growth differentiation factor 1
GDF1	ENSG00000223802	growth differentiation factor 1
GDF6	ENSG00000156466	growth differentiation factor 6
GDNF	ENSG00000168621	glial cell derived neurotrophic factor
GFRA1	ENSG00000151892	GDNF family receptor alpha 1
GFRA4	ENSG00000125861	GDNF family receptor alpha 4
GHSR	ENSG00000121853	growth hormone secretagogue receptor
GJD2	ENSG00000159248	gap junction protein delta 2
GLRA3	ENSG00000145451	glycine receptor alpha 3
GLRB	ENSG00000109738	glycine receptor beta
GNA11	ENSG00000088256	G protein subunit alpha 11
GNAI1	ENSG00000127955	G protein subunit alpha i1
GNAL	ENSG00000141404	G protein subunit alpha L
GNAO1	ENSG00000087258	G protein subunit alpha o1
GNAS	ENSG00000087460	GNAS complex locus
GNB4	ENSG00000114450	G protein subunit beta 4
GNG7	ENSG00000176533	G protein subunit gamma 7
GPC6	ENSG00000183098	glypican 6
GPR39	ENSG00000183840	G protein-coupled receptor 39
GRIA2	ENSG00000120251	glutamate ionotropic receptor AMPA type subunit 2
GRIA4	ENSG00000152578	glutamate ionotropic receptor AMPA type subunit 4
GRID1	ENSG00000182771	glutamate ionotropic receptor delta type subunit 1
GRID2	ENSG00000152208	glutamate ionotropic receptor delta type subunit 2
GRIK1	ENSG00000171189	glutamate ionotropic receptor kainate type subunit 1
GRIK2	ENSG00000164418	glutamate ionotropic receptor kainate type subunit 2
GRIK3	ENSG00000163873	glutamate ionotropic receptor kainate type subunit 3
GRIN2A	ENSG00000183454	glutamate ionotropic receptor NMDA type subunit 2A
GRIN2D	ENSG00000105464	glutamate ionotropic receptor NMDA type subunit 2D
GRIN3B	ENSG00000116032	glutamate ionotropic receptor NMDA type subunit 3B
GRM1	ENSG00000152822	glutamate receptor, metabotropic 1
GRM6	ENSG00000113262	glutamate receptor, metabotropic 6
GRM7	ENSG00000196277	glutamate receptor, metabotropic 7
GRP	ENSG00000134443	gastrin releasing peptide
GUCY1B3	ENSG00000061918	guanylate cyclase 1, soluble, beta 3
HAND1	ENSG00000113196	heart and neural crest derivatives expressed 1
HAPLN4	ENSG00000187664	hyaluronan and proteoglycan link protein 4
HCN2	ENSG00000099822	hyperpolarization activated cyclic nucleotide gated potassium channel 2
HCRTR2	ENSG00000137252	hypocretin receptor 2
HEBP1	ENSG00000013583	heme binding protein 1
HPSE2	ENSG00000172987	heparanase 2 (inactive)
HRH2	ENSG00000113749	histamine receptor H2

HTR1A	ENSG00000178394	5-hydroxytryptamine receptor 1A
HTR1E	ENSG00000168830	5-hydroxytryptamine receptor 1E
HTR5A	ENSG00000157219	5-hydroxytryptamine receptor 5A
IRX4	ENSG00000113430	iroquois homeobox 4
ISL1	ENSG00000016082	ISL LIM homeobox 1
ITGA4	ENSG00000115232	integrin subunit alpha 4
ITGA8	ENSG00000077943	integrin subunit alpha 8
ITGAD	ENSG00000156886	integrin subunit alpha D
ITGAM	ENSG00000169896	integrin subunit alpha M
ITGAX	ENSG00000140678	integrin subunit alpha X
JAM2	ENSG00000154721	junctional adhesion molecule 2
JAM3	ENSG00000166086	junctional adhesion molecule 3
KAZALD1	ENSG00000107821	Kazal type serine peptidase inhibitor domain 1
KCNA1	ENSG00000111262	potassium voltage-gated channel subfamily A member 1
KCNA3	ENSG00000177272	potassium voltage-gated channel subfamily A member 3
KCNA4	ENSG00000182255	potassium voltage-gated channel subfamily A member 4
KCNA5	ENSG00000130037	potassium voltage-gated channel subfamily A member 5
KCNA6	ENSG00000130035	potassium voltage-gated channel subfamily A member 6
KCNA6	ENSG00000151079	potassium voltage-gated channel subfamily A member 6
KCNAB1	ENSG00000169282	potassium voltage-gated channel subfamily A member regulatory beta subunit 1
KCNAB2	ENSG00000069424	potassium voltage-gated channel subfamily A regulatory beta subunit 2
KCNAB3	ENSG00000170049	potassium voltage-gated channel subfamily A regulatory beta subunit 3
KCNC2	ENSG00000166006	potassium voltage-gated channel subfamily C member 2
KCNH5	ENSG00000140015	potassium voltage-gated channel subfamily H member 5
KCNK1	ENSG00000135750	potassium two pore domain channel subfamily K member 1
KCNN1	ENSG00000105642	potassium calcium-activated channel subfamily N member 1
KCNQ4	ENSG00000117013	potassium voltage-gated channel subfamily Q member 4
KCNQ5	ENSG00000185760	potassium voltage-gated channel subfamily Q member 5
KCNS1	ENSG00000124134	potassium voltage-gated channel modifier subfamily S member 1
KCNV1	ENSG00000164794	potassium voltage-gated channel modifier subfamily V member 1
KDR	ENSG00000128052	kinase insert domain receptor
KITLG	ENSG00000049130	KIT ligand
KY	ENSG00000174611	kyphoscoliosis peptidase
LAMA1	ENSG00000101680	laminin subunit alpha 1
LAMB1	ENSG00000091136	laminin subunit beta 1
LAMB3	ENSG00000196878	laminin subunit beta 3
LEFTY2	ENSG00000143768	left-right determination factor 2
LEP	ENSG00000174697	leptin
LEPR	ENSG00000116678	leptin receptor
LHCGR	ENSG00000138039	luteinizing hormone/choriogonadotropin receptor
LHX5	ENSG00000089116	LIM homeobox 5
LIN28B	ENSG00000187772	lin-28 homolog B
LPAR1	ENSG00000198121	lysophosphatidic acid receptor 1
LTBP3	ENSG00000168056	latent transforming growth factor beta binding protein 3
LTBP4	ENSG00000090006	latent transforming growth factor beta binding protein 4
MADCAM1	ENSG00000099866	mucosal vascular addressin cell adhesion molecule 1

MATN3	ENSG00000132031	matrilin 3
MBP	ENSG00000197971	myelin basic protein
MCHR2	ENSG00000152034	melanin concentrating hormone receptor 2
MED16	ENSG00000175221	mediator complex subunit 16
MED16	ENSG00000282092	mediator complex subunit 16
MED30	ENSG00000164758	mediator complex subunit 30
MLNR	ENSG00000102539	motilin receptor
MMP16	ENSG00000156103	matrix metallopeptidase 16
MMP9	ENSG00000100985	matrix metallopeptidase 9
MNX1	ENSG00000130675	motor neuron and pancreas homeobox 1
MSX1	ENSG00000163132	msh homeobox 1
MSX2	ENSG00000120149	msh homeobox 2
MTNR1B	ENSG00000134640	melatonin receptor 1B
MYF5	ENSG00000111049	myogenic factor 5
MYF6	ENSG00000111046	myogenic factor 6
MYH14	ENSG00000105357	myosin, heavy chain 14, non-muscle
MYH9	ENSG00000100345	myosin, heavy chain 9, non-muscle
MYOD1	ENSG00000129152	myogenic differentiation 1
NCOR2	ENSG00000196498	nuclear receptor corepressor 2
NEFL	ENSG00000277586	neurofilament, light polypeptide
NELL1	ENSG00000165973	neural EGFL like 1
NEUROD1	ENSG00000162992	neuronal differentiation 1
NKD1	ENSG00000140807	naked cuticle homolog 1 (Drosophila)
NKX2-2	ENSG00000125820	NK2 homeobox 2
NKX2-5	ENSG00000183072	NK2 homeobox 5
NMBR	ENSG00000135577	neuromedin B receptor
NOS1	ENSG00000089250	nitric oxide synthase 1
NOS3	ENSG00000164867	nitric oxide synthase 3
NOV	ENSG00000136999	nephroblastoma overexpressed
NPBWR1	ENSG00000183729	neuropeptides B/W receptor 1
NPFFR2	ENSG00000056291	neuropeptide FF receptor 2
NPY	ENSG00000122585	neuropeptide Y
NPY2R	ENSG00000185149	neuropeptide Y receptor Y2
NPY4R	ENSG00000204174	neuropeptide Y receptor Y4
NR2F2	ENSG00000185551	nuclear receptor subfamily 2 group F member 2
NR5A2	ENSG00000116833	nuclear receptor subfamily 5 group A member 2
NRG1	ENSG00000157168	neuregulin 1
NRG3	ENSG00000185737	neuregulin 3
NRXN1	ENSG00000179915	neurexin 1
NTF3	ENSG00000185652	neurotrophin 3
NTSR2	ENSG00000169006	neurotensin receptor 2
OLIG2	ENSG00000205927	oligodendrocyte lineage transcription factor 2
OLIG3	ENSG00000177468	oligodendrocyte transcription factor 3
ONECUT1	ENSG00000169856	one cut homeobox 1
OPRK1	ENSG00000082556	opioid receptor, kappa 1
OPRM1	ENSG00000112038	opioid receptor, mu 1
ORAI3	ENSG00000175938	ORAI calcium release-activated calcium modulator 3

P2RX2	ENSG00000187848	purinergic receptor P2X 2
P2RY6	ENSG00000171631	pyrimidinergic receptor P2Y6
P4HA3	ENSG00000149380	prolyl 4-hydroxylase subunit alpha 3
P4HB	ENSG00000185624	prolyl 4-hydroxylase subunit beta
PAX3	ENSG00000135903	paired box 3
PAX6	ENSG00000007372	paired box 6
PAX7	ENSG00000009709	paired box 7
PCDH1	ENSG00000156453	protocadherin 1
PCDH10	ENSG00000138650	protocadherin 10
PCDH8	ENSG00000136099	protocadherin 8
PCDH9	ENSG00000184226	protocadherin 9
PCDHA1	ENSG00000204970	protocadherin alpha 1
PCDHA10	ENSG00000250120	protocadherin alpha 10
PCDHA11	ENSG00000249158	protocadherin alpha 11
PCDHA12	ENSG00000251664	protocadherin alpha 12
PCDHA13	ENSG00000239389	protocadherin alpha 13
PCDHA2	ENSG00000204969	protocadherin alpha 2
PCDHA3	ENSG00000255408	protocadherin alpha 3
PCDHA4	ENSG00000204967	protocadherin alpha 4
PCDHA5	ENSG00000204965	protocadherin alpha 5
PCDHA6	ENSG00000081842	protocadherin alpha 6
PCDHA7	ENSG00000204963	protocadherin alpha 7
PCDHA8	ENSG00000204962	protocadherin alpha 8
PCDHA9	ENSG00000204961	protocadherin alpha 9
PCDHB12	ENSG00000120328	protocadherin beta 12
PCDHB13	ENSG00000187372	protocadherin beta 13
PCDHB15	ENSG00000113248	protocadherin beta 15
PCDHB16	ENSG00000272674	protocadherin beta 16
PCDHB4	ENSG00000081818	protocadherin beta 4
PCDHB5	ENSG00000113209	protocadherin beta 5
PCDHB6	ENSG00000113211	protocadherin beta 6
PCDHGA1	ENSG00000204956	protocadherin gamma subfamily A, 1
PCDHGA10	ENSG00000253846	protocadherin gamma subfamily A, 10
PCDHGA11	ENSG00000253873	protocadherin gamma subfamily A, 11
PCDHGA12	ENSG00000253159	protocadherin gamma subfamily A, 12
PCDHGA2	ENSG00000081853	protocadherin gamma subfamily A, 2
PCDHGA3	ENSG00000254245	protocadherin gamma subfamily A, 3
PCDHGA4	ENSG00000262576	protocadherin gamma subfamily A, 4
PCDHGA5	ENSG00000253485	protocadherin gamma subfamily A, 5
PCDHGA6	ENSG00000253731	protocadherin gamma subfamily A, 6
PCDHGA7	ENSG00000253537	protocadherin gamma subfamily A, 7
PCDHGA8	ENSG00000253767	protocadherin gamma subfamily A, 8
PCDHGA9	ENSG00000261934	protocadherin gamma subfamily A, 9
PCDHGB1	ENSG00000254221	protocadherin gamma subfamily B, 1
PCDHGB2	ENSG00000253910	protocadherin gamma subfamily B, 2
PCDHGB3	ENSG00000262209	protocadherin gamma subfamily B, 3
PCDHGB4	ENSG00000253953	protocadherin gamma subfamily B, 4

PCDHGB5	ENSG00000276547	protocadherin gamma subfamily B, 5
PCDHGB6	ENSG00000253305	protocadherin gamma subfamily B, 6
PCDHGB7	ENSG00000254122	protocadherin gamma subfamily B, 7
PCDHGC3	ENSG00000240184	protocadherin gamma subfamily C, 3
PCDHGC4	ENSG00000242419	protocadherin gamma subfamily C, 4
PCDHGC5	ENSG00000240764	protocadherin gamma subfamily C, 5
PCSK5	ENSG00000099139	proprotein convertase subtilisin/kexin type 5
PDE11A	ENSG00000128655	phosphodiesterase 11A
PDE4C	ENSG00000105650	phosphodiesterase 4C
PDE4D	ENSG00000113448	phosphodiesterase 4D
PDE8B	ENSG00000113231	phosphodiesterase 8B
PDGFD	ENSG00000170962	platelet derived growth factor D
PDX1	ENSG00000139515	pancreatic and duodenal homeobox 1
PENK	ENSG00000181195	proenkephalin
PHOX2B	ENSG00000109132	paired like homeobox 2b
PIK3CD	ENSG00000171608	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta
PIK3R2	ENSG00000105647	phosphoinositide-3-kinase regulatory subunit 2
PIK3R5	ENSG00000141506	phosphoinositide-3-kinase regulatory subunit 5
PIP5K1C	ENSG00000186111	phosphatidylinositol-4-phosphate 5-kinase, type I, gamma
PLCB1	ENSG00000182621	phospholipase C beta 1
PLD1	ENSG00000075651	phospholipase D1
PLEC	ENSG00000178209	plectin
PLOD2	ENSG00000152952	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2
PLXDC2	ENSG00000120594	plexin domain containing 2
PMP22	ENSG00000109099	peripheral myelin protein 22
PPP2R5C	ENSG00000078304	protein phosphatase 2 regulatory subunit B', gamma
PRKAR1B	ENSG00000188191	protein kinase cAMP-dependent type I regulatory subunit beta
PRKCB	ENSG00000166501	protein kinase C beta
PRKCZ	ENSG00000067606	protein kinase C zeta
PRLHR	ENSG00000119973	prolactin releasing hormone receptor
PROKR2	ENSG00000101292	prokineticin receptor 2
PRSS3	ENSG00000010438	protease, serine 3
PTGFR	ENSG00000122420	prostaglandin F receptor
PTH2R	ENSG00000144407	parathyroid hormone 2 receptor
PXDNL	ENSG00000147485	peroxidasin like
PYGO1	ENSG00000171016	pygopus family PHD finger 1
QRFPR	ENSG00000186867	pyroglutamylated RFamide peptide receptor
RASGRF1	ENSG00000058335	Ras protein specific guanine nucleotide releasing factor 1
ROBO1	ENSG00000169855	roundabout guidance receptor 1
ROR2	ENSG00000169071	receptor tyrosine kinase-like orphan receptor 2
RSPO2	ENSG00000147655	R-spondin 2
RSPO3	ENSG00000146374	R-spondin 3
RXFP3	ENSG00000277069	relaxin/insulin-like family peptide receptor 3
RXFP3	ENSG00000182631	relaxin/insulin-like family peptide receptor 3
RYR3	ENSG00000198838	ryanodine receptor 3
S1PR1	ENSG00000170989	sphingosine-1-phosphate receptor 1
S1PR3	ENSG00000213694	sphingosine-1-phosphate receptor 3

SCN8A	ENSG00000196876	sodium voltage-gated channel alpha subunit 8
SCTR	ENSG00000080293	secretin receptor
SDC2	ENSG00000169439	syndecan 2
SEMA3C	ENSG00000075223	semaphorin 3C
SEMA6B	ENSG00000167680	semaphorin 6B
SERPINB9	ENSG00000170542	serpin peptidase inhibitor, clade B (ovalbumin), member 9
SFRP1	ENSG00000104332	secreted frizzled-related protein 1
SFRP2	ENSG00000145423	secreted frizzled-related protein 2
SHANK1	ENSG00000161681	SH3 and multiple ankyrin repeat domains 1
SHANK2	ENSG00000162105	SH3 and multiple ankyrin repeat domains 2
SLC12A5	ENSG00000124140	solute carrier family 12 member 5
SLC17A6	ENSG00000091664	solute carrier family 17 member 6
SLC18A3	ENSG00000187714	solute carrier family 18 member A3
SLC2A2	ENSG00000163581	solute carrier family 2 member 2
SLC32A1	ENSG00000101438	solute carrier family 32 member 1
SLC5A7	ENSG00000115665	solute carrier family 5 member 7
SLC6A1	ENSG00000157103	solute carrier family 6 member 1
SLC6A11	ENSG00000132164	solute carrier family 6 member 11
SLC6A3	ENSG00000142319	solute carrier family 6 member 3
SLC6A3	ENSG00000276996	solute carrier family 6 member 3
SLC8A1	ENSG00000183023	solute carrier family 8 member A1
SLC8A2	ENSG00000118160	solute carrier family 8 member A2
SLIT2	ENSG00000145147	slit guidance ligand 2
SLIT3	ENSG00000184347	slit guidance ligand 3
SNAI2	ENSG00000019549	snail family zinc finger 2
SNAP25	ENSG00000132639	synaptosome associated protein 25kDa
SOX10	ENSG00000100146	SRY-box 10
SOX17	ENSG00000164736	SRY-box 17
SOX5	ENSG00000134532	SRY-box 5
SPOCK1	ENSG00000152377	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1
SPOCK3	ENSG00000196104	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3
SST	ENSG00000157005	somatostatin
SSTR1	ENSG00000139874	somatostatin receptor 1
SSTR2	ENSG00000180616	somatostatin receptor 2
SSTR4	ENSG00000132671	somatostatin receptor 4
STX2	ENSG00000111450	syntaxin 2
SYN2	ENSG00000157152	synapsin II
SYN3	ENSG00000185666	synapsin III
TACR3	ENSG00000169836	tachykinin receptor 3
TAS1R3	ENSG00000169962	taste 1 receptor member 3
TBX20	ENSG00000164532	T-box 20
TBX5	ENSG00000089225	T-box 5
TCF4	ENSG00000196628	transcription factor 4
TDGF1	ENSG00000241186	teratocarcinoma-derived growth factor 1
TFAP2A	ENSG00000137203	transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha)
TFAP2B	ENSG00000008196	transcription factor AP-2 beta (activating enhancer binding protein 2 beta)
TGFB2	ENSG00000092969	transforming growth factor beta 2

TGFBI	ENSG00000120708	transforming growth factor beta induced
THBS4	ENSG00000113296	thrombospondin 4
THRA	ENSG00000126351	thyroid hormone receptor, alpha
THY1	ENSG00000154096	Thy-1 cell surface antigen
TIMP4	ENSG00000157150	TIMP metallopeptidase inhibitor 4
TLL1	ENSG00000038295	tolloid like 1
TLX2	ENSG00000115297	T-cell leukemia homeobox 2
TNFSF11	ENSG00000120659	tumor necrosis factor superfamily member 11
TNFSF9	ENSG00000125657	tumor necrosis factor superfamily member 9
TNNC2	ENSG00000101470	troponin C2, fast skeletal type
TNNI3	ENSG00000129991	troponin I3, cardiac type
TNR	ENSG00000116147	tenascin R
TRH	ENSG00000170893	thyrotropin releasing hormone
TRPC4	ENSG00000133107	transient receptor potential cation channel subfamily C member 4
TRPC6	ENSG00000137672	transient receptor potential cation channel subfamily C member 6
TSPEAR	ENSG00000175894	thrombospondin-type laminin G domain and EAR repeats
TWIST1	ENSG00000122691	twist family bHLH transcription factor 1
UNC5C	ENSG00000182168	unc-5 netrin receptor C
UNC5D	ENSG00000156687	unc-5 netrin receptor D
VCAN	ENSG00000038427	versican
VEGFC	ENSG00000150630	vascular endothelial growth factor C
VIPR2	ENSG00000106018	vasoactive intestinal peptide receptor 2
WIF1	ENSG00000156076	WNT inhibitory factor 1
WNT3	ENSG00000108379	wingless-type MMTV integration site family member 3
WNT3	ENSG00000277626	wingless-type MMTV integration site family member 3
WNT3	ENSG00000277641	wingless-type MMTV integration site family member 3
WNT5B	ENSG00000111186	wingless-type MMTV integration site family member 5B
ZIC1	ENSG00000152977	Zic family member 1

Supporting Information Table 7: Early biomarkers panel of 74 CpG islands putative genes regulators significantly altered in CRC and adenomas

	Discovery set							In silico validation set TCGA						
Gene	**Adenomas**	**Controls**	**Adenomas Δβ value**	**CRC**	**Peritumoral**	**CRC Δβ value**	**CRC Adenomas ΔΔβ value**	**AUC**	**Specificity**	**Sensitivity**	**comb.p.value Discovery.set**	**comb.p.value GSE48684**	**comb.p.value TCGA**	**Subcellular locations**
ADAMTS2	0,35	0,12	0,22	0,52	0,1	0,42	0,2	0.96	0.98	0.91	<10E-8	<10E-8	<10E-8	E
ADAMTS5	0,49	0,25	0,24	0,55	0,21	0,33	0,09	0.97	0.95	0.92	<10E-8	<10E-8	<10E-8	E, ER
ADCY8	0,49	0,29	0,2	0,6	0,35	0,25	0,05	0.98	0.98	0.89	<10E-8	<10E-8	<10E-8	P, C
ADRA1B	0,33	0,14	0,18	0,48	0,15	0,33	0,15	0.97	0.95	0.94	4E-08	<10E-8	<10E-8	P, N
CACNA1A	0,48	0,24	0,24	0,58	0,21	0,38	0,13	1	0.95	0.97	<10E-8	<10E-8	<10E-8	P, N
CACNA1E	0,51	0,29	0,22	0,59	0,35	0,25	0,03	0.96	1	0.86	<10E-8	<10E-8	<10E-8	P
CDH4	0,48	0,17	0,3	0,64	0,15	0,49	0,19	0.96	1	0.93	<10E-8	<10E-8	<10E-8	P
COL23A1	0,34	0,13	0,2	0,51	0,12	0,39	0,19	0.96	1	0.89	<10E-8	<10E-8	<10E-8	P, ER, E
COL4A1	0,45	0,19	0,26	0,59	0,19	0,4	0,14	0.98	0.95	0.94	<10E-8	<10E-8	<10E-8	E, ER
CPLX1	0,47	0,28	0,19	0,59	0,28	0,32	0,13	0.98	1	0.96	<10E-8	<10E-8	<10E-8	C
CREB3L1	0,5	0,33	0,18	0,61	0,18	0,43	0,25	0.92	0.91	0.79	<10E-8	<10E-8	<10E-8	N, ER
CREB5	0,35	0,13	0,22	0,48	0,15	0,33	0,11	0.96	1	0.91	<10E-8	<10E-8	<10E-8	E, N
CRHR2	0,37	0,15	0,23	0,49	0,13	0,36	0,13	0.96	1	0.91	<10E-8	<10E-8	<10E-8	P
CRISPLD1	0,38	0,2	0,17	0,51	0,24	0,27	0,1	0.96	0.95	0.91	<10E-8	<10E-8	<10E-8	E
DKK 2	0,45	0,15	0,3	0,54	0,19	0,35	0,05	0.97	1	0.91	<10E-8	<10E-8	<10E-8	E
EDNRB	0,53	0,27	0,26	0,59	0,36	0,23	-0,03	0.9	0.89	0.85	<10E-8	<10E-8	<10E-8	P
EN2	0,49	0,26	0,23	0,59	0,26	0,33	0,1	0.98	1	0.96	<10E-8	<10E-8	<10E-8	N

FBN1	0,35	0,11	0,24	0,49	0,08	0,41	0,16	0.95	1	0.91	<10E-8	<10E-8	<10E-8	E, P
FBN2	0,45	0,18	0,27	0,59	0,18	0,41	0,14	0.98	1	0.94	<10E-8	<10E-8	<10E-8	E
FNDC1	0,43	0,16	0,27	0,52	0,13	0,39	0,12	0.96	1	0.9	<10E-8	<10E-8	<10E-8	N, E
GAD2	0,47	0,21	0,26	0,55	0,24	0,31	0,05	0.99	1	0.95	<10E-8	<10E-8	<10E-8	P, C, GA
GHSR	0,43	0,26	0,17	0,56	0,3	0,26	0,09	0.97	0.98	0.88	<10E-8	<10E-8	<10E-8	P
GLRB	0,42	0,2	0,22	0,56	0,16	0,4	0,18	0.97	1	0.96	<10E-8	<10E-8	<10E-8	P
GNG7	0,65	0,85	-0,2	0,6	0,89	-0,29	-0,09	0.94	0.95	0.86	<10E-8	<10E-8	<10E-8	E, P
GRIA4	0,53	0,17	0,36	0,63	0,18	0,45	0,08	0.97	0.98	0.92	<10E-8	<10E-8	<10E-8	P, E
GRIK3	0,49	0,27	0,21	0,58	0,29	0,29	0,08	0.97	0.98	0.94	<10E-8	<10E-8	<10E-8	P
GRIN2D	0,42	0,17	0,25	0,58	0,17	0,42	0,16	0.98	0.98	0.94	3E-07	<10E-8	<10E-8	P
GRIN2D_1	0,63	0,39	0,23	0,73	0,46	0,26	0,03	0.93	0.95	0.81	1E-06	<10E-8	<10E-8	P
GRIN2D_2	0,53	0,34	0,19	0,63	0,33	0,31	0,11	0.97	1	0.9	<10E-8	<10E-8	<10E-8	P
GRIN3B	0,64	0,83	-0,19	0,63	0,87	-0,24	-0,05	0.98	1	0.97	<10E-8	<10E-8	<10E-8	P
GRM6	0,62	0,46	0,16	0,7	0,5	0,2	0,04	0.93	1	0.7	<10E-8	<10E-8	<10E-8	P, ER, GA
GRM7	0,62	0,38	0,24	0,7	0,47	0,24	0	0.96	1	0.79	<10E-8	<10E-8	<10E-8	P
HAPLN4	0,52	0,32	0,2	0,63	0,32	0,31	0,11	0.99	1	0.96	<10E-8	<10E-8	<10E-8	E
HTR1E	0,46	0,22	0,25	0,59	0,27	0,32	0,08	0.96	1	0.88	<10E-8	<10E-8	<10E-8	P
ITGA4	0,48	0,16	0,32	0,61	0,19	0,42	0,1	0.93	0.98	0.75	<10E-8	<10E-8	<10E-8	P, E
ITGA8	0,4	0,16	0,24	0,53	0,17	0,36	0,12	0.98	1	0.93	<10E-8	<10E-8	<10E-8	P, ER
JAM3	0,41	0,18	0,23	0,56	0,19	0,37	0,14	0.97	1	0.9	<10E-8	<10E-8	<10E-8	P, E
KCNA1	0,46	0,25	0,21	0,55	0,23	0,32	0,11	0.96	0.98	0.9	<10E-8	<10E-8	<10E-8	P, ER, C
KCNA3	0,55	0,35	0,2	0,67	0,39	0,29	0,08	0.98	0.98	0.91	<10E-8	<10E-8	<10E-8	P

KCNQ5	0,44	0,15	0,29	0,57	0,14	0,44	0,15	0.98	1	0.94	<10E-8	<10E-8	<10E-8	P
LAMA1	0,49	0,23	0,26	0,58	0,24	0,34	0,08	0.97	0.98	0.93	<10E-8	<10E-8	<10E-8	E
LHCGR	0,56	0,37	0,19	0,66	0,39	0,28	0,09	0.89	0.98	0.77	<10E-8	<10E-8	<10E-8	P, EN
MMP9	0,59	0,43	0,16	0,69	0,46	0,23	0,07	0.97	1	0.91	<10E-8	<10E-8	<10E-8	E, P, N
NELL1	0,49	0,24	0,25	0,6	0,24	0,36	0,11	0.99	0.98	0.97	<10E-8	<10E-8	<10E-8	E, N
NOS1	0,4	0,19	0,21	0,5	0,19	0,3	0,09	0.97	1	0.93	<10E-8	<10E-8	<10E-8	P, N, ER, C, CS, M
NPBWR1	0,45	0,22	0,23	0,57	0,25	0,33	0,09	0.99	1	0.97	<10E-8	<10E-8	<10E-8	P
NPY2R	0,38	0,16	0,22	0,47	0,16	0,32	0,09	0.97	0.95	0.91	<10E-8	<10E-8	<10E-8	P
OPRK1	0,52	0,25	0,27	0,59	0,27	0,32	0,05	0.96	0.95	0.92	<10E-8	<10E-8	<10E-8	P
PCDHG@	0,64	0,38	0,26	0,74	0,44	0,31	0,05	0.98	0.98	0.96	<10E-8	<10E-8	<10E-8	P
PCDHG@ 1	0,5	0,23	0,28	0,65	0,22	0,43	0,16	1	1	0.96	<10E-8	<10E-8	<10E-8	P
PLCB1	0,44	0,18	0,26	0,57	0,16	0,41	0,14	0.99	0.98	0.96	<10E-8	<10E-8	<10E-8	E, N, C
PRKAR1B	0,51	0,17	0,34	0,67	0,27	0,4	0,06	0.99	1	0.96	<10E-8	<10E-8	<10E-8	C
PRKCB	0,25	0,06	0,19	0,38	0,05	0,32	0,13	0.96	1	0.89	<10E-8	<10E-8	<10E-8	P, E, N, C
PTGFR	0,5	0,34	0,16	0,58	0,39	0,2	0,04	0.98	1	0.96	<10E-8	<10E-8	<10E-8	P, E
RSPO2	0,44	0,18	0,26	0,56	0,18	0,38	0,12	0.99	0.98	0.95	<10E-8	<10E-8	<10E-8	E
RSPO3	0,48	0,25	0,23	0,59	0,28	0,31	0,08	0.95	0.98	0.89	4E-06	<10E-8	<10E-8	E
SCTR	0,5	0,26	0,24	0,64	0,24	0,4	0,17	0.91	0.95	0.78	<10E-8	<10E-8	<10E-8	P, CS
SDC2	0,33	0,08	0,25	0,43	0,08	0,35	0,1	0.96	0.98	0.94	<10E-8	<10E-8	<10E-8	P, E, L, GA
SFRP1	0,52	0,28	0,24	0,62	0,34	0,29	0,04	0.98	1	0.94	<10E-8	<10E-8	<10E-8	P, E, C
SFRP2	0,41	0,19	0,22	0,54	0,16	0,38	0,16	0.96	1	0.93	<10E-8	<10E-8	<10E-8	E
SHANK1	0,43	0,23	0,2	0,52	0,26	0,27	0,06	0.93	0.95	0.82	<10E-8	<10E-8	<10E-8	P, C, N

SLC12A5	0,58	0,41	0,17	0,69	0,44	0,25	0,08	0.94	0.95	0.86	<10E-8	<10E-8	<10E-8	P
SLC18A3	0,5	0,28	0,22	0,61	0,3	0,31	0,09	0.94	0.91	0.87	<10E-8	<10E-8	<10E-8	P, GA
SLC32A1	0,48	0,28	0,19	0,57	0,3	0,27	0,08	0.99	0.98	0.95	<10E-8	<10E-8	<10E-8	P
SLC5A7	0,44	0,25	0,18	0,56	0,25	0,31	0,12	0.98	1	0.92	<10E-8	<10E-8	<10E-8	P
SLC6A1	0,49	0,31	0,18	0,57	0,31	0,25	0,07	0.97	1	0.92	<10E-8	<10E-8	<10E-8	P
SLC6A11	0,55	0,39	0,16	0,65	0,42	0,22	0,06	0.97	0.98	0.88	<10E-8	<10E-8	<10E-8	P
SLC6A3	0,43	0,23	0,2	0,58	0,25	0,33	0,13	0.98	1	0.92	<10E-8	<10E-8	<10E-8	P
SLC8A1	0,56	0,25	0,3	0,69	0,31	0,38	0,08	0.99	0.98	0.96	3E-06	<10E-8	<10E-8	P, N
SLIT3	0,49	0,19	0,3	0,6	0,19	0,4	0,1	0.96	1	0.92	<10E-8	<10E-8	<10E-8	E, M
SPOCK1	0,39	0,17	0,23	0,53	0,17	0,37	0,14	0.98	1	0.95	<10E-8	<10E-8	<10E-8	E
SYN3	0,4	0,19	0,22	0,51	0,17	0,34	0,13	0.92	1	0.8	<10E-8	<10E-8	<10E-8	C
VCAN	0,36	0,11	0,25	0,49	0,1	0,39	0,15	0.93	1	0.86	<10E-8	<10E-8	<10E-8	E, L, GA
VIPR2	0,48	0,23	0,26	0,64	0,23	0,41	0,15	0.99	1	0.97	<10E-8	<10E-8	<10E-8	P

Extracellular = E

Endoplasmic reticulum = ER

Plasma membrane = P

Nucleus = N

Cytosol = C

Golgi apparatus = GA

Endosome = EN

Cytoskeleton = CS

Mitochondrion = M

Lysosome = L

Supporting Information Table 8: statistical performance measures relative to 74 identified CpG islands biomarkers

ADAMTS2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 178770724-178772794	Estimate	0.961	0.909	0.977	0.819	0.997	0.581	40.012	0.093	1	31	0.887
	95%_CI_lower_limit	0.943	0.874	0.882	-	0.982	0.467	5.762	0.066	-	-	-
	95%_CI_upper_limit	0.979	0.935	0.996	-	0.999	0.687	277.834	0.130	-	-	-
ADAMTS5		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr21: 28337856-28340237	Estimate	0.989	0.953	0.977	0.866	0.997	0.729	41.942	0.048	1	16	0.930
	95%_CI_lower_limit	0.98	0.925	0.882	-	0.983	0.604	6.041	0.030	-	-	-
	95%_CI_upper_limit	0.998	0.971	0.996	-	0.999	0.826	291.192	0.077	-	-	-
ADCY8		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 132052203-132054749	Estimate	0.976	0.889	0.977	0.916	0.997	0.531	39.111	0.114	1	38	0.866
	95%_CI_lower_limit	0.962	0.851	0.882	-	0.982	0.423	5.632	0.084	-	-	-
	95%_CI_upper_limit	0.99	0.918	0.996	-	0.999	0.636	271.601	0.154	-	-	-
ADRA1B		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 159399004-159399928	Estimate	0.968	0.939	0.955	0.699	0.994	0.667	20.649	0.064	2	21	0.893
	95%_CI_lower_limit	0.951	0.908	0.849	-	0.978	0.544	5.330	0.042	-	-	-
	95%_CI_upper_limit	0.985	0.959	0.987	-	0.998	0.771	79.997	0.098	-	-	-
CACNA1A		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 13616752-13617267	Estimate	0.952	0.906	1.000	0.878	1.000	0.579	Inf	0.094	0	32	0.906
	95%_CI_lower_limit	0.93	0.871	0.920	-	0.988	0.467	NaN	0.067	-	-	-
	95%_CI_upper_limit	0.973	0.933	1.000	-	1.000	0.684	Inf	0.130	-	-	-
CACNA1E		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr1: 181452706-181453073	Estimate	0.934	0.810	0.955	0.909	0.993	0.393	17.819	0.199	2	65	0.764
	95%_CI_lower_limit	0.908	0.765	0.849	-	0.974	0.305	4.596	0.159	-	-	-
	95%_CI_upper_limit	0.96	0.848	0.987	-	0.998	0.487	69.080	0.250	-	-	-
CDH4		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr20: 59826977-59828978	Estimate	0.99	0.959	0.977	0.764	0.997	0.754	42.199	0.042	1	14	0.936
	95%_CI_lower_limit	0.981	0.932	0.882	-	0.983	0.629	6.078	0.025	-	-	-
	95%_CI_upper_limit	0.998	0.975	0.996	-	0.999	0.848	292.973	0.070	-	-	-
COL23A1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 178016558-178017670	Estimate	0.961	0.889	1.000	0.891	1.000	0.537	Inf	0.111	0	38	0.889
	95%_CI_lower_limit	0.943	0.851	0.920	-	0.988	0.429	NaN	0.082	-	-	-
	95%_CI_upper_limit	0.98	0.918	1.000	-	1.000	0.640	Inf	0.150	-	-	-
COL4A1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr13: 110958891-110960590	Estimate	0.989	0.968	0.977	0.682	0.997	0.796	42.585	0.033	1	11	0.945
	95%_CI_lower_limit	0.98	0.943	0.882	-	0.983	0.671	6.134	0.018	-	-	-

	95%_CI_upper_limit	0.997	0.982	0.996	-	0.999	0.882	295.644	0.059	-	-	-
CPLX1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr4: 778661-780592	Estimate	0.984	0.965	1.000	0.724	1.000	0.786	Inf	0.035	0	12	0.965
	95%_CI_lower_limit	0.972	0.940	0.920	-	0.988	0.662	NaN	0.020	-	-	-
	95%_CI_upper_limit	0.996	0.980	1.000	-	1.000	0.873	Inf	0.061	-	-	-
CREB3L1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr11: 46316875-46317485	Estimate	0.94	0.868	0.909	0.922	0.987	0.471	9.553	0.145	4	45	0.778
	95%_CI_lower_limit	0.915	0.828	0.788	-	0.966	0.368	3.749	0.109	-	-	-
	95%_CI_upper_limit	0.966	0.900	0.964	-	0.995	0.576	24.340	0.193	-	-	-
CREB5		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr7: 28448716-28450028	Estimate	0.962	0.906	1.000	0.880	1.000	0.579	Inf	0.094	0	32	0.906
	95%_CI_lower_limit	0.943	0.871	0.920	-	0.988	0.467	NaN	0.067	-	-	-
	95%_CI_upper_limit	0.98	0.933	1.000	-	1.000	0.684	Inf	0.130	-	-	-
CRHR2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr7: 30721372-30722445	Estimate	0.956	0.909	1.000	0.881	1.000	0.587	Inf	0.091	0	31	0.909
	95%_CI_lower_limit	0.937	0.874	0.920	-	0.988	0.474	NaN	0.065	-	-	-
	95%_CI_upper_limit	0.976	0.935	1.000	-	1.000	0.691	Inf	0.127	-	-	-
CRISPLD1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 75896528-75897116	Estimate	0.961	0.912	0.955	0.849	0.994	0.583	20.070	0.092	2	30	0.867
	95%_CI_lower_limit	0.942	0.878	0.849	-	0.977	0.468	5.180	0.065	-	-	-
	95%_CI_upper_limit	0.98	0.938	0.987	-	0.998	0.690	77.764	0.130	-	-	-
DKK2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr4: 107956555-107957453	Estimate	0.973	0.906	1.000	0.892	1.000	0.579	Inf	0.094	0	32	0.906
	95%_CI_lower_limit	0.958	0.871	0.920	-	0.988	0.467	NaN	0.067	-	-	-
	95%_CI_upper_limit	0.988	0.933	1.000	-	1.000	0.684	Inf	0.130	-	-	-
EDNRB		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr13: 78492425-78493382	Estimate	0.916	0.787	0.909	0.900	0.985	0.354	8.652	0.235	4	73	0.696
	95%_CI_lower_limit	0.883	0.740	0.788	-	0.963	0.272	3.393	0.188	-	-	-
	95%_CI_upper_limit	0.949	0.827	0.964	-	0.994	0.446	22.061	0.294	-	-	-
EN2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr7: 155255098-155255311	Estimate	0.984	0.959	1.000	0.737	1.000	0.759	Inf	0.041	0	14	0.959
	95%_CI_lower_limit	0.973	0.932	0.920	-	0.988	0.635	NaN	0.025	-	-	-
	95%_CI_upper_limit	0.996	0.975	1.000	-	1.000	0.850	Inf	0.068	-	-	-
FBN1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr15: 48936810-48938577	Estimate	0.967	0.930	1.000	0.780	1.000	0.647	Inf	0.070	0	24	0.930
	95%_CI_lower_limit	0.95	0.898	0.920	-	0.988	0.528	NaN	0.048	-	-	-
	95%_CI_upper_limit	0.984	0.952	1.000	-	1.000	0.750	Inf	0.103	-	-	-

FBN2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 127872563-127874945	Estimate	0.979	0.942	1.000	0.768	1.000	0.688	Inf	0.058	0	20	0.942
	95%_CI_lower_limit	0.965	0.911	0.920	-	0.988	0.566	NaN	0.038	-	-	-
	95%_CI_upper_limit	0.992	0.962	1.000	-	1.000	0.788	Inf	0.089	-	-	-
FNDC1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr6: 159589636-159591319	Estimate	0.957	0.898	1.000	0.886	1.000	0.557	Inf	0.102	0	35	0.898
	95%_CI_lower_limit	0.937	0.861	0.920	-	0.988	0.447	NaN	0.075	-	-	-
	95%_CI_upper_limit	0.977	0.925	1.000	-	1.000	0.661	Inf	0.140	-	-	-
GAD2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr10: 26504383-26507434	Estimate	0.961	0.863	1.000	0.929	1.000	0.484	Inf	0.137	0	47	0.863
	95%_CI_lower_limit	0.943	0.822	0.920	-	0.987	0.384	NaN	0.105	-	-	-
	95%_CI_upper_limit	0.979	0.895	1.000	-	1.000	0.585	Inf	0.179	-	-	-
GHSR		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr3: 172165372-172166738	Estimate	0.965	0.880	0.977	0.901	0.997	0.512	38.725	0.123	1	41	0.857
	95%_CI_lower_limit	0.948	0.841	0.882	-	0.981	0.407	5.576	0.092	-	-	-
	95%_CI_upper_limit	0.982	0.910	0.996	-	0.999	0.616	268.929	0.164	-	-	-
GLRB		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr4: 157997166-157997686	Estimate	0.972	0.956	1.000	0.724	1.000	0.746	Inf	0.044	0	15	0.956
	95%_CI_lower_limit	0.955	0.929	0.920	-	0.988	0.622	NaN	0.027	-	-	-
	95%_CI_upper_limit	0.988	0.973	1.000	-	1.000	0.839	Inf	0.072	-	-	-
GNG7		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 2513215-2513696	Estimate	0.971	0.930	0.977	0.810	0.997	0.642	40.912	0.072	1	24	0.907
	95%_CI_lower_limit	0.956	0.898	0.882	-	0.982	0.522	5.892	0.049	-	-	-
	95%_CI_upper_limit	0.986	0.952	0.996	-	0.999	0.746	284.068	0.106	-	-	-
GRIA4		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr11: 105481126-105481422	Estimate	0.984	0.962	1.000	0.762	1.000	0.772	Inf	0.038	0	13	0.962
	95%_CI_lower_limit	0.973	0.936	0.920	-	0.988	0.648	NaN	0.022	-	-	-
	95%_CI_upper_limit	0.995	0.978	1.000	-	1.000	0.862	Inf	0.065	-	-	-
GRIK3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr1: 37498377-37500624	Estimate	0.968	0.901	1.000	0.903	1.000	0.564	Inf	0.099	0	34	0.901
	95%_CI_lower_limit	0.951	0.864	0.920	-	0.988	0.454	NaN	0.072	-	-	-
	95%_CI_upper_limit	0.984	0.928	1.000	-	1.000	0.669	Inf	0.137	-	-	-
GRIN2D		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 48918115-48918340	Estimate	0.983	0.974	1.000	0.730	1.000	0.830	Inf	0.026	0	9	0.974
	95%_CI_lower_limit	0.97	0.951	0.920	-	0.989	0.708	NaN	0.014	-	-	-
	95%_CI_upper_limit	0.996	0.986	1.000	-	1.000	0.908	Inf	0.050	-	-	-

GRIN2D_1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 48919173-48919412	Estimate	0.995	0.974	0.955	0.785	0.994	0.824	21.421	0.028	2	9	0.928
	95%_CI_lower_limit	0.991	0.951	0.849	-	0.978	0.697	5.530	0.014	-	-	-
	95%_CI_upper_limit	1	0.986	0.987	-	0.998	0.904	82.974	0.053	-	-	-
GRIN2D_2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 48945800-48947636	Estimate	0.987	0.965	1.000	0.836	1.000	0.786	Inf	0.035	0	12	0.965
	95%_CI_lower_limit	0.976	0.940	0.920	-	0.988	0.662	NaN	0.020	-	-	-
	95%_CI_upper_limit	0.997	0.980	1.000	-	1.000	0.873	Inf	0.061	-	-	-
GRIN3B		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 1004324-1005605	Estimate	0.899	0.848	0.886	0.841	0.983	0.429	7.462	0.172	5	52	0.734
	95%_CI_lower_limit	0.867	0.806	0.760	-	0.961	0.332	3.265	0.131	-	-	-
	95%_CI_upper_limit	0.932	0.882	0.950	-	0.993	0.531	17.052	0.225	-	-	-
GRM6		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 178421225-178422337	Estimate	0.928	0.699	1.000	0.956	1.000	0.299	Inf	0.301	0	103	0.699
	95%_CI_lower_limit	0.897	0.648	0.920	-	0.984	0.231	NaN	0.256	-	-	-
	95%_CI_upper_limit	0.958	0.745	1.000	-	1.000	0.378	Inf	0.354	-	-	-
GRM7		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr3: 6902823-6903516	Estimate	0.956	0.789	1.000	0.941	1.000	0.379	Inf	0.211	0	72	0.789
	95%_CI_lower_limit	0.934	0.743	0.920	-	0.986	0.296	NaN	0.171	-	-	-
	95%_CI_upper_limit	0.977	0.829	1.000	-	1.000	0.470	Inf	0.258	-	-	-
HAPLN4		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 19371675-19372393	Estimate	0.962	0.895	1.000	0.872	1.000	0.550	Inf	0.105	0	36	0.895
	95%_CI_lower_limit	0.944	0.858	0.920	-	0.988	0.441	NaN	0.077	-	-	-
	95%_CI_upper_limit	0.98	0.923	1.000	-	1.000	0.654	Inf	0.143	-	-	-
HTR1E		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr6: 87647253-87647707	Estimate	0.955	0.883	1.000	0.900	1.000	0.524	Inf	0.117	0	40	0.883
	95%_CI_lower_limit	0.936	0.845	0.920	-	0.987	0.418	NaN	0.087	-	-	-
	95%_CI_upper_limit	0.975	0.913	1.000	-	1.000	0.627	Inf	0.156	-	-	-
ITGA4		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr2: 182321761-182323029	Estimate	0.991	0.962	0.977	0.833	0.997	0.768	42.327	0.039	1	13	0.939
	95%_CI_lower_limit	0.982	0.936	0.882	-	0.983	0.642	6.097	0.023	-	-	-
	95%_CI_upper_limit	0.999	0.978	0.996	-	0.999	0.859	293.863	0.066	-	-	-
ITGA8		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr10: 15761423-15762101	Estimate	0.975	0.912	0.977	0.869	0.997	0.589	40.140	0.090	1	30	0.890
	95%_CI_lower_limit	0.961	0.878	0.882	-	0.982	0.474	5.781	0.064	-	-	-
	95%_CI_upper_limit	0.989	0.938	0.996	-	0.999	0.695	278.725	0.127	-	-	-

JAM3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr11: 133938850-133939681	Estimate	0.975	0.927	1.000	0.818	1.000	0.638	Inf	0.073	0	25	0.927
	95%_CI_lower_limit	0.961	0.894	0.920	-	0.988	0.520	NaN	0.050	-	-	-
	95%_CI_upper_limit	0.989	0.950	1.000	-	1.000	0.741	Inf	0.107	-	-	-
KCNA1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr12: 5018585-5021171	Estimate	0.974	0.904	1.000	0.922	1.000	0.571	Inf	0.096	0	33	0.904
	95%_CI_lower_limit	0.96	0.868	0.920	-	0.988	0.460	NaN	0.070	-	-	-
	95%_CI_upper_limit	0.988	0.930	1.000	-	1.000	0.676	Inf	0.133	-	-	-
KCNA3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr1: 111216244-111217937	Estimate	0.979	0.939	0.977	0.875	0.997	0.672	41.298	0.063	1	21	0.916
	95%_CI_lower_limit	0.966	0.908	0.882	-	0.983	0.550	5.948	0.041	-	-	-
	95%_CI_upper_limit	0.992	0.959	0.996	-	0.999	0.774	286.739	0.095	-	-	-
KCNQ5		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr6: 73330942-73333109	Estimate	0.983	0.942	1.000	0.852	1.000	0.688	Inf	0.058	0	20	0.942
	95%_CI_lower_limit	0.972	0.911	0.920	-	0.988	0.566	NaN	0.038	-	-	-
	95%_CI_upper_limit	0.995	0.962	1.000	-	1.000	0.788	Inf	0.089	-	-	-
LAMA1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr18: 7116852-7118241	Estimate	0.977	0.939	0.955	0.852	0.994	0.667	20.649	0.064	2	21	0.893
	95%_CI_lower_limit	0.963	0.908	0.849	-	0.978	0.544	5.330	0.042	-	-	-
	95%_CI_upper_limit	0.991	0.959	0.987	-	0.998	0.771	79.997	0.098	-	-	-
LHCGR		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr2: 48982621-48982958	Estimate	0.923	0.798	1.000	0.911	1.000	0.389	Inf	0.202	0	69	0.798
	95%_CI_lower_limit	0.896	0.753	0.920	-	0.986	0.305	NaN	0.163	-	-	-
	95%_CI_upper_limit	0.951	0.837	1.000	-	1.000	0.482	Inf	0.249	-	-	-
MMP9		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr20: 44640288-44641210	Estimate	0.909	0.781	0.955	0.898	0.993	0.359	17.175	0.230	2	75	0.735
	95%_CI_lower_limit	0.878	0.734	0.849	-	0.973	0.278	4.429	0.186	-	-	-
	95%_CI_upper_limit	0.94	0.821	0.987	-	0.998	0.449	66.599	0.283	-	-	-
NELL1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr11: 20690579-20691845	Estimate	0.985	0.953	1.000	0.872	1.000	0.733	Inf	0.047	0	16	0.953
	95%_CI_lower_limit	0.974	0.925	0.920	-	0.988	0.610	NaN	0.029	-	-	-
	95%_CI_upper_limit	0.995	0.971	1.000	-	1.000	0.829	Inf	0.075	-	-	-
NOS1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr12: 117798076-117799448	Estimate	0.967	0.924	0.977	0.837	0.997	0.623	40.655	0.078	1	26	0.901
	95%_CI_lower_limit	0.951	0.891	0.882	-	0.982	0.505	5.855	0.054	-	-	-
	95%_CI_upper_limit	0.984	0.948	0.996	-	0.999	0.728	282.287	0.113	-	-	-

NPBWR1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 53851701-53854426	Estimate	0.988	0.971	1.000	0.787	1.000	0.815	Inf	0.029	0	10	0.971
	95%_CI_lower_limit	0.978	0.947	0.920	-	0.989	0.692	NaN	0.016	-	-	-
	95%_CI_upper_limit	0.998	0.984	1.000	-	1.000	0.896	Inf	0.054	-	-	-
NPY2R		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr4: 156129168-156130209	Estimate	0.97	0.909	0.955	0.851	0.994	0.575	20.006	0.095	2	31	0.864
	95%_CI_lower_limit	0.953	0.874	0.849	-	0.977	0.461	5.163	0.067	-	-	-
	95%_CI_upper_limit	0.986	0.935	0.987	-	0.998	0.682	77.515	0.134	-	-	-
OPRK1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 54163303-54164443	Estimate	0.963	0.918	0.955	0.848	0.994	0.600	20.199	0.086	2	28	0.873
	95%_CI_lower_limit	0.946	0.884	0.849	-	0.977	0.483	5.213	0.060	-	-	-
	95%_CI_upper_limit	0.981	0.943	0.987	-	0.998	0.707	78.260	0.123	-	-	-
PCDHG.		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 140857864-140858065	Estimate	0.978	0.965	0.977	0.798	0.997	0.782	42.456	0.036	1	12	0.942
	95%_CI_lower_limit	0.964	0.940	0.882	-	0.983	0.656	6.115	0.021	-	-	-
	95%_CI_upper_limit	0.992	0.980	0.996	-	0.999	0.871	294.754	0.063	-	-	-
PCDHG._1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 140864527-140864748	Estimate	0.996	0.962	1.000	0.922	1.000	0.772	Inf	0.038	0	13	0.962
	95%_CI_lower_limit	0.99	0.936	0.920	-	0.988	0.648	NaN	0.022	-	-	-
	95%_CI_upper_limit	1.001	0.978	1.000	-	1.000	0.862	Inf	0.065	-	-	-
PLCB1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr20: 8112884-8113592	Estimate	0.892	0.769	0.977	0.846	0.996	0.352	33.836	0.236	1	79	0.746
	95%_CI_lower_limit	0.86	0.721	0.882	-	0.979	0.273	4.870	0.194	-	-	-
	95%_CI_upper_limit	0.924	0.811	0.996	-	0.999	0.441	235.090	0.288	-	-	-
PRKAR1B		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr7: 751712-752150	Estimate	0.986	0.956	1.000	0.853	1.000	0.746	Inf	0.044	0	15	0.956
	95%_CI_lower_limit	0.974	0.929	0.920	-	0.988	0.622	NaN	0.027	-	-	-
	95%_CI_upper_limit	0.997	0.973	1.000	-	1.000	0.839	Inf	0.072	-	-	-
PRKCB		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr16: 23846941-23848102	Estimate	0.961	0.904	0.977	0.810	0.997	0.566	39.754	0.099	1	33	0.881
	95%_CI_lower_limit	0.943	0.868	0.882	-	0.982	0.454	5.725	0.071	-	-	-
	95%_CI_upper_limit	0.979	0.930	0.996	-	0.999	0.671	276.053	0.137	-	-	-
PTGFR		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr1: 78956624-78957516	Estimate	0.932	0.819	0.955	0.906	0.993	0.404	18.012	0.190	2	62	0.773
	95%_CI_lower_limit	0.906	0.774	0.849	-	0.975	0.315	4.646	0.150	-	-	-
	95%_CI_upper_limit	0.959	0.856	0.987	-	0.998	0.500	69.824	0.240	-	-	-

RSPO2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 109094485-109095849	Estimate	0.987	0.947	0.977	0.882	0.997	0.705	41.684	0.054	1	18	0.925
	95%_CI_lower_limit	0.977	0.918	0.882	-	0.983	0.581	6.004	0.034	-	-	-
	95%_CI_upper_limit	0.996	0.966	0.996	-	0.999	0.804	289.411	0.085	-	-	-
RSPO3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr6: 127441553-127441760	Estimate	0.947	0.895	0.977	0.871	0.997	0.544	39.368	0.108	1	36	0.872
	95%_CI_lower_limit	0.925	0.858	0.882	-	0.982	0.435	5.669	0.079	-	-	-
	95%_CI_upper_limit	0.969	0.923	0.996	-	0.999	0.650	273.382	0.147	-	-	-
SCTR		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr2: 120281661-120282211	Estimate	0.964	0.930	1.000	0.771	1.000	0.647	Inf	0.070	0	24	0.930
	95%_CI_lower_limit	0.946	0.898	0.920	-	0.988	0.528	NaN	0.048	-	-	-
	95%_CI_upper_limit	0.982	0.952	1.000	-	1.000	0.750	Inf	0.103	-	-	-
SDC2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 97505747-97507607	Estimate	0.962	0.936	0.977	0.615	0.997	0.662	41.170	0.066	1	22	0.913
	95%_CI_lower_limit	0.944	0.905	0.882	-	0.983	0.540	5.929	0.044	-	-	-
	95%_CI_upper_limit	0.981	0.957	0.996	-	0.999	0.765	285.849	0.099	-	-	-
SFRP1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr8: 41165852-41167140	Estimate	0.983	0.944	1.000	0.862	1.000	0.698	Inf	0.056	0	19	0.944
	95%_CI_lower_limit	0.972	0.915	0.920	-	0.988	0.576	NaN	0.036	-	-	-
	95%_CI_upper_limit	0.994	0.964	1.000	-	1.000	0.798	Inf	0.086	-	-	-
SFRP2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr4: 154709512-154710827	Estimate	0.964	0.933	1.000	0.754	1.000	0.657	Inf	0.067	0	23	0.933
	95%_CI_lower_limit	0.945	0.901	0.920	-	0.988	0.537	NaN	0.045	-	-	-
	95%_CI_upper_limit	0.982	0.955	1.000	-	1.000	0.759	Inf	0.100	-	-	-
SHANK1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr19: 51169659-51172023	Estimate	0.939	0.860	0.955	0.873	0.993	0.467	18.912	0.147	2	48	0.814
	95%_CI_lower_limit	0.915	0.819	0.849	-	0.976	0.367	4.880	0.112	-	-	-
	95%_CI_upper_limit	0.962	0.892	0.987	-	0.998	0.569	73.298	0.193	-	-	-
SLC12A5		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr20: 44685771-44687610	Estimate	0.93	0.751	0.977	0.936	0.996	0.336	33.064	0.254	1	85	0.729
	95%_CI_lower_limit	0.901	0.703	0.882	-	0.978	0.260	4.759	0.210	-	-	-
	95%_CI_upper_limit	0.959	0.794	0.996	-	0.999	0.421	229.747	0.307	-	-	-
SLC18A3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr10: 50817601-50820356	Estimate	0.97	0.939	0.977	0.815	0.997	0.672	41.298	0.063	1	21	0.916
	95%_CI_lower_limit	0.955	0.908	0.882	-	0.983	0.550	5.948	0.041	-	-	-
	95%_CI_upper_limit	0.986	0.959	0.996	-	0.999	0.774	286.739	0.095	-	-	-

SLC32A1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr20: 37352130-37357372	Estimate	0.978	0.915	1.000	0.893	1.000	0.603	Inf	0.085	0	29	0.915
	95%_CI_lower_limit	0.965	0.881	0.920	-	0.988	0.488	NaN	0.060	-	-	-
	95%_CI_upper_limit	0.991	0.940	1.000	-	1.000	0.707	Inf	0.120	-	-	-
SLC5A7		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr2: 108602824-108603467	Estimate	0.944	0.863	0.955	0.870	0.993	0.472	18.977	0.144	2	47	0.817
	95%_CI_lower_limit	0.921	0.822	0.849	-	0.976	0.372	4.896	0.110	-	-	-
	95%_CI_upper_limit	0.966	0.895	0.987	-	0.998	0.575	73.546	0.189	-	-	-
SLC6A1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr3: 11034446-11035384	Estimate	0.973	0.918	1.000	0.808	1.000	0.611	Inf	0.082	0	28	0.918
	95%_CI_lower_limit	0.959	0.884	0.920	-	0.988	0.496	NaN	0.057	-	-	-
	95%_CI_upper_limit	0.988	0.943	1.000	-	1.000	0.715	Inf	0.117	-	-	-
SLC6A11		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr3: 10857687-10858447	Estimate	0.973	0.883	0.977	0.907	0.997	0.518	38.854	0.120	1	40	0.860
	95%_CI_lower_limit	0.958	0.845	0.882	-	0.982	0.412	5.595	0.089	-	-	-
	95%_CI_upper_limit	0.988	0.913	0.996	-	0.999	0.622	269.820	0.161	-	-	-
SLC6A3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 1444678-1446648	Estimate	0.984	0.921	1.000	0.874	1.000	0.620	Inf	0.079	0	27	0.921
	95%_CI_lower_limit	0.973	0.888	0.920	-	0.988	0.503	NaN	0.055	-	-	-
	95%_CI_upper_limit	0.994	0.945	1.000	-	1.000	0.724	Inf	0.113	-	-	-
SLC8A1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr2: 40678477-40678717	Estimate	0.965	0.918	0.955	0.828	0.994	0.600	20.199	0.086	2	28	0.873
	95%_CI_lower_limit	0.949	0.884	0.849	-	0.977	0.483	5.213	0.060	-	-	-
	95%_CI_upper_limit	0.982	0.943	0.987	-	0.998	0.707	78.260	0.123	-	-	-
SLIT3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 168727429-168728275	Estimate	0.964	0.918	1.000	0.870	1.000	0.611	Inf	0.082	0	28	0.918
	95%_CI_lower_limit	0.946	0.884	0.920	-	0.988	0.496	NaN	0.057	-	-	-
	95%_CI_upper_limit	0.981	0.943	1.000	-	1.000	0.715	Inf	0.117	-	-	-
SPOCK1		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 136834016-136835146	Estimate	0.98	0.953	1.000	0.760	1.000	0.733	Inf	0.047	0	16	0.953
	95%_CI_lower_limit	0.967	0.925	0.920	-	0.988	0.610	NaN	0.029	-	-	-
	95%_CI_upper_limit	0.993	0.971	1.000	-	1.000	0.829	Inf	0.075	-	-	-
SYN3		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr22: 33453892-33454505	Estimate	0.969	0.906	1.000	0.896	1.000	0.579	Inf	0.094	0	32	0.906
	95%_CI_lower_limit	0.952	0.871	0.920	-	0.988	0.467	NaN	0.067	-	-	-
	95%_CI_upper_limit	0.985	0.933	1.000	-	1.000	0.684	Inf	0.130	-	-	-

VCAN		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr5: 82768387-82769268	Estimate	0.927	0.860	1.000	0.854	1.000	0.478	Inf	0.140	0	48	0.860
	95%_CI_lower_limit	0.901	0.819	0.920	-	0.987	0.379	NaN	0.108	-	-	-
	95%_CI_upper_limit	0.952	0.892	1.000	-	1.000	0.579	Inf	0.182	-	-	-

VIPR2		AUC	Se	Sp	cutoff	PPV	NPV	DLR Positive	DLR Negative	FP	FN	Optimal criterion
chr7: 158936507-158938492	Estimate	0.99	0.971	1.000	0.867	1.000	0.815	Inf	0.029	0	10	0.971
	95%_CI_lower_limit	0.981	0.947	0.920	-	0.989	0.692	NaN	0.016	-	-	-
	95%_CI_upper_limit	0.998	0.984	1.000	-	1.000	0.896	Inf	0.054	-	-	-

AUC: Area Under the ROC curve

Se: sensitivity

Sp: specificity

cutoff: individuals with a test value lower than the cutoff are classified as healthy (negative test), whereas patients with a test value greater than (or equal to) the cutoff are classified as diseased (positive test). The estimation of the optimal cutoff has been evalueted using the Youden index method that gives equal weight to sensitivity and specificity.

PPV: positive predictive value

NPV: negative predictive value

DLR.Positive: positive diagnostic likelihood ratio

DLR.Negative: negative diagnostic likelihood ratio

FP: false positive

FN: false negative

Optimal_criterion: the numerical value of the criterion (Youden index) at the optimal cutpoint

Supporting Information Table 9: Clinical features of the forty five CRC cases enrolled for cell free DNA samples analysed by methyl-BEAMing[1]

Patient ID[2]	Gender	Age	Stage at diagnosis	Patient category	CEA	Genetic alterations	% of DNA KRAS mutated in plasma	Number of metastases	Resection of primary tumor
CP26	M	58	II	Metastatic patient	1.1	BRAF mutated	0	5+	
CP15	M	71	III	Metastatic patient	2.2	Wilde Type	X	5+	no
CP28	M	68	IV	Metastatic patient	2.9	Wilde Type	X	5+	yes
CP27	M	80	II	Metastatic patient	1.7	KRAS mutated	0	5+	no
CP47	F	72	IV	Metastatic patient	2.9	Wilde Type	X	5+	yes
CP25	M	80	III	Metastatic patient	0.9	KRAS mutated	0		no
CP18	M	57	IV	Metastatic patient	1.9	Wilde Type	X	5	no
CP16	F	59	IV	Metastatic patient	1.8	KRAS mutated	0	1	no
CP52	F	64	III	Metastatic patient	4.7	BRAF mutated	0	5+	no
CP19	M	74	IV	Metastatic patient	1.4	KRAS mutated	0	5+	yes
CP30	M	69	II	Metastatic patient	1.9	Wilde Type	X	4	no
CP24	M	69	IV	Metastatic patient	4.1	KRAS mutated	0	5+	no
CP06	M	63	III	Metastatic patient	3.7	KRAS mutated	0	2	no
CP01	M	71	IV	Metastatic patient	4.4	Wilde Type	X	5+	no
CP51	F	62	III	Metastatic patient	1.5	KRAS mutated	0		no
CP02	M	49	IV	Metastatic patient	64.7	KRAS mutated	17	5+	no
CP04	M	70	IV	Metastatic patient	50.1	KRAS mutated	12.4	5+	no
CP05	M	62	IV	Metastatic patient	47.5	Wilde Type	X	4	no
CP11	M	74	IV	Metastatic patient	36.2	KRAS mutated	0.11	11+	yes
CP12	M	63	IV	Metastatic patient	73.7	KRAS mutated	13.2	5 +	yes
CP13	F	67	IV	Metastatic patient	8.3	KRAS mutated	0	5+	no
CP14	F	71	IV	Metastatic patient	896.3	KRAS mutated	0	5+	no
CP31	F	71	IV	Metastatic patient	46.6	Wilde Type	X	5+	no
CP33	F	56	IV	Metastatic patient	9.7	Wilde Type	X	5+	yes
CP35	F	69	IV	Metastatic patient	18.6	Wilde Type	X	5+	yes
CP36	M	69	IV	Metastatic patient	262.5	KRAS mutated	33	5+	yes
CP38	F	83	IV	Metastatic patient	146.5	KRAS mutated	0	5+	no
CP39	F	70	III	Metastatic patient	15.4	Wilde Type	X		no
CP40	M	43	IIIc	Metastatic patient	10.5	KRAS mutated	0	4 (+ lymph nodes)	yes
CP49	F	76	IV	Metastatic patient	2396	Wilde Type	X		no
CP50	M	66	IV	Metastatic patient	46.7	KRAS mutated	2	5+	no
CP53	M	69	IV	Metastatic patient	417.9	KRAS mutated	3	5+	no
CP54	F	67	IV	Metastatic patient	25.1	KRAS mutated	1.9	5+	yes
CP03	F	64	III	NED patient	1.2	Wilde Type	X		no

CP07	F	69	III	NED patient	1.8	Wilde Type	X		no
CP09	M	59	II	NED patient	2.9	Wilde Type	X		no
CP10	M	65	IV	NED patient	20.4	KRAS mutated	0	2	yes
CP17	F	59	III	NED patient	3.5	BRAF mutated	0		no
CP20	M	72	IV	NED patient	2.3	Wilde Type	X		no
CP21	M	56	III	NED patient	3.8	KRAS mutated	0		no
CP23	F	56	II	NED patient	3.1	Wilde Type	X		no
CP29	M	52	IV	NED patient	3.1	KRAS mutated	0	1	no
CP32	M	57	III	NED patient	24.3	Wilde Type	X		yes
CP34	M	46	II	NED patient	NA	NA	X		no
CP55	M	79	III	NED patient	28.2	NA	X		

[1] The primer sequences are reported in **Appendix Table S8.** More protocol details can be found in references 16 and 17.

[2] All DNA samples were treated with sodium bisulfite using the EZ DNA Methylation Gold Kit ™ (Zymo Research).

Supporting Information Table 10: Association between methylation status and cancer of some genes associated to the CpG islands found altered in this study

Gene Name	Associated Disease	Reference
ADAMTS5	CRC	Kim, Young-Ho, et al. Annals of surgical oncology 18.8 (2011): 2338-2347. doi: 10. 1245/ s10434-011-1573-y
ADRA1B	Bipolar disorder	Sugawara, Hiroko, et al. Progress in Neuro-Psychopharmacology and Biological Psychiatry 56 (2015): 117-121. doi: 10.1016/j.pnpbp.2014.08.010
	Cell line (SK-N-MC and DU145)	Michelotti, Gregory A., et al. The FASEB Journal 21.9 (2007): 1979-1993. doi: 10.1096/fj.06-7118com
	Gastric cancer	Noda, H., et al. British journal of cancer 96.2 (2007): 383-390. doi:10.1038/sj.bjc.6603555
CACNA1	Bladder cancer	García-Baquero, Rodrigo, et al. Tumor Biology 35.6 (2014): 5777-5786. doi:10. 1007/ s13277-014-1767-6
	Bladder cancer	García-Baquero, Rodrigo, et al. The Journal of urology 190.2 (2013): 723-730. doi:10.1016/j.juro.2013.01.105
	Ovarian clear cell adenocarcinoma (OCCA)	Ho, Chih-Ming, et al. Mol Cancer 11 (2012): 53. doi: 10.1186/1476-4598-11-53
	Triple-negative breast carcinomas	Branham, M. T., et al. Oncogenesis 1.7 (2012): e17. doi:10.1038/oncsis.2012.17
	Lung cancer	Castro, Mónica, et al. J Transl Med 8.86 (2010): 1479-5876.doi:10. 1186/ 1479-5876-8-86
	Prostate cancer	Kinney, Shannon R. Morey, et al. Molecular Cancer Research 6.8 (2008): 1365-1374. doi:10.1158/1541-7786.MCR-08-0040
CDH4	Gastric cancer	Zhou, Lin, et al. Cancer biology & therapy 16.8 (2015): 1241-1251. doi:10.1080/15384047.2015.1056411
	CRC	Nishioka, Y., et al. The International journal of biological markers 30.1 (2014): e81-7. doi:10.5301/jbm.5000099
	Gastric cancer	Zhou, Lin, et al. The FASEB Journal 27.12 (2013): 4929-4939. doi:10.1096/fj.13-233387
	Nasopharyngeal carcinoma	Du, Chunping, et al. Cancer letters 309.1 (2011): 54-61. doi:10.1016/j.canlet.2011.05.016
	Head and neck squamous cell carcinoma (HNSCC) and Salivary rinses	Demokan, Semra, et al. International Journal of Cancer 127.10 (2010): 2351-2359. doi: 10.1002/ijc.25248 doi: 10.1002/ijc.25248
	Gastric cancer	Tani, Nobuyuki, et al. Gan to kagaku ryoho. Cancer & chemotherapy 33.12 (2006): 1720-1722.
	Colorectal and Gastric cancer	Miotto, Elena, et al. Cancer research 64.22 (2004): doi: 10.1158/0008-5472.CAN-04-30008156-8159.
CHAT	Astrocytomas	Wu, Xiwei, et al. Cancer research 70.7 (2010): 2718-2727. doi: 10.1158/0008-5472.CAN-09-3631
	Cell lines (NE1-115 cell; NG108-15)	Quirin-Stricker, C., et al. Molecular brain research 23.3 (1994): 253-265. doi:10.1016/0169-328X(94)90232-1
COL23A1	Gastric carcinomas	Yamanoi, Kazuhiro, et al. Cancer Research 74.19 Supplement (2014): 1370-1370. doi: 10.1093/carcin/bgv013
COL4A1	Uterine leiomyomas	Maekawa, Ryo, et al. PLoS ONE 8(6): e66632. (2013): e66632. doi:10.1371/journal.pone.0066632
COL4A2	CRC	Mitchell, Susan M., et al. BMC cancer 14.1 (2014): 54. doi:10.1186/1471-2407-14-54
	Uterine leiomyomas	Maekawa, Ryo, et al. PLoS ONE 8(6): e66632. (2013): e66632. doi:10.1371/journal.pone.0066632
CREB3L1	Bladder cancer	Rose, Michael, et al. Epigenetics 9.12 (2014): 1626-1640. doi: 10.1002/path.4210
	Human papillomavirus (hrHPV)	Steenbergen, Renske DM, et al. The Journal of pathology 231.1 (2013): 53-62. doi: 10.1002/path.4210
DKK 2	Nasopharyngeal carcinoma	Li, Lili, et al. Epigenomics 0 (2014): 155-173. doi:10.2217/epi.14.79
	CRC	Silva, Ana-Luisa, et al. BMC cancer 14.1 (2014): 891. doi: 10.1186/1471-2407-14-891

	- - -	Sinha, Shriprakash. Integrative Biology 6.11 (2014): 1034-1048. doi: 10.1039/C4IB00124A
	CRC	Harada, Taku, et al. Cancer Prevention Research 7.10 (2014): 1002-1010. doi:10.1158/1940-6207.CAPR-14-0162
	CRC	Farkas, Sanja A., et al. Epigenomics 6.2 (2014): 179-191. doi:10.2217/epi.14.7
	CRC	Silva, Tiago Donizetti, et al. Oncology letters 6.6 (2013): 1687-1692. doi: 10.3892/ol.2013.1606
	Cholangiocarcinomas	Goeppert, Benjamin, et al. Hepatology 59.2 (2014): 544-554. doi: 10.1002/hep.26721
	Gastric cancer	Suzuki, Ryo, et al. Journal of gastroenterology 49.7 (2014): 1135-1144. doi: 10.1007/s00535-013-0861-7
	Medulloblastoma	Valdora, Francesca, et al. International journal of molecular sciences 14.4 (2013): 7492-7505. doi: 10.3390/ijms14047492
	Ovarian carcinoma	Zhu, Jing, et al. Carcinogenesis (2012): bgs278. doi: 10.1093/carcin/bgs278
	Primary B cell chronic lymphocytic leukaemia	Moskalev, Evgeny A., et al. BMC cancer 12.1 (2012): 213. doi: 10.1186/1471-2407-12-213
	Cylindromas/Spiradenomas	Rajan, Neil, et al. The Journal of pathology 224.3 (2011): 309-321. doi: 10.1002/path.2896
	Lung adenocarcinoma	Jung, Il Lae, et al. International journal of molecular medicine 26.1 (2010): 33-38. doi: 10.1002/path.2896
	Renal cancer	Hirata, Hiroshi, et al. Clinical Cancer Research 15.18 (2009): 5678-5687. doi: 10.1158/1078-0432
	Gastrointestinal cancer	Maehata, Tadateru, et al. World journal of gastroenterology: WJG 14.17 (2008): 2702. doi: 10.3748/wjg.14.2702
	Gastrointestinal cancer	Sato, Hironobu, et al. Carcinogenesis 28.12 (2007): 2459-2466. doi: 10.1093/carcin/bgm178
	- - -	Katoh, Masuko, and Masaru Katoh. Oncology reports 14.3 (2005): 783-787.
EDNRB	Hepatocellular Carcinoma	Chen, Haiyan, et al. Journal of Cancer 6.8 (2015): 740. doi: 10.7150/jca.11691
	head and neck squamous cell carcinoma	Hayashi, Masamichi, et al. Cancer (2015). DOI: 10.1002/cncr.29303
	head and neck squamous cell carcinoma	Mydlarz, Wojciech K., et al. Head & neck (2014). doi: 10.1002/hed.23842
	head and neck squamous cell carcinoma	Hayashi, Masamichi, et al. Annals of surgical oncology 21.9 (2014): 3124-3131. DOI 10.1245/s10434-014-3661-2
	cervical intraepithelial neoplasia	Vasiljević, Nataša, et al. Gynecologic oncology 132.3 (2014): 709-714. doi: 10.1016/j.ygyno.2014.02.001
	CRC	Chen, Cheng, et al. Diagn Pathol 8.1 (2013): 199. doi: 10.1186/1746-1596-8-199
	prostate/bladder/clear cell renal carcinomas	Watanabe, Junko, et al. Journal of cancer research and clinical oncology 140.1 (2014): 99-107. doi: 10.1007/s00432-013-1546-6
	Gatric cancer	Qu, Yiping, Siwen Dang, and Peng Hou. "Gene methylation in gastric cancer." Clinica Chimica Acta 424 (2013): 53-65. doi: 10.1016/j.cca.2013.05.002
	Oral cancer/dysplasia	Schussel, Juliana, et al. Clinical Cancer Research 19.12 (2013): 3268-3275. doi: 10.1158/1078-0432.CCR-12-3496
	head neck cancer	Roh, Jong-Lyel, et al. Clinical Cancer Research (2013).doi: 10.1158/1078-0432.CCR-12-3047
	oral/oropharyngeal squamous cell carcinoma	de Freitas Cordeiro-Silva, Melissa, et al. Molecular biology reports 39.12 (2012): 10111-10119. doi: 10.1007/s11033-012-1885-4
	head and neck squamous cell carcinoma	Roh, Jong–Lyel, et al. Head & neck 34.11 (2012): 1529-1536. doi: 10.1002/hed.21982
	oral/oropharyngeal squamous cell carcinoma	Viet, Chi T., et al. PAIN® 152.10 (2011): 2323-2332. doi: 10.1016/j.pain.2011.06.025
	Non-muscle-invasive bladder cancer	Zuiverloon, Tahlita, et al. BJU international 109.6 (2012): 941-948. doi: 10.1111/j.1464-410X.2011.10428.x
	Prostate cancer	Vasiljeviš, Nataša, et al. Disease markers 30.4 (2011): 151-161. doi: 10.3233/DMA-2011-0790
	oral/oropharyngeal squamous cell carcinoma	Guerrero-Preston, Rafael, et al. Cancer prevention research 4.7 (2011): 1061-1072. doi: 10.1158/1940-6207.CAPR-11-0006

oral/oropharyngeal squamous cell carcinoma	de Freitas Cordeiro-Silva, Melissa, et al. Molecular biology reports 38.8 (2011): 5435-5441. doi: 10.1007/s11033-011-0698-1.
Gastric cancer	Tao, Kaixiong, et al. Medical Oncology 29.1 (2012): 107-112. doi: 10.1007/s12032-010-9805-8
Glioma	De Tayrac, Marie, et al. Clinical Cancer Research 17.2 (2011): 317-327. doi: 10.1158/1078-0432.CCR-10-1126
Oral cancer	Pattani, Kavita Malhotra, et al. Cancer Prevention Research 3.9 (2010): 1093-1103. doi: 10.1158/1940-6207.CAPR-10-0115.
oral/oropharyngeal squamous cell carcinoma	Kaur, Jatinder, et al. International Journal of Cancer 127.10 (2010): 2367-2373. doi: 10.1002/ijc.25377.
head and neck carcinogenesis	Demokan, Semra, et al. International Journal of Cancer 127.10 (2010): 2351-2359. doi: 10.1002/ijc.25248.
Prostate cancer	Phé, Veronique, Olivier Cussenot, and Morgan Rouprêt. BJU international 105.10 (2010): 1364-1370. doi: 10.1111/j.1464-410X.2009.09167.x
Urothelial cell carcinomas	Phé, Véronique, Olivier Cussenot, and Morgan Rouprêt. BJU international 104.7 (2009): 896-901. doi: 10.1111/j.1464-410X.2009.08696.x.
Bladder cancer	Wolff, Erika M., et al. Cancer research 68.15 (2008): 6208-6214. doi: 10.1158/0008-5472.CAN-07-6616
esophageal squamous cell carcinoma	Zhao, B-J., et al. Diseases of the Esophagus 22.1 (2009): 55-61. doi: 10.1111/j.1442-2050.2008.00848.x
leukemia	Hsiao, P., et al. Chinese Journal of Physiology 51.1 (2008): 27.
Prostate cancer	Ellinger, Jörg, et al. Urology 71.1 (2008): 161-167. doi: 10.1016/j.urology.2007.09.056
Nasopharyngeal carcinoma	Zhou, Liang, et al. Oncology 72.5-6 (2007): 357-363. doi: 10.1159/000113146.
Prostate cancer	Bastian, Patrick J., et al. The Journal of urology 179.2 (2008): 529-535. doi:10.1016/j.juro.2007.09.038
Non-muscle-invasive bladder cancer	Friedrich, M. G., et al. Der Urologe. Ausg. A 46.7 (2007): 761-768. DOI: 10.1007/s00120-007-1360-3
Bladder cancer	Yates, David R., et al. Clinical Cancer Research 13.7 (2007): 2046-2053. doi: 10.1158/1078-0432.CCR-06-2476
Prostate cancer	Rogers, Craig G., et al. The Journal of urology 176.5 (2006): 2280-2284. doi:10.1016/j.juro.2006.07.047
endometrial carcinoma	Zhu, J. Z., et al. Zhonghua bing li xue za zhi Chinese journal of pathology 35.8 (2006): 489.
Prostate cancer	Bastian, Patrick J., et al. European urology 51.3 (2007): 665-674. doi:10.1016/j.eururo.2006.08.008
Renal cancer	Pflug, Beth R., et al. Cancer letters 246.1 (2007): 139-148. doi:10.1016/j.canlet.2006.02.007
Hepatocellular Carcinoma	Hsu, Li-Sung, et al. Oncology reports 15.2 (2006): 507-511.DOI: 10.3892/or.15.2.507
Lung cancer	Chen, Shu-Chen, et al. Oncology reports 15.1 (2006): 167-172. DOI: 10.3892/or.15.1.167
Non-muscle-invasive bladder cancer	Friedrich, Martin G., et al. European Journal of Cancer 41.17 (2005): 2769-2778. doi:10.1016/j.ejca.2005.07.019
Bladder cancer	Friedrich, Martin G., et al. Clinical Cancer Research 10.22 (2004): 7457-7465. doi: 10.1158/1078-0432.CCR-04-0930
Prostate cancer	Singal, Rakesh, et al. Oncology reports 12.3 (2004): 631-637. DOI: 10.3892/or.12.3.631
Prostate cancer	Woodson, Karen, et al. The Prostate 60.1 (2004): 25-31. DOI: 10.1002/pros.20013
Prostate cancer	Woodson, Karen, Jeffrey Hanson, and Joseph Tangrea. Cancer letters 205.2 (2004): 181-188. doi:10.1016/j.canlet.2003.11.027
Prostate cancer	Yegnasubramanian, Srinivasan, et al. Cancer research 64.6 (2004): 1975-1986. doi: 10.1158/0008-5472.CAN-03-3972
Medulloblastoma	Lindsey, Janet C., et al. Carcinogenesis 25.5 (2004): 661-668. doi: 10.1093/carcin/bgh055
Prostate adenocarcinoma	Jeronimo, C., et al. Journal of clinical pathology 56.1 (2003): 52-55. doi:10.1136/jcp.56.1.52
Nasopharyngeal carcinoma	Lo, Kwok-Wai, et al. International journal of cancer 98.5 (2002): 651-655. doi:10.1136/jcp.56.1.52

	Bladder/Colon/Prostate cancer	Pao, Martha M., et al. Human molecular genetics 10.9 (2001): 903-910. doi:10.1136/jcp.56.1.52
	Prostate cancer	Nelson, Joel B., et al. Cancer research 57.1 (1997): 35-37.
EN2	Myelogenous leukemia	Abollo-Jiménez, Fernando, et al. Cell Cycle 13.11 (2014): 1717-1726. doi: 10.4161/cc.28629
	Pilocytic astrocytomas	Lambert, Sally R., et al. Acta neuropathologica 126.2 (2013): 291-301. doi: 10.1007/s00401-013-1124-7
	renal cancer	Slater, Amy A., et al. Epigenetics 8.3 (2013): 252-267. doi: 10.4161/epi.23817
	follicular limphoma	Bennett, Lynda B., et al. Genes, Chromosomes and Cancer 48.9 (2009): 828-841. doi: 10.1002/gcc.20687.
FBN1	CRC	Li, Wen-han, et al. Disease markers (2015). doi: 10.1155/2015/657570
	cholangiocarcinoma	Andresen, Kim, et al. Hepatology 61.5 (2015): 1651-1659. doi: 10.1002/hep.27707
	CRC	Bethge, Nicole, et al. Epigenetics 9.3 (2014): 428-436. doi: 10.4161/epi.27554
	CRC	Guo, Qi, et al. Medical Oncology 30.4 (2013): 1-5. doi: 10.1007/s12032-013-0695-4
	CRC/Adenomas	Lind, Guro E., et al. Mol Cancer 10 (2011): 85. doi: 10.1186/1476-4598-10-85.
	Prostate cancer	WANG, YIPENG, et al. Annals of the New York Academy of Sciences 1058.1 (2005): 162-185. DOI: 10.1196/annals.1359.024
	Prostate cancer cell lines	Wang, Yipeng, et al. Neoplasia 7.8 (2005): 748-IN7. doi:10.1593/neo.05289
FBN2	clear cell renal cell carcinoma	Ricketts, Christopher J., Victoria K. Hill, and W. Marston Linehan. (2014): e85621. doi: 10.1371/journal.pone.0085621
	CRC	Kim, Tae-Oh, et al. International journal of molecular medicine 31.5 (2013): 1255-1261. doi: 10.3892/ijmm.2013.1317.
	CRC	Hibi, Kenji, et al. Anticancer research 32.10 (2012): 4371-4374.
	Breast carcinoma	Kikuyama, Mizuho, et al. Cancer letters 322.2 (2012): 204-212. doi: 10.1016/j.canlet.2012.03.016
	CRC	Yi, Joo Mi, et al. Tumor Biology 33.2 (2012): 363-372. doi: 10.1007/s13277-011-0302-2
	Renal cell carcinoma	Morris, Mark R., et al. Oncogene 30.12 (2011): 1390-1401. doi: 10.1038/onc.2010.525
	CRC	Yagi, Koichi, et al. Clinical Cancer Research 16.1 (2010): 21-33. doi: 10.1158/1078-0432.CCR-09-2006
	esophageal squamous cell carcinoma	Tsunoda, Shigeru, et al. Oncology reports 21.4 (2009): 1067-1073.
	Lung cancer	Cortese, Rene, et al. The international journal of biochemistry & cell biology 40.8 (2008): 1494-1508. doi: 10.1016/j.biocel.2007.11.018
	non-small cell lung cancer	Chen, Hong, et al. Lung Cancer 50.1 (2005): 43-49.
	pancreatic cancer	Hagihara, Atsushi, et al. Oncogene 23.53 (2004): 8705-8710.
GAD2	CRC	Li, Hai, et al. Oncology reports 28.1 (2012): 99-104. doi: 10.3892/or.2012.1779
	follicular limphoma	Bennett, Lynda B., et al. Genes, Chromosomes and Cancer 48.9 (2009): 828-841. doi: 10.1002/gcc.20687
GHSR	lung/breast/prostate/pancreas/CRC/glioblastoma/B cell chronic lymphocytic leukaemia	Moskalev, Evgeny A., et al. Oncotarget 6.6 (2015): 4418.
	Breast carcinoma	de Groot, Jolien S., et al. Cellular Oncology 37.4 (2014): 297-303.doi: 10.1007/s13402-014-0189-1
	Breast carcinoma	Botla, Sandeep Kumar, et al. Breast cancer research and treatment 135.3 (2012): 705-713. doi: 10.1007/s10549-012-2197-z
	Breast carcinoma	Ordway, Jared M., et al. PLoS One 2.12 (2007): e1314.
GNG7	head and neck cancer	Demokan, Semra, et al. International journal of oncology 42.4 (2013): 1427-1436. doi: 10.3892/ijo.2013.1808

	head and neck squamous cell carcinoma	Hartmann, Sylvia, et al. Journal of applied genetics 53.2 (2012): 167-174. doi: 10.1007/s13353-011-0079-4
	oesophageal cancer	Ohta, M., et al. British journal of cancer 98.2 (2008): 410-417. doi: 10.1038/sj.bjc.6604124
GRIA4	oropharyngeal squamous cell carcinomas	Kostareli, Efterpi, et al. The Journal of clinical investigation 123.6 (2013): 2488. doi: 10.1172/JCI67010
	follicular limphoma	Bennett, Lynda B., et al. Genes, Chromosomes and Cancer 48.9 (2009): 828-841. doi: 10.1002/gcc.20687
GRIK3	lung adenocarcinoma	Pradhan, Meeta P., Akshay Desai, and Mathew J. Palakal. BMC systems biology 7.1 (2013): 141. doi: 10.1186/1752-0509-7-141
GRM6	clear cell renal cell carcinoma	Arai, Eri, et al. Carcinogenesis 33.8 (2012): 1487-1493. doi: 10.1093/carcin/bgs177
GRM7	Chronic lymphocytic leukemia	Rush, Laura J., et al. Cancer research 64.7 (2004): 2424-2433.
HRH2	Gastric cancer	Nomura, Tomoe, et al. BMC gastroenterology 13.1 (2013): 1. doi: 10.1186/1471-230X-13-1
	pancreatic cancer	Ginestà, Mireia M., et al. Journal of clinical pathology (2012): jclinpath-2012. doi: 10.1136/jclinpath-2012-201123
ITGA4	Colitis-associated cancer	Gerecke, Christian, et al. Journal of cancer research and clinical oncology (2015): 1-11.
	CRC/Adenomas	Zhang, Xie, et al. World journal of gastroenterology: WJG 21.9 (2015): 2629. doi: 10.3748/wjg.v21.i9.2629
	Prostate cancer	Mostafavi-Pour, Z., et al. Pathology & Oncology Research (2015): 1-7. doi: 10.1007/s12253-015-9917-8
	Breast carcinoma	Do, Sung-Im, et al. Tumor Biology 35.7 (2014): 7079-7084. doi: 10.1007/s13277-014-1952-7
	cervical cancer	Hansel, Alfred, et al. PloS one 9.3 (2014): e91905. doi: 10.1371/journal.pone.0091905
	lymphocytic leukemia cells	Zucchetto, Antonella, et al. Blood 122.19 (2013): 3317-3321. doi: 10.1182/blood-2013-06-507335
	Gastric cancer	Qu, Yiping, Siwen Dang, and Peng Hou. Clinica Chimica Acta 424 (2013): 53-65. doi: 10.1016/j.cca.2013.05.002
	CRC	Ahmed, Deeqa, et al. Clinical and translational gastroenterology 3.12 (2012): e27. doi: 10.1038/ctg.2012.21
	Breast carcinoma	Lian, Zhen-Qiang, et al. International journal of oncology 41.2 (2012): 629-638. doi: 10.3892/ijo.2012.1464
	CRC	Chang, Eugene, et al. Hepato-gastroenterology 57.101 (2009): 720-727.
	cholangiocarcinoma	Uhm, Kyung-Ok, et al. Journal of cancer research and clinical oncology 136.2 (2010): 187-194. doi: 10.1007/s00432-009-0646-9.
	acute myeloid leukemia (AML) and chronic myeloid leukemia (CML)	Uhm, Kyung-Ok, et al. Journal of Korean medical science 24.3 (2009): 493-497. doi: 10.3346/jkms.2009.24.3.493.
	Colorectal polyp	Ausch, Christoph, et al. Clinical chemistry 55.8 (2009): 1559-1563. doi: 10.1373/clinchem.2008.122937
	Gastric cancer	Kim, Ji Hun, et al. Oncology reports 21.5 (2009): 1251-1259.
	hepatocellular carcinoma	Ko, Eunkyung, et al. Cancer Epidemiology Biomarkers & Prevention 17.9 (2008): 2260-2267. doi: 10.1158/1055-9965.EPI-08-0236
	esophageal squamous cell carcinoma	Lee, Eun Ju, et al. International journal of cancer 123.9 (2008): 2073-2079. doi: 10.1002/ijc.23598.
	bladder cancer	Yu, Jian, et al. Clinical Cancer Research 13.24 (2007): 7296-7304.
	Gastric cancer	Park, Jinah, et al. Oncogene 23.19 (2004): 3474-3480.
	fibrosarcoma cell lines	Sato, Shuichi, et al. Oncogene 17.1 (1998): 105-113.
	Mouse liver tumors	Akama, Tomoya O., et al. Cancer research 57.15 (1997): 3294-3299.
	- - -	AUDET, JEAN-FRANCOIS, et al. DNA and cell biology 13.11 (1994): 1071-1085.
ITGA8	CRC	Kok-Sin, Teow, et al. Oncology reports 34.1 (2015): 22-32. doi: 10.3892/or.2015.3993

	Ovarian cancer	Cai, Li-yi, et al. Life sciences 80.16 (2007): 1458-1465.
JAM3	cervical cancer	Boers, A., et al. British journal of cancer 111.6 (2014): 1095-1101. doi: 10.1038/bjc.2014.392
	cervical cancer	Vasiljević, Nataša, et al. Gynecologic oncology 132.3 (2014): 709-714. doi: 10.1016/j.ygyno.2014.02.001
	cervical cancer	Eijsink, J. J. H., et al. International Journal of Cancer 130.8 (2012): 1861-1869. doi: 10.1002/ijc.26326
	cervical cancer	Eijsink, J. J. H., et al. Gynecologic oncology 120.2 (2011): 280-283. doi: 10.1016/j.ygyno.2010.10.029
KCNQ5	CRC	Ashktorab, Hassan, et al. Epigenetics 8.8 (2013): 807-815. doi: 10.4161/epi.25497.
LAMA1	Breast carcinoma	Simonova, O. A., et al. Molecular Biology 49.4 (2015): 598-607. doi: 10.7868/S0026898415040163
	small intestinal neuroendocrine tumors	Fotouhi, Omid, et al. Epigenetics 9.7 (2014): 987-997. doi: 10.4161/epi.28936
	CRC	Ashktorab, Hassan, et al. Epigenetics 8.8 (2013): 807-815. doi: 10.4161/epi.25497.
	CRC	Kim, Young-Ho, et al. Annals of surgical oncology 18.8 (2011): 2338-2347. doi: 10.1245/s10434-011-1573-y
LHCGR	polycystic ovary syndrome	Wang, Peng, et al. Endocrinology 155.4 (2014): 1445-1452. doi: 10.1210/en.2013-1764
MMP9	Lung cancer	Cheng, X., et al. Oncogene (2015). doi: 10.1038/onc.2015.14
	Gastric cancer	Wu, Feng-lei, et al. Cancer letters 363.1 (2015): 7-16. doi: 10.1016/j.canlet.2015.01.006
	Skin cancer	Han, Yantao, et al. Mutagenesis 30.2 (2015): 287-296. doi: 10.1093/mutage/geu071
	Breast carcinoma	Salz, Tal, et al. Molecular Cancer Research 13.3 (2015): 461-469. doi: 10.1158/1541-7786.MCR-14-0389.
	Breast carcinoma	Rizwani, Wasia, et al. (2014): e100888. doi: 10.1371/journal.pone.0100888.
	Type 2 Diabetes Mellitus	Singh, Kanhaiya, et al. The international journal of lower extremity wounds (2014): 1534734614534980. doi: 10.1177/1534734614534980
	Renal cell carcinoma	Chen, Xuanyu, et al. Journal of cancer research and clinical oncology 140.8 (2014): 1295-1304. doi: 10.1007/s00432-014-1690-7.
	Stomach cancer	Nemtsova, M. V., et al. Klinicheskaia laboratornaia diagnostika 11 (2013): 12-15.
	CRC	Oh, Jong-Tae Kim, et al. (2014). DOI: 10.18632/oncotarget.1714
	Extravillous trophoblast (EVT) cell invasion	Takahashi, H., et al. Placenta 35.3 (2014): 163-170. doi: 10.1016/j.placenta.2013.12.009
	diabetic retinopathy	Kowluru, Renu A., Julia M. Santos, and Manish Mishra. BioMed research international 2013 (2013). doi: 10.1155/2013/635284.
	CRC	Kuhmann, Christine, et al. Human molecular genetics 23.8 (2014): 2043-2054. doi: 10.1093/hmg/ddt599
	Lung cancer	Xu, Shun, et al. Oncology reports 31.1 (2014): 79-86. doi: 10.3892/or.2013.2799
	Gastric cancer	Hong, Seung-Hyun, et al. Oncotarget 4.10 (2013): 1791.
	Human fibroblast cells	Poplineau, Mathilde, et al. Molecular carcinogenesis 54.1 (2015): 24-34. doi: 10.1002/mc.22071
	Ulcerative colitis	Koukos, Georgios, et al. Gastroenterology 145.4 (2013): 842-852. doi: 10.1053/j.gastro.2013.07.001.
	Medulloblastoma	Asuthkar, S., et al. Oncogene 33.15 (2014): 1922-1933. doi: 10.1038/onc.2013.151.
	fibrosarcoma cell lines	Poplineau, M., et al. Cell proliferation 46.2 (2013): 127-136.doi: 10.1111/cpr.12021
	CRC	Lin, Yan-Wei, et al. The Journal of pathology 230.3 (2013): 277-290. doi: 10.1002/path.4179.
	acute myeloid leukemia and myelodysplasia	Bernal, Teresa, et al. Experimental hematology 41.2 (2013): 172-179. doi: 10.1016/j.exphem.2012.10.005.

	Prostate cancer	Patra, Aditi, et al. Clinical epigenetics 2.2 (2011): 339-348. doi: 10.1007/s13148-010-0019-x.
	non-small cell lung cancer	Yoo, Jin Young, et al. Journal of neuro-oncology 109.2 (2012): 219-227. doi: 10.1007/s11060-012-0890-8.
	prostate cancer cell line	Purcell, Daniel J., et al. Journal of cellular biochemistry 113.7 (2012): 2406-2414. doi: 10.1002/jcb.24114
	acute myeloid leukemia and myelodysplasia	Eberle, Franziska C., et al. Haematologica 96.4 (2011): 558-566. doi: 10.3324/haematol.2010.033167
	Hepatocellular Carcinoma	Huang, Jian, et al. The Journal of clinical investigation 120.1 (2010): 223. doi: 10.1172/JCI38012
	MLL-associated leukemias	Li, Zejuan, et al. Cancer research 69.3 (2009): 1109-1116. doi: 10.1158/0008-5472.CAN-08-3381.
NOS1	neuroblastoma cell differentiation	Das, Sudipto, et al. Cancer research 70.20 (2010): 7874-7881. doi: 10.1158/0008-5472.CAN-10-1534
NPBWR1	non-small cell lung cancer	Sandoval, Juan, et al. Journal of Clinical Oncology 31.32 (2013): 4140-4147. doi: 10.1200/JCO.2012.48.5516
	Prostate cancer	Cottrell, Susan, et al. The Journal of urology 177.5 (2007): 1753-1758. doi:10.1016/j.juro.2007.01.010
NPY2R	Urothelial cell carcinomas	Yang, Tse-Yen, et al. Environmental research 128 (2014): 57-63. doi: 10.1016/j.envres.2013.10.006
PDE8B	cervical cancer	Huang, Rui-Lan, et al. PLoS One 7.7 (2012): e41060. doi: 10.1371/journal.pone.0041060.
PLCB1	head and neck cancer	Guerrero-Preston, Rafael, et al. Epigenetics 9.7 (2014): 1031-1046. doi: 10.4161/epi.29025.
PRKCB	Adult T cell leukemia/lymphom	Kataoka, Keisuke, et al. Nature genetics (2015). doi: 10.1038/ng.3415
	lung adenocarcinoma	Pradhan, Meeta P., Akshay Desai, and Mathew J. Palakal. BMC systems biology 7.1 (2013): 141. doi: 10.1186/1752-0509-7-141
PTGFR	CRC	Cebola, Inês, et al. Clinical epigenetics 7.1 (2015): 1-11. doi: 10.1186/s13148-015-0110-4
SCTR	Breast carcinoma	Kang, Seongeun, et al. International journal of oncology 47.5 (2015): 1923-1931. doi: 10.3892/ijo.2015.3164.
	sporadic CRC	Karpinski, Pawel, et al. Tumor Biology 32.4 (2011): 653-659. doi: 10.1007/s13277-011-0165-6
	Prostate cancer	Devaney, James, et al. Cancer Epidemiology Biomarkers & Prevention 20.1 (2011): 148-159. doi: 10.1158/1055-9965.EPI-10-0719.
	CRC	Mayor, R., et al. British journal of cancer 100.10 (2009): 1534-1539. doi: 10.1038/sj.bjc.6605045
	CRC	Karpinski, Pawel, et al. Molecular Cancer Research 6.4 (2008): 585-591. doi: 10.1158/1541-7786.MCR-07-2158
SDC2	Gastric cancer	Chong, Yosep, et al. Oncology reports 31.6 (2014): 2535-2544. doi: 10.3892/or.2014.3133.
	CRC	Mitchell, Susan M., et al. BMC cancer 14.1 (2014): 54. doi: 10.1186/1471-2407-14-54
	osteonecrosis of the jaw	Gambino, A., et al. Annali di stomatologia 5.2 Suppl (2014): 11. doi: 10.1016/j.mrgentox.2013.07.003.
	CRC	Oh, TaeJeong, et al. The Journal of Molecular Diagnostics 15.4 (2013): 498-507. doi: 10.1016/j.jmoldx.2013.03.004
	head and neck squamous cell carcinoma	Worsham, Maria J., et al. Otolaryngology--Head and Neck Surgery 149.3 (2013): 409-416. doi: 10.1177/0194599813490895
SFRP1_2	CRC/Adenomas	Patai, Árpád V., et al. PloS one 10.8 (2015): e0133836. doi: 10.1371/journal.pone.0133836
SLC32A1	pancreatic cancer	Shimizu, Hideyuki, et al. Biochemical and biophysical research communications 411.1 (2011): 162-167. doi: 10.1016/j.bbrc.2011.06.121.
VCAN	CRC	Torabi, Keyvan, et al. Carcinogenesis 36.10 (2015): 1103-1110. doi: 10.1093/carcin/bgv115.
	Ovarian clear cell carcinoma	Yamaguchi, K., et al. Oncogene 29.12 (2010): 1741-1752. doi: 10.1038/onc.2009.470
	Prostate cancer	Read, Jason T., et al. Journal of Biological Chemistry 282.44 (2007): 31954-31963. doi: 10.1074/jbc.M702099200
	ulcerative colitis–associated colorectal cancers	Maia, Lara, et al. Cancer genetics and cytogenetics 162.1 (2005): 68-73. doi:10.1016/j.cancergencyto.2005.02.017

Ulcerative colitis	Issa, Jean-Pierre J., et al. Cancer research 61.9 (2001): 3573-3577.
CRC	Toyota, Minoru, et al. Cancer research 59.10 (1999): 2307-2312.
CRC	Adany, Roza, and Renato V. Iozzo. Biochemical and biophysical research communications 171.3 (1990): 1402-1413. doi:10.1016/0006-291X(90)90841-A

Supplementary Materials and methods

DNA extraction and bisulfite conversion from tissue samples

Genomic DNA was extracted from tumoral and peritumoral tissue using the DNeasy Blood & Tissue Kit (Qiagen) and bisulfite converted using EZ DNA Methylation Gold Kit ™ (Zymo Research) according to the manufacturer's instructions. Quality control and quantification of DNA were performed before and after bisulfite conversion. DNA was quantified with NanoDrop (NanoDrop Products Thermo Scientific Wilmington, DE) and by fluorometric reading (Quant-iT™ PicoGreen® dsDNA Assay Kit); quality was assessed by visualisation of genomic DNA on 1% agarose gel electrophoresis. Only DNA samples not fragmented and with a concentration higher than 50 ng/µl were subsequently processed.

Samples for pyrosequencing methylation validation

A validation of *GRIA4*, *SLC8A1*, *SYN3* CpG islands hypermethylation was performed in 78 CRCs and the respective 78 peritumoral samples. Tissue samples were collected at the Department of Surgical Sciences, University of Cagliari (Italy) (**Figure 1**).

Samples for transcriptome analysis

RNA was extracted from remaining tissue available from the first methylome step using the RNeasy Lipid Tissue Mini Kit (Qiagen). Seventeen tumoral and 2 peritumoral samples achieved an optimal RIN (>7) and were processed for further analyses. In addition, 2 new peritumoral samples were recruited for the whole genome gene expression analysis (**Figure 1**).

Pyrosequencing analysis

Primers were designed for *SYN3*, *SLC8A1* and *GRIA4* using the PyroMark® Assay Design SW 2.0 (Qiagen) (**Supporting Information Table 3**). Amplification was carried out using the Platinum® Taq (Life technologies). PCR products were purified on the PyroMark Q24 Vacuum Workstation according to manufacturer protocol and annealed with the sequencing primer before being run on the PyroMark Q24 (Qiagen). Pyrograms were analyzed using PyroMark Q24 Software.

Whole genome gene expression analysis

The transcriptome analysis was performed using the HumanHT-12 v4 Expression BeadChip Kit (Illumina, San Diego, CA) according to manufacturer’s protocol. The RNA, quantified by spectrophotometric (NanoDrop) and fluorometric (PicoGreen) reading, is qualitatively checked by

means of the tool Bioanalyzer2100 (Agilent Technologies), which provides an index of integrity of the RNA (RNA Integrity Number, RIN), ranging between 0 (complete degradation) and 10 (excellent quality); in our study, were further processed only the samples with RIN>7.

200 ng of RNA were amplified (Illumina TotalPrep RNA Amplification Kit), labeled and hybridized on Illumina microarray. Even with data from transcriptome analysis an Unsupervised Hierarchical Clustering is carried out and the expression values are often shown by HeatMap

qRT-PCR analysis

Retro-transcription was performed starting from 1µg RNA/sample using the High Capacity Kit (Applied Biosystems, Carlsbad, CA, USA). Gene expression was assessed by quantitative RT-PCR using Express Sybr Green (Invitrogen, Paisley, UK) for each gene tested and for the endogenous TFRC. The primer sequences are reported in **Supporting Information Table 3.**

Microsatellite instability analysis

The microsatellite instability analysis was conducted by using the "MSI Analysis System, Version 1.2" (Promega Italia). Amplification were carried out according to manufacturer's protocol. The PCR products were then run on an ABI PRISM® 3100 - Applied Biosystems® 3130 Genetic Analyzer, and the output data were analyzed with GeneMapper® software (Applied Biosystems).

Genetic mutations screening

The genetic mutation screening in adenomas was conducted as previously reported [11,48].

To search for *KRAS* mutations in CRCs, we amplified two fragments corresponding to the exons 2 and 3 (codon 1-97) (annealing 60 °C). Primer sequences are reported in **Supporting Information Table 3**.

PCR products were amplified with High-Fidelity Taq polymerase (Platinum® Taq DNA Polymerase High Fidelity, Invitrogen), purified (by exonuclease 1 and shrimp alkaline phosphatase) and sequenced by fluorescent based Sanger's direct sequencing in an ABI 3130 DNA capillary sequencer.

CIMP phenotype definition

To infer a CIMP phenotype definition from the output data obtained by using Infinium® HumanMethylation450 BeadChip in the 18 CRCs and 21 adenomas, we referred to the SALSA® MLPA® kit (MRC-Holland, Amsterdam, the Netherlands) as reference panel. This ME042-C1 CIMP MS-MLPA probemix contains 31 MS-MLPA probes which detect the methylation status of

promoter regions of the following 8 genes: *CACNA1G*, *CDKN2A*, *CRABP1*, *IGF2*, *MLH1*, *NEUROG1*, *RUNX3* and *SOCS1*. In the present study, several CpG loci were analyzed for each of these gene promoters, mapping inside the MS-MLPA probe sequence. All positions/probes were scored using two different thresholds of methylation (≥ 20% or ≥ 30%), as suggested by Berg et al. [49] for defining a specific position/probe as methylated. To define a gene as methylated, at least one probe/position has to be methylated. To assign the CIMP positivity to a sample, at least 60% of the genes in the panel must be labelled as CIMP positive.

Statisical analysis

Transcriptome

The quantile normalization algorithm was applied on the data set. Differential expression analyses were carried out between CRCs (T) and peritumoral samples (P). The differential expression is given in fold change (FC). False Discovery Rate of 0.05 was chosen as a threshold for significance.

RNAseq data and differential expression analysis

To facilitate cross-sample comparison and differential expression analysis, the Upper Quartile normalized FPKM (UQ-FPKM) values has been obtained from TCGA open-access data for 478 CRC solid tumor tissues and 41 solid normal tissues from the GDC (Genomic Data Commons) data portal. The data downloading and differential expression analysis has been conducted using the Bioconductor "*TCGAbiolinks*" package [50].

qRT-PCR

Data are expressed as mean ± standard deviation (SD) or mean ± standard error (SEM). Analysis of significance was done by t Student's test and by One-Way ANOVA using the GraphPad software (La Jolla, California). P-values were considered significant at $p < 0.05$.