UNIVERSITÀ DEGLI STUDI DELL'INSUBRIA

PhD Program - XXIX Cycle
Experimental and Translational Medicine

UHRF1 coordinates DNA methylation and histone post-translational modifications in colon cancer

Supervisor:

Dr. Ian Marc Bonapace

Candidate:

Dr. Filippo Macchi
Uni ID: 614067

Academic Year 2015-2016

Table of content

Abstract

In Colorectal cancer (CRC) genetic and epigenetic alterations are tightly connected, although these interactions on the patient's outcome are not clearly understood.
A peculiar subclass of sporadic CRC tumors, is characterized by Microsatellite instability (MSI), due to hyper-methylated promoter of MHL1 gene and subsequent inactivation of the mismatch repair (MMR) mechanism, and the hypermethylation of CpG islands phenotype (CIMP), mediated by the hyper-methylation of promoter regions of several tumor suppressor genes (TSGs). This subclass carries a good prognosis and presents an inverse correlation with genomic and chromosome instability (CIN), together with a higher levels of DNA methylation at global level, compared to other CRCs. Among the epigenetic alterations, DNA methylation and histone modifications rearrangements are extremely important steps during tumorigenesis. UHRF1 is a key master epigenetic regulator that couples the maintenance of DNA methylation through the cell cycle with the histone-modification pattern. It monoubiquitinates H3K18/23 enabling the correct localization and activation of DNMT1 on the specific sites. UHRF1 is overexpressed in several cancer types mediating the hypermethylation of promoter regions of the TSGs and coordinating their heterochromatic silencing.
Relying on the idea that UHRF1 could play a crucial role in the modulation of DNA methylation changes, the overall aim of this PhD thesis was to evaluate the role of UHRF1 in the coordination of DNA methylation and histone post-translational modifications both at genome-wide and at locus specific level in CRCs.
Unexpectedly, we found that in CRC tissues UHRF1 was higher in tumors with microsatellite instability (MSI CRC), which have a better prognosis, compared to the stable ones (MSS CRC). MSI tumors were also characterized by higher levels of DNA methylation compared to the MSS. The UHRF1 knock-down in a MSI CRC cell line (RKO cells) induced an overall decrease in DNA methylation (RRB-seq analysis, pyrosequencing and MS-MLPA) both at global level and at gene promoters without affecting DNMTs levels, as observed by WB and RNA-seq analysis. ChIP experiments showed that UHRF1 depletion reduces DNMT1 binding to both repetitive elements (LINE-1) and specific gene promoters (MLH1, CDH1), decreasing H3K9me3 and increasing H3K4me3 on those hypo-methylated loci. RNA-seq data analysis showed that UHRF1 loss interferes with several important pathways, among others cell cycle, growth and proliferation. SILAC LC-MS/MS analysis showed that in RKO cells, UHRF1 loss decreases the overall presence of H3K23ub (≈ 30%) and H3K18ub (≈ 8%).

These results, together with the published findings, led us to hypothesize a model in which the loss of UHRF1 directly impairs the DNA methylation maintenance by reducing H3K18/K23ub and consequently DNMT1 activity and, indirectly, impairs the binding of Suv39H1, the histone methyl transferase (HMT) responsible for H3K9me3, to both genome-wide and promoter specific loci. These changes led to a severe chromatin rearrangement of heterochromatic signatures toward a more open and transcriptionally accessible structure, probably due to the disruption of the axis UHRF1-H3ub-DNMT1-HMTs. Our molecular data, together with the analysis performed on CRC samples, led us to speculate that the better prognosis correlated with MSI-CRC model, could reside in the UHRF1-high levels that result in a sort of protective condition for the genome integrity, maintaining the global DNA methylation level closer to the normal mucosae, and probably counteracting the hyper-methylation of TSGs.

1. Introduction

1.1. Colorectal Cancer (CRC)

1.1.1. Epidemiology and risk factors of CRC

Colorectal cancer (CRC) is the third most common cancer in men (746,000 cases, 10.0% of the total, after lung and prostate cancers), and the second in women (614,000 cases, 9.2% of the total, after breast cancer) worldwide. Almost 55% of the CRCs cases occur in more developed regions, and represent respectively the fourth leading cause of cancer death in men, and the third in women. However, the mortality is considerably lower (694,000 deaths in both sexes) compared to totality of cancer deaths (8.5% of the total) with the peculiarity of a variable rate in cancer mortality, from highest in Central and Eastern Europe (20.3 per 100,000 for men, 11.7 per 100,000 for women), to lowest in Western Africa (3.5 and 3.0, respectively) [Ferlay J., et al.; Int. J Cancer., 2015].
CRCs in Italy has increased from 18,000 new cases per year in 1990 to 25,000 new cases per year in the end of 90's, with a relative 5-years survival rate assessed around 40%-50% [Capocaccia R, et al.; Int J Epidemiol, 1997]. Between 2003-2007, the Istituto Superiore di Sanita' established a national health program regarding the CRC screening using the fecal occult blood test (FOBT), that resulted effective in reducing cause-specific mortality. Moreover, although it detects pre-cancerous adenomas, the screening with immunochemical FOBT (FIT) leads to a decrease in the CRC incidence and mortality [Giorgi Rossi P, et al.; The American Journal of Gastroenterology, 2015].

The probability of suffering from CRC is about 4%–5%. The most affected anatomic parts of Colon are the Sigmoid (25%) and the Rectum (39%) sections (Figure 1.1).
The developing risk of CRC is associated with personal predisposing features or habits such as age, chronic disease history and lifestyle [Mármol I; et al.; Int. J. Mol. Sci., 2017]:

- age; the CRC incidence increases proportionally with the individual age; early development of CRC is general related to inherited cancers;
- chronic disease history; chronic inflammation mechanism, such as Crohn's bowel syndrome or ulcerative colitis can induce colonic carcinogenesis; in this context, the gut microbiota and dysbiosis situations have a relevant role;

- lifestyle; smoking, environmental factors (such as pollution), and unbalanced diets (rich in fat and protein, and poor in fiber and fruit) predispose to colon tumor development;
- personal predisposing features; familial or inherited aberrant alterations of genetic and/or epigenetic factors, i.e. familial adenomatous polyposis (FAP), hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome, Hamartomatous Polyposis Syndromes, and Familial CRC without identifiable gene mutations [Stoffel EM, et al.; Clin Gastroenterol Hepatol. 2014].

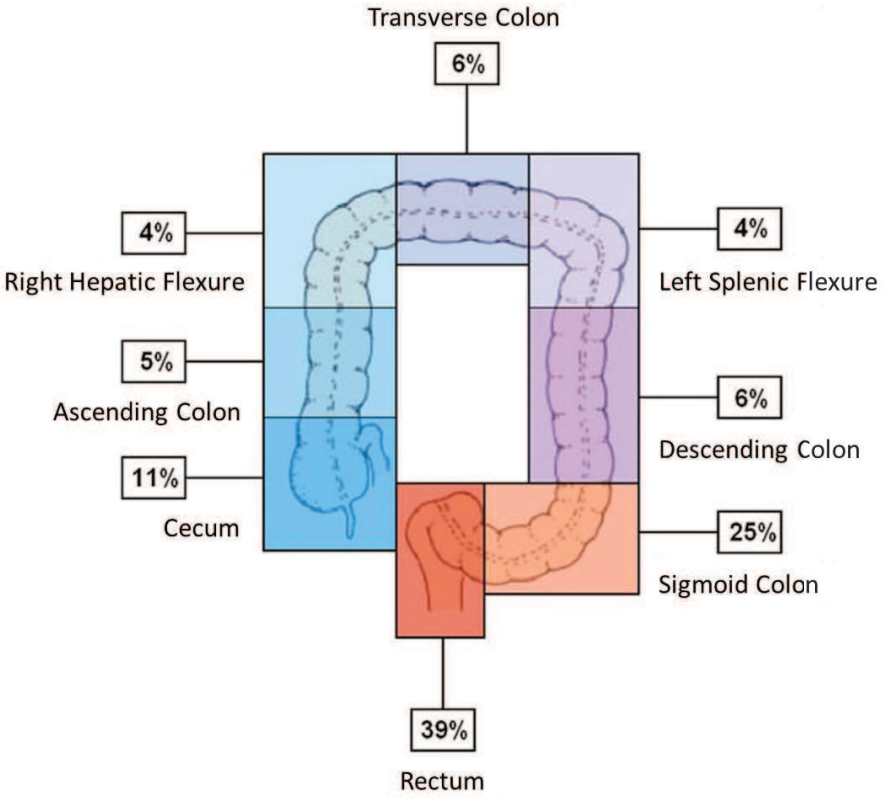


Figure 1.1. Percentages of the relative tumor onset in the different anatomic trait of large intestine. Sigmoid and Rectum show the highest percentage of tumor onset, respectively 25% and 39%. The image is adapted from "Dionigi, Basi teoriche e Chirurgia generale - Chirurgia specialistica, 4ª ed., Padova, Elsevier Masson, 2006".

1.1.2. Pathogenesis and classification of CRC

The CRC is a multistep cancer, that involves several molecular pathways leading to genomic and epigenomic instability. Mutations and/or epigenetic alterations can target oncogenes, tumor suppressor genes (TSGs) and genes related to DNA repair mechanisms [Mármol I; et al.; Int. J. Mol. Sci., 2017].

Etiological classification

From an etiological point of view, colorectal carcinomas can be classified as sporadic (70%), inherited (5%) and familial (25%) among all cases, in relation to the origin of the mutation/epigenetic alteration.

Sporadic cancers derived from point mutations that can target different heterogeneous genes [Fearon ER and Vogelstein B; Cell, 1990]. Generally, in approximately 70% of the CRC with point mutations, the first mutation occurs in the tumor suppressor gene APC (adenomatous polyposis coli), that sustains the formation of non-malignant adenomas (polyps). Upon further mutations in KRAS, TP53 and DCC genes, these polyps tend to be promoted to a malignant carcinoma [Fearon ER and Vogelstein B; Cell, 1990].

Inherited cancers are caused by hereditary mutations. The mutated gene are affected in one of the alleles, and a subsequent point mutation in the other allele could promote the onset of the tumor. The hereditary cancers are further classified in two groups, called polyposis and non-polyposis forms. Polyposis variant is mainly represented by familial adenomatous polyposis (FAP), characterized by the inherited mutation of APC gene, that promotes the formation of multiple potentially malignant polyps in the colon [Lynch HT, and de la Chapelle A; N. Engl. J. Med., 2003]. On the other side, non-polyposis forms are related to mutations in DNA repair mechanisms. The main form of hereditary non-polyposis colorectal cancer (HNPCC) is Lynch syndrome, arisen from inherited mutations in one of the alleles coding for DNA repair proteins (e.g. MSH2, MLH1, MLH6, PMS1 and PMS2). Lynch Syndrome is the most common among the inherited cancer syndromes and can be found in 2–4% of all CRC cases [Hampel H, et al.; J Clin Oncol., 2008] [Umar A, et al.; J. Natl. Cancer Inst., 2004].

Familial colorectal cancers, approximately 25% of all CRC cases, include patients with a strong familial history of cancer and/or CRC diagnosis at a young age. However, the tests of inherited mutation, in genes known as CRC-related, cannot meet clinical

criteria for any of the hereditary CRC syndromes. Although these familial CRCs are not included in any inherited cancer variant, and their etiology remains unclear, exhaustive examinations of whole genome and epigenome, by next generation sequencing (NGS) techniques, might shed light on this significant fee of CRCs [Stoffel EM, et al.; Clin Gastroenterol Hepatol. 2014].

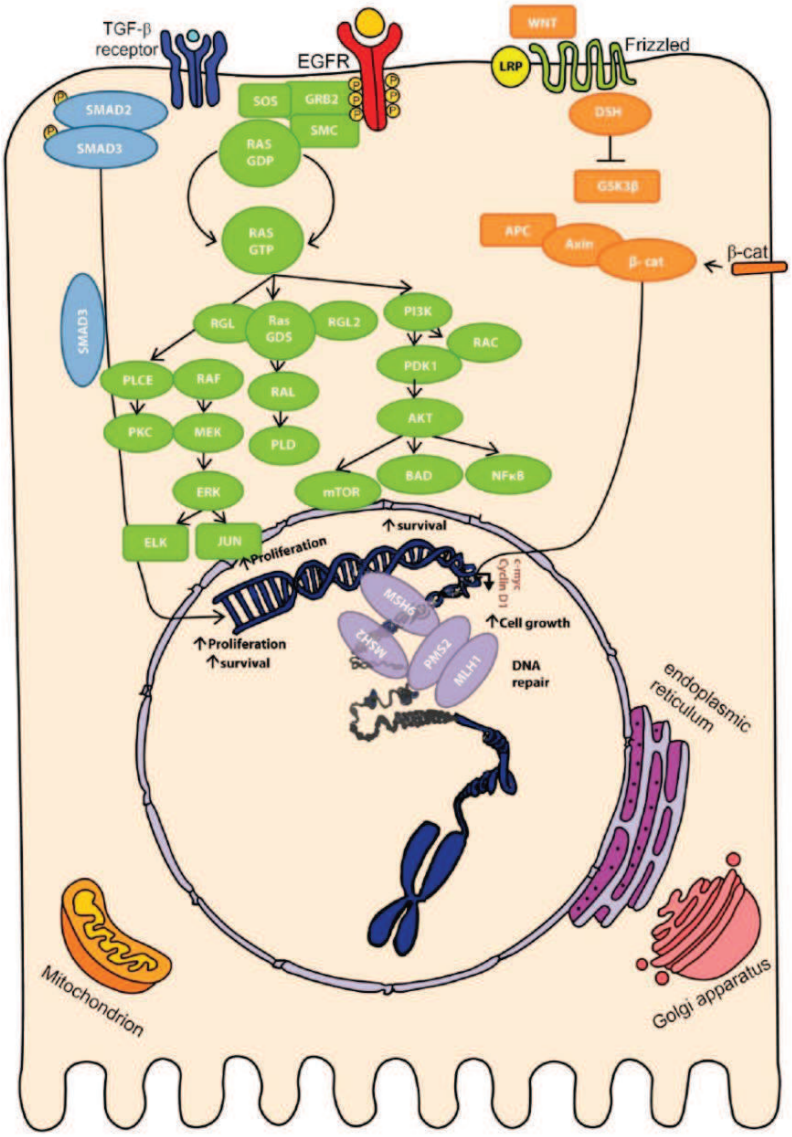


Figure 1.2. Molecular pathways involved in colorectal carcinogenesis. Mutations or epigenetic aberrations affect genes (and related encoded proteins) involved in WNT (orange), EGFR/MAPK/PI3K (green), TGF-β/SMAD (blue) or DNA repair (purple) pathways [Mármol I; et al.; Int. J. Mol. Sci., 2017].

Molecular Classification

From a molecular point of view, colorectal carcinomas can be classified in three types, named chromosomal instability (CIN), microsatellite instability (MSI) and CpG island methylator phenotype (CIMP), relying on the pathogenetic-molecular mechanisms that lead to genomic and epigenomic instability [Ogino S, and Goel A; J. Mol. Diagn., 2008]. These pathogenic mechanisms can target oncogenes, TSGs and genes related to DNA repair mechanisms (e.g. MLH1, MSH2, MSH6, PMS1 and PMS2), by mutation in common genes (e.g. APC, TP53, c-MYC, KRAS, BRAF, PIK3CA, PTEN, SMAD2/SMAD4), by epigenetic alteration (i.e. promoter hyper-methylation or chromatin silencing), and by chromosomal changes and translocations (i.e. aneuploidy and LOH), affecting important pathways such as WNT, MAPK/PI3K, TGF-β and TP53 (Figure 1.2) [Mármol I; et al.; Int. J. Mol. Sci., 2017].

The chromosomal instability (CIN) is the most common type of genomic instability observed in colon cancer, representing up to 80%-85% of all CRC cases. The hallmark of CIN is the imbalances in the number of chromosomes, that lead to aneuploidy and loss of heterozygosity (LOH) [Grady WM, and Carethers JM; Gastroenterology, 2008]. Injurious mechanisms in CIN establishment include alterations in chromosome segregation, DNA damage response and telomere dysfunction, affecting critical genes (e.g. APC, KRAS, PI3K and TP53) involved in the maintenance of correct cell function. The well-known aberrant pathway (Figure 1.3), in this kind of tumors, occurs in components of the Wnt signaling. Generally, the chain of events includes somatic or inherited mutation in one allele of APC gene, and subsequently LOH of the second normal allele. APC mutations triggers the translocation of β-catenin to the nucleus, that enhances cell growth and proliferation, laying the fundamental step in tumorigenesis. Subsequent mutational activation of KRAS and PI3K leads to a constant activation of MAP kinase pathway, forcing a higher cell proliferation rate and driving to carcinoma formation, upon mutation or LOH in TP53, the main cell-cycle checkpoint protein [Pino MS, and Chung DC; Gastroenterology 2010].

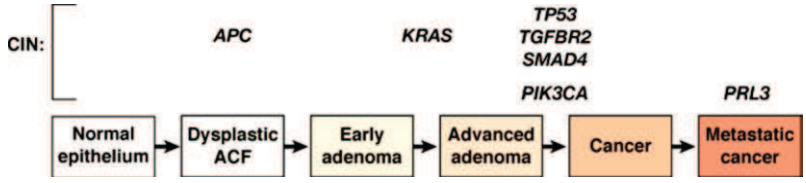


Figure 1.3. Progression of colorectal tumors with CIN [Grady WM, and Carethers JM; Gastroenterology, 2008].

In CIN progression, TGF-β signaling molecules has been reported to be severely affected: noteworthy, loss of heterozygosity at chromosome locus 18q, that harbors SMAD4 and SMAD2, and mutations of the kinase domain of TGFBR2 are often found. In metastasis processes are involved several genes, among the other PRL3 [Grady WM, and Carethers JM; Gastroenterology, 2008].

The microsatellite instability (MSI) tumors, either sporadic or from patients with Lynch syndrome, are characterized by the loss of Mismatch Repair (MMR) function, generally as one of the main early step in cancerogenesis, and by histological serrated adenoma (non-polyposis). Almost the totality of sporadic tumors loses MMR function due to the specific hypermethylation of the promoter of MLH1 gene. In contrast, patients with Lynch syndrome present a germline mutation in one of the MMR genes (e.g. MLH1, MSH2, MSH6, PMS1 and PMS2). The MMR is involved also in repairing insertion/deletion loop in short DNA chains or tandem repeats (two to five base-pair repeats), called micro-satellites. Thus, in the context of MMR absence, a hypermutable phenotype develops occurring in DNA multiple mutations [Boland CR, and Goel A; Gastroenterology, 2010]. The majority of microsatellites affected by mutation due to MMR absence are located in noncoding sequences, such as intronic regions. Nevertheless, there are several genes that harbor microsatellites in their coding sequences (e.g. TGFBR2, ACVR2, BAX, hMSH3, hMSH6), resulting in frameshift aberrations. These mutations and subsequent other additional mutations, due to the hypermutable phenotype, lead to malignant progression that generally involves Wnt signaling alterations and, in Lynch syndrome, CTNNB1 (β-catenin) mutation (Figure 1.4). Other important mutations involve EGFR signaling pathway: in sporadic tumor BRAF mutations are principally found, on the contrary in Lynch syndrome KRAS mutations are often found but not BRAF mutation. These features are used to differentiate between these 2 groups [Ogino S, et al.; Gut, 2009].
Furthermore, TGF-β signaling is altered in a not completely understood way either in MSI and non-MSI (MSS, microsatellite stable) tumors. Finally, MSI tumors harbor less LOH and are mainly diploid (non-CIN, chromosomal stable), having a better prognosis [Umar A, et al.; J. Natl. Cancer Inst., 2004]. On the contrary, MSS tumors are generally correlate to CIN-positive and worst prognosis [Ogino S, and Goel A; J. Mol. Diagn., 2008] [Ahmed D, et al.; Oncogenesis, 2013].

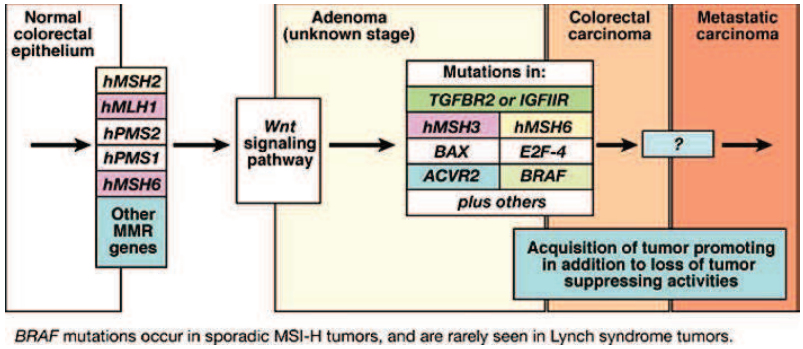


Figure 1.4. Progression of colorectal tumors with MSI [Grady WM, and Carethers JM; Gastroenterology, 2008].

Epigenetic instability is another important feature in CRCs. One of the major studied epigenetic aberrations in CRCs are the alteration in DNA methylation profile, which results paradoxically in local hypermethylation of the CpG islands (CGI) in gene promoters and global LINE-1 hypomethylation [Lao VV and Grady WM; Nat. Rev. Gastroenterol. Hepatol., 2011]. CpG island hypermethylation is the main effect of the so-called CpG island methylator phenotype (CIMP). The principal characteristic of CIMP-positive tumors is the abnormal methylation of promoter sequences of several tumor suppressor genes, leading to epigenetic silencing of these genes, by tightly allied regulation of chromatin structure [Grady WM, and Carethers JM; Gastroenterology, 2008]. CIMP tumors in part overlap with, or are comprehensive of, sporadic MSI tumors due to hypermethylation of MLH1 gene promoter. A full understanding of the progression steps in CIMP-related CRC is not clear (Figure 1.5), however seems that hyperplastic Aberrant Crypt Foci (ACF) may be the initial lesion in this kind of tumors, in contrast with the dysplastic ACF, typically arisen from an altered Wnt signaling. In this initiation phase the hypermethylation of promoter genes (e.g. MGMT, EVL, HLTF, SFRP2, SLC5A8, and MINT1) drives the colorectal development. Subsequently hypermethylation of MLH1 promoter might lead to the development of a serrated adenoma, in accordance with MSI-CRCs progression.

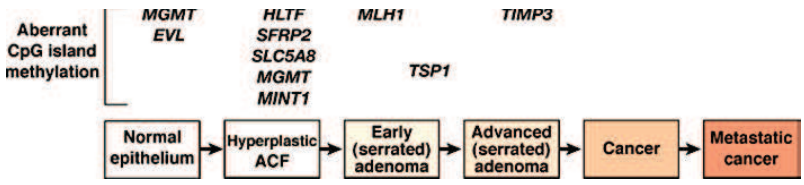


Figure 1.5. Conceptual progression of colorectal tumors with CpG island Methylator Phenotype (CIMP) [Grady WM, and Carethers JM; Gastroenterology, 2008].

Hypermethylation of TSP1 and TIMP3 promoters could help to drive the progression of CIMP tumors. Beyond the not completely understanding of the progression steps, CIMP tumors are particularly interesting due to their tight relationship or overlapping with other epigenetic and genetic aberration. An example of the combined effect of epigenetics and genetics in CIMP-CRCs development is the strong association observed in many CIMP-positive tumors with microsatellite instability (MSI) and BRAF mutations [Weisenberger DJ, et al.; Nat. Genet., 2006].

Genetics and epigenetics alteration are not mutual exclusive events in CRCs, as well as in other tumor types. In CRCs both alterations cooperate in its development and progression, and notably more aberrant methylation events are found than point mutations [Lao VV and Grady WM; Nat. Rev. Gastroenterol. Hepatol., 2011]. This is the case of the CRC-derived cell lines RKO (MSI CRC model, CIMP-high, CIN-negative, BRAF-V600E, PI3KCA-H1047R), object of this PhD thesis [Ahmed D., et al.; Oncogenesis, 2013]. Nevertheless, the effect of the interaction between genetics and epigenetics alteration on the patient's outcome is not completely understood [Wu C and Bekaii-Saab T; Chemotherapy Research and Practice, 2012].

1.2. Concepts of Epigenetic

In the past, the term 'epigenetic', originally coined by Waddington in 1942, referred to the causal mechanisms by which the environment alters the manifestation of a phenotype modifying the subtended genotype outcome. Nowadays, 'epigenetics' defines the study of inheritable modifications of gene expression that do not involve changes in the primary nucleotide sequence and/or in the number of copies of the DNA [Bird, A, et al.; Genes & Development, 2002] [Riggs AD, et al.; Regulation in Epigenetic Mechanisms of Gene; Cold Spring Harbor, 1996]. Conventionally, all the modifications and/or molecular mechanisms involved in the regulation of gene expression have considered the epigenome, including either DNA methylation (CpG methylation), and chromatin modifications (i.e. hPTMs, post-translational modifications of histones, and/or replacement with histone variants). Furthermore, recent data indicate that even non-coding RNAs (e.g. miRNAs, lncRNAs, ceRNAs) play a pivotal role in the fine-tuning of the epigenetic regulation of gene expression. Altogether these epigenetic marks define a peculiar chromatin profile that delineates the cell identity. In other term, the heritable instructions that determine spatial and temporal changes of gene activation and repression, leads to functional distinct cell types characterized by different phenotypes arisen from the same genotype [Probst AV, et al.; Nat Rev Mol Cell Biol, 2009].

1.2.1. Chromatin organization and regulation of gene expression

The epigenetic information relies on a particular structural organization named chromatin. Chromatin is the structure in which the DNA is packaged within the cell nuclei. The nucleosome is the basic unit of the chromatin and it is composed of an octamer of the four core histones (H3, H4, H2A, H2B, each one as dimer) around which 147 base pairs of DNA are wrapped [Kouzarides T; Cell, 2007]. There are 14 contact points between histones and DNA. These multiple interactions make the nucleosome one of the most stable protein-DNA complex under physiological conditions. For this reason, the histone-DNA complex is evolutionarily conserved and adapted to package, organize and regulate the DNA macromolecule (Figure 1.6). In fact the nucleosome is not a simple static packaging unit. It possesses dynamic properties which are strictly regulated by chemical modification mediated by several protein complexes orchestrating the chromatin remodeling and determining regulation of the gene expression [Li B, et al.; Cell, 2007].

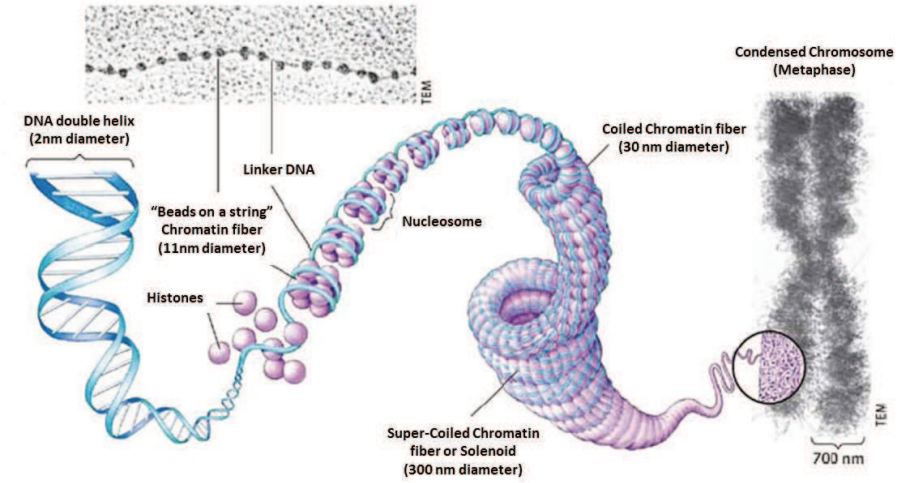


Figure 1.6. Representation of chromatin spatial organization. The image is adapted from online notes of Arianna Macrì Blogspot website (Mar 2013).

Transcriptional regulation resides in a complex and fine-tuned protein network that results in gene active transcription by the RNA polymerase II. Basically, the transcription initiation is controlled by at least three different architectural levels of regulation:

- DNA consensus sequences of transcription factors (common-sequence or sequence-specific) [Malik S, and Roeder RG; Trends Biochem. Sci., 2005];
- architectural complexity of the core promoter (regulatory elements, multiple start sites and alternative promoter) [Smale ST; Biochimica et Biophysica Acta; 1997];
- control of DNA accessibility by overlying epigenetic processes (DNA methylation, and chromatin structural organization like modifications of histone and nucleosome positioning) [Berger SL, et al.; Nature, 2007].

The explosion of epigenetic data in the recent years shows clearly that different structures of chromatin impose profound effects on all cell-life processes (i.e. transcription, recombination, DNA repair, replication, formation of centromere and kinetochore, etc.) [Li B, et al.; Cell; 2007]. Chromatin can be subdivided into two major status: heterochromatin, characterized by highly condensed state, late timing during replication, and the presence of inactive genes; and euchromatin, which is characterized, on the contrary, by open and accessible state and the presence of transcriptionally active genes (Figure 1.7). Different distribution of the genome from

the first to the second chromatin configuration leads to delineate the cell identity and the different cellular commitment during development [Reik W; Nature, 2007].

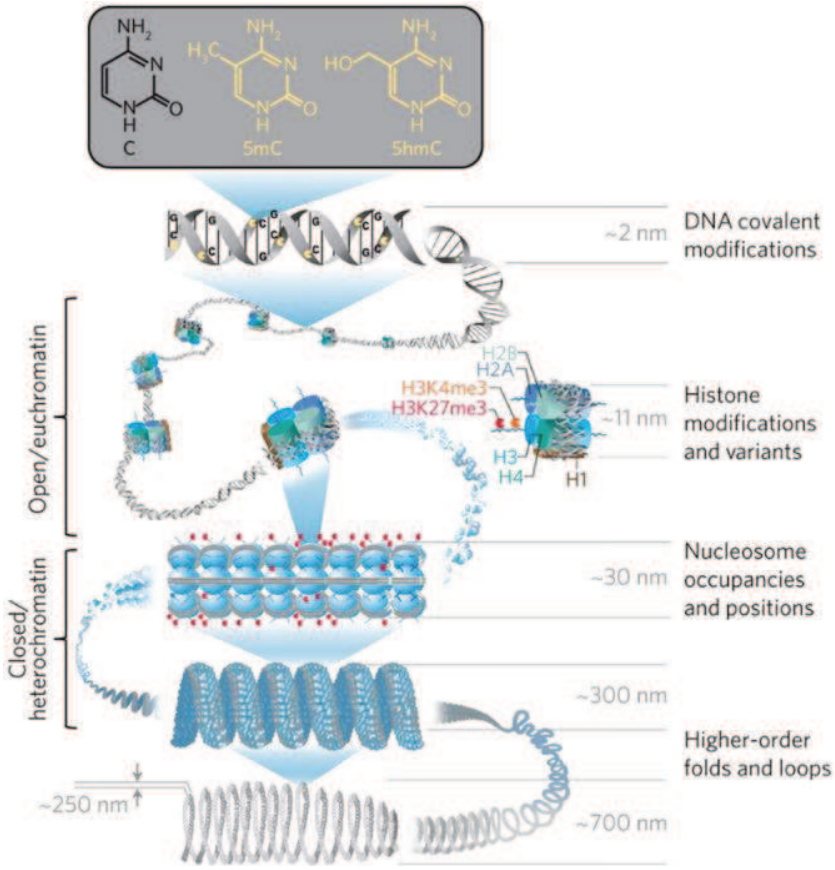


Figure 1.7. Representation of chromatin organization and principal modifications [Aguilar CA, and Craighead HG; Nature Biotech, 2013].

1.2.2. Chromatin modifications and functions

The dynamics of chromatin structure is regulated by several mechanisms including histone post translational modifications (hPTMs), presence of different histone variants, chromatin remodeling, and DNA methylation (different CpG methylation profile) [Li B, et al.; Cell; 2007]. All these dynamic mechanisms are orchestrated by a plethora of chromatin-modifying enzymes that cooperate to chromatin organization in a highly-regulated manner [Allis CD, et al.; Cell, 2007]. Chromatin-modifying

enzymes could be classified in 'writers', 'erasers' and 'readers' relying on their abilities respectively to add, remove or read different histone and DNA modifications [Treviño LS, et al., Prog Biophys. Mol. Biol., 2015].

Chromatin modifications, readers and their functions

Chromatin Modification	Nomenclature	Chromatin-Reader Motif	Attributed Function
DNA Modifications			
5-methylcytosine	5mC	MBD domain	transcription
5-hydroxymethylcytosine	5hmC	unknown	transcription
5-formylcytosine	5fC	unknown	unknown
5-carboxylcytosine	5caC	unknown	unknown
Histone Modifications			
Acetylation	K-ac	Bromodomain Tandem PHD fingers	Transcription, repair, replication and condensation
Methylation (lysine)	K-me1 K-me2 K-me3	Chromodomain, Tudor domain, MBT domain, PWWP domain, PHD fingers, WD40/β propeller	transcription and repair
Methylation (arginine)	R-me1 R-me2s R-me2a	Tudor domain	transcription
Phosphorylation (serine and threonine)	S-ph T-ph	14-3-3, BRCT	Transcription, repair and condensation
Phosphorylation (tyrosine)	Y-ph	SH2[a]	transcription and repair
Ubiquitylation	K-ub	UIM, IUIM	transcription and repair
Sumoylation	K-su	SIM[a]	transcription and repair
ADP ribosylation	E-ar	Macro domain, PBZ domain	transcription and repair
Deimination	R → Cit	unknown	transcription and decondensation
Proline isomerisation	P-cis↔P-trans	unknown	transcription
Crotonylation	K-cr	unknown	transcription
Propionylation	K-pr	unknown	unknown
Butyrylation	K-bu	unknown	unknown
Formylation	K-fo	unknown	unknown
Hyroxylation	Y-oh	unknown	unknown
O-GlcNAcylation (serine and threonine)	S-GlcNAc; T-GlcNAc	unknown	transcription

Modifications: me1, monomethylation; me2, dimethylation; me3, trimethylation; me2s, symmetrical dimethylation; me2a, asymmetrical methylation; and Cit, citrulline. Reader domains: MBD, methyl-CpG-binding domain; PHD, plant homeodomain; MBT, malignant brain tumor domain; PWWP, proline-tryptophan-tryptophan-proline domain; BRCT, BRCA1 C terminus domain; UIM, ubiquitin interaction motif; IUIM, inverted ubiquitin interaction motif; SIM, sumo interaction motif; and PBZ, poly ADP-ribose binding zinc finger. [a]These are established binding modules for the posttranslational modification; however, binding to modified histones has not been firmly established.

Figure 1.8. Table of chromatin modifications and their functions [Dawson M., and Kouzarides T.; Cell, 2012].

Nowadays, at least 4 different DNA modifications [Wu H, et al.; Genes Dev., 2011] and 16 classes of histone modifications [Kouzarides T; Cell, 2007] are known, even if the attribution a biological function remain unclear for some of them (Figure 1.8).

DNA Methylation

DNA methylation is the first epigenetic modification identified and has been intensively studied for half a century. In particular, among the DNA modifications, the most studied is the addition of methyl groups to the C5 position of cytosines (5mC) catalyzed by a family of DNA methyltransferases, named DNMTs.

DNA methylation occurs essentially in the context of CpG dinucleotides that tend to cluster in regions called CpG islands (CGIs). In general, CGIs, which constitute about 1-2% of the genome, are DNA stretches of more than 200 base pairs in length (from 0.2 to 3 kb), characterized by an elevated C/G content (at least 50 %) and a high ratio of CpG dinucleotide (at least 0.6) [Illingworth RS, and Bird AP; FEBS Lett., 2009]. About the half of CpG islands are locate in the proximity of transcription start sites (Deaton AM, and Bird AP; Genes & Dev., 2011] and many are associated with housekeeping genes and developmental regulators [Meissner A, Cell Stem Cell, 2011]. About 60% of human genes contains CpG islands in their promoter, which are usually unmethylated at all stages of development and in all tissue types (Antequera F, and Bird AP; PNAS, 1993), nevertheless some of them (around 6%) become methylated in a tissue-specific manner during early development or in differentiated tissues [Straussman R, et al. Nat Struct Mol Biol., 2009]. Microarray analyses showed that unmethylated regions seem to be established during early embryogenesis, mainly as a consequence of transcription factors recognition and localization on the specific sequence motifs closely associated with transcription start sites (TSS) [Straussman R, et al. Nat Struct Mol Biol., 2009]. In general, DNA methylation is mainly observed at centromeres, telomeres, inactive X-chromosomes, repetitive sequences and at some tissue-specific CpG island genes, as a stable gene-silencing mechanism, and it is associated to other repressive marks of chromatin [Baylin S, and Jones P; Nat Rev Cancer, 2011].

In the DNMTs family, three active methyltransferases have been identified in higher eukaryotes: DNMT1, DNMT3A and DNMT3B. DNMT1 is the principal methyltransferase of maintenance [Li E, et al. Cell, 1992] that, together with UHRF1, recognizes the hemi-methylated DNA, generated during DNA replication [Sharif J, et al.; Nature, 2007]. DNMT3A and DNMT3B, act primarily as de novo DNA

methyltransferases in establishment of DNA methylation during embryogenesis [Okano M, et al.; Cell, 1999]. Interestingly, it has been observed that DNMT3A/3B corroborate DNMT1 in DNA methylation maintenance [Jones PA and Liang G; Nature Rev. Genet., 2009]. In fact, DNMT3A/3B can anchor nucleosomes containing methylated CpGs, in particular nucleosomes containing methylated SINE or LINE elements and CpG islands [Jeong S, et al.; Mol Cell Biol., 2009]. The reason of these observations could reside in the fact that DNMT3A/3B, conversely to DNMT1, does not require the presence of interacting-chromatin-actors (e.g. HP1, MeCp2, EZH2, UHRF1, HDAC1, PCNA), or particular histone PTMs profiles, to participate in DNA methylation maintenance, after the passage of the replication fork [Jones PA and Liang G; Nature Rev. Genet., 2009].

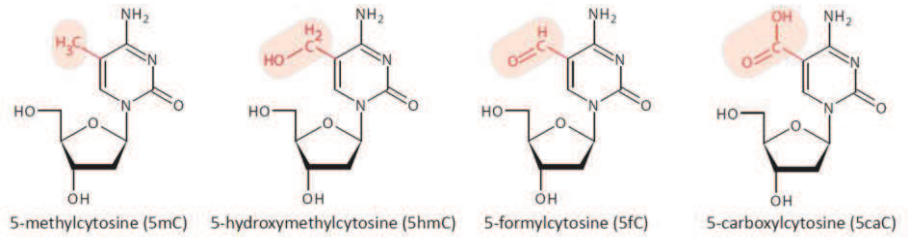


Figure 1.9. Overview of DNA modifications. The molecules were drawn with MarvinSketch software.

Notably, m5C can be actively oxidized to 5-hydroxy-methylcytosine (5hmC) by TET family proteins and subsequent deamination and replacement via BER pathway (Guo et al., 2011). Although the role of these enzymes in the global DNA de-methylation processes are still unclear [von Meyenn F, et al.; Mol. Cell, 2016].

Histone modifications

In the last 20 years, a large variety of histone modifications has been found. As expected these hPTMs are tightly regulated by enzymes that mediate acetylation [Sterner and Berger, 2000], methylation [Zhang and Reinberg, 2001], phosphorylation [Nowak and Corces, 2004], ubiquitination [Shilatifard, 2006], SUMOylation [Nathan et al., 2003], ADP-ribosylation [Hassa et al., 2006], deamination [Cuthbert et al., 2004], proline isomerization [Nelson et al., 2006] and crotonylation [Tan et al., 2011]. In general, post-translational histone modifications can affect chromatin structure by a direct and an indirect mechanism. In fact they can directly alters the condensation of chromatin, by changing the electric charges.

For example, the acetylation of lysine residues has the highest potential to de-condense chromatin since it neutralizes the basic charge of histones and the phosphorylation of Serine 10 on H3 is crucial for chromosome condensation and cell-cycle progression during mitosis and meiosis [Ahn et al., 2005; Fischle et al., 2005; Krishnamoorthy et al., 2006]. By indirect mechanisms the composition of histone modifications mediates the binding of important protein complexes to the chromatin. For example, HP1 (heterochromatin-associated protein 1) binds specifically the H3K9me3 [Jacobs and Khorasanizadeh, 2002; Jacobs et al., 2001] and is released upon phosphorylation of the adjacent serine 10 during M phase of the cell cycle [Fischle et al., 2005]. The most studied histone modifications are lysine and arginine methylation, lysine acetylation, serine and threonine phosphorylation and lysine ubiquitination. Generally, histone methylations lead to a more condensed chromatin, while histone acetylation induces electrostatic repulsions that guide chromatin opening [Chen et al., 2014]. Pattern of specific histone modifications define heterochromatin and euchromatin domains. In particular, acetylation of histone 3 and histone 4 (H3 and H4) or di- or trimethylation (me2 or me3) of H3K4, are commonly related to euchromatic regions, while H3K9 and H3K27 methylation, are often associated to heterochromatin. In particular, H3K9me3 is associated to constitutive heterochromatin, while H3K27me3 is associated to facultative heterochromatin [Li B, et al., 2007]. Inside these two "status" of the chromatin, the hPTMS are not casually distributed. In fact, its localization is finely regulated and generally they localized in specific patterns in promoters, upstream regions, the 5' end of the open reading frame (ORF) and the 3' end of the ORF.

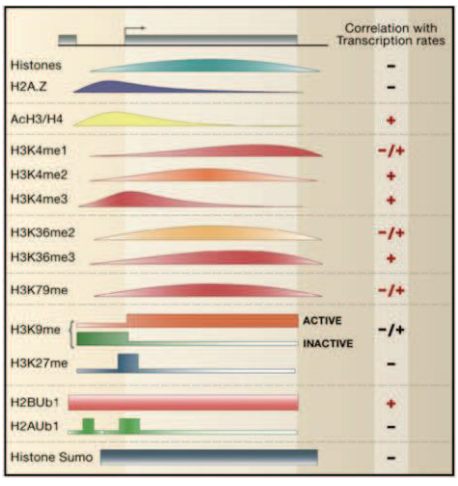


Figure 1.10. Schematic representation of the distribution at genomic level of the histone modifications in relation to transcriptional activity [Li B, et al., 2007].

The information conveyed by epigenetic modifications plays a critical role in the regulation of all DNA-based processes, such as transcription, DNA repair, and replication. Consequently, abnormal expression patterns or genomic alterations in chromatin regulators can have profound results and can lead to the induction and maintenance of various cancers [Dawson et al., 2012].

Crosstalk between Epigenetic Modifications

The two major epigenetic features, the DNA methylation and the histone modifications, are deeply linked. In particular, there are three proteins family that specifically recognize methylated CpG sites (mCpG): the methyl-CpG binding domain (MBD), the Kaiso and the UHRF protein family (Rottach et al., 2009). The methyl binding proteins can recognize specific DNA methylation status and recruit protein complexes containing HDACs and histone methyltransferases. [Nan et al., 1998; Rottach et al., 2009; Hendrich and Tweedie, 2003; Bird, 2002]. The different histone modification patterns, the DNA methylation status and higher ordered structures act together mediating their interaction, interference, and/or an inter-functional dependency (cross-talk). Due the redundancy and the complexity of this epigenetic code and the multiplicity of the protein complexes involved, most of the details of these interactions are still under investigation [Kouzarides T; Cell; 2007].

However, it is possible to report the following considerations (Figure 1.11):

- the presence of different types of modifications on the same lysine residues undoubtedly results in some form of antagonism between these different modifications;
- the binding of a protein to a specific site can be impaired by an adjacent modification, e.g. the phosphorylation of serine H3S10 influences the binding of HP1 protein to methylated H3K9 [Fischle et al.; Nature; 2005];
- the catalytic activity of an enzyme could be compromised by the modification of its recognition site, for example, the isomerization of the proline H3P38 influences the ability of Set2 to methylate H3K36 [Nelson, C.J., el al.; Cell; 2006];
- an enzyme, in contrast, may be able to more effectively recognize its substrate, if there is another specific modification, e.g. the GCN5-acetyltransferase can recognize H3 more effectively when is present the phosphorylation of serine H3S10 [Clements, A., et al.; Mol. Cell; 2003];
- cross-talk between the various modifications can occur even when the markers are on the histone tails of different histones, but belonging to the same nucleosome, e.g. the ubiquitination of lysine in position 123 of histone H2B, is critical for the tri-methylation of H3K4;

- some histone modifications could interfere with the maintenance of DNA methylation, e.g. H3K18ub, ubiquitinated by UHRF1, is essential for the maintenance of DNA methylation, enabling methylation activity of DNMT1, through interaction with its ubiquitin interacting motif (UIM) [Qin W., et al; Cell Research, 2015].

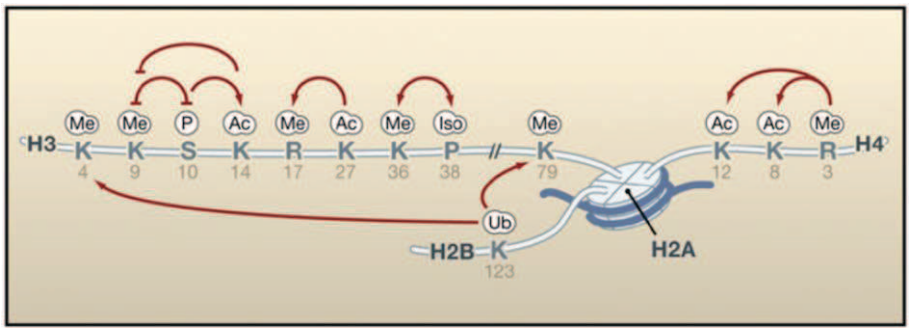


Figure 1.11. Crosstalk between histone modifications [Kouzarides T; Cell; 2007].

Epigenetic modifications in cancer

For a long period, alterations in the genetic code was considered the major cause of cancer development. As long as the knowledge about the epigenetic mechanisms increased, it became clear that, not only there were a deep rearrangement of the epigenome during cancer progression, but that the modulation of epigenetic modifications could affect cancer outcome. As for the genetic modifications, the epigenetic alterations can be cancer type-specific (Figure 1.12), and they can regard many aspect, such as DNA methylation, histone modification, nucleosome positioning and micro-RNAs expression [Esteller, 2010].

The first epigenetic traits that has been found altered in cancer cells is the DNA methylation [Feinberg, 1983; Goelz, 1985]. Cancer cells are characterized by a specific epigenome: the CpGs interspersed in the genome, usually highly methylated in normal tissues, are heavily hypo-methylated, while the CpGs clustered in the CGIs, typically located in gene promoters and usually unmethylated, are hyper-methylated. Form a functional point of view, DNA hypomethylation correlates with chromosome instability, activation of transposable elements and loss of imprinting. In fact, the DNA hypomethylation at the DNA satellite loci or in the pericentromeric associates with increased chromosomal rearrangements, mitotic recombination, and aneuploidy [Eden, 2003; Karpf, 2005]. Repetitive elements, such as L1 (long

interspersed nuclear elements) and Alu (recombinogenic sequence), are silenced in somatic cells and become reactivated in human cancer [Schulz, 2006]. In gene promoters, high DNA methylation levels affect the gene expression of specific pathways, such us, cell cycle, DNA repair, apoptosis and many others [Esteller, 2008]. Although the ultimate causes of aberrant DNA methylation remain to be determined, several studies showed that alterations in the DNA methylome could be directly affected by diet, xenobiotic chemicals, and exogenous stimuli, such as inflammation or viral/bacterial infection.

Table 1. Epigenetic Aberrations among Different Tumor Types.*

Type of Cancer	Epigenetic Disruption
Colon cancer	CpG-island hypermethylation (*hMLH1*, *p16^{INK4a}*, *p14^{ARF}*, *RARB2*, *SFRP1*, and *WRN*), hypermethylation of miRNAs (*miR-124a*), global genomic hypomethylation, loss of imprinting of *IGF2*, mutations of histone modifiers (*EP300* and *HDAC2*), diminished monoacetylated and trimethylated forms of histone H4
Breast cancer	CpG-island hypermethylation (*BRCA1*, E-cadherin, *TMS1*, and estrogen receptor), global genomic hypomethylation
Lung cancer	CpG-island hypermethylation (*p16^{INK4a}*, *DAPK*, and *RASSF1A*), global genomic hypomethylation, genomic deletions of *CBP* and the chromatin-remodeling factor *BRG1*
Glioma	CpG-island hypermethylation (DNA-repair enzyme *MGMT*, *EMP3*, and *THBS1*)
Leukemia	CpG-island hypermethylation (*p15^{INK4b}*, *EXT1*, and *ID4*), translocations of histone modifiers (*CBP*, *MOZ*, *MORF*, *MLL1*, *MLL3*, and *NSD1*)
Lymphoma	CpG-island hypermethylation (*p16^{INK4a}*, *p73*, and DNA-repair enzyme *MGMT*), diminished monoacetylated and trimethylated forms of histone H4
Bladder cancer	CpG-island hypermethylation (*p16^{INK4a}* and *TPEF/HPP1*), hypermethylation of miRNAs (*miR-127*), global genomic hypomethylation
Kidney cancer	CpG-island hypermethylation (*VHL*), loss of imprinting of *IGF2*, global genomic hypomethylation
Prostate cancer	CpG-island hypermethylation (*GSTP1*), gene amplification of polycomb histone methyltransferase *EZH2*, aberrant modification pattern of histones H3 and H4
Esophageal cancer	CpG-island hypermethylation (*p16^{INK4b}* and *p14^{ARF}*), gene amplification of histone demethylase *JMJD2C/GASC1*
Stomach cancer	CpG-island hypermethylation (*hMLH1* and *p14^{ARF}*)
Liver cancer	CpG-island hypermethylation (*SOCS1* and *GSTP1*), global genomic hypomethylation
Ovarian cancer	CpG-island hypermethylation (*BRCA1*)

Figure 1.12. Table of the major epigenetic aberrations reported in different tumors [Esteller M; N Engl J Med, 2008].

The changes in the histone modifications are more complex to track, due to the complexity of all the possible permutations and combinations (Figure 1.13). In colon cancer, it has been shown that the hypermethylation of CGi in the promoters of tumor-suppressor genes is associated with a particular combination of histone markers: deacetylation of histones H3 and H4, loss of H3K4 tri-methylation, and gain of H3K9 methylation and H3K27 tri-methylation [Jones, 2007].

Nevertheless, as DNA methylation and histone modifications act as mechanisms for controlling cellular differentiation, allowing the expression only of tissue-specific and

housekeeping genes in somatic differentiated cells, it is possible that the inappropriate (re)activation of tissue-specific genes can also plays a role in cancer [Fan et al., 2008].

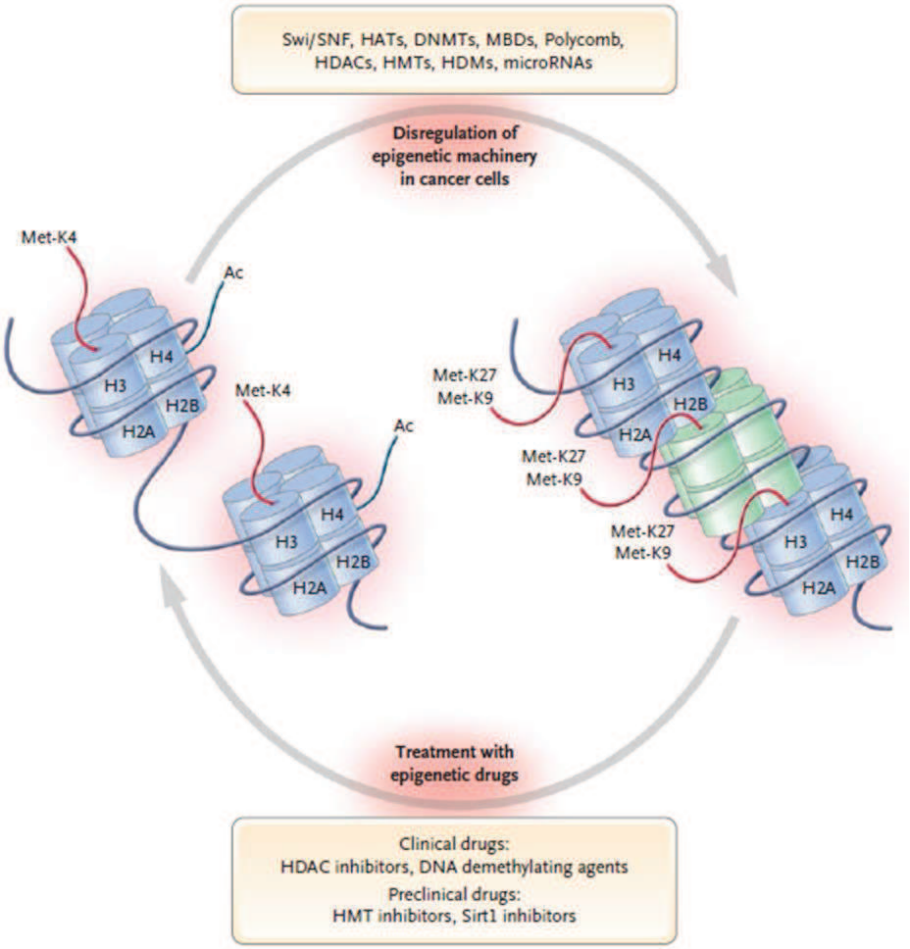


Figure 1.13. Aberrant epigenetic inactivation of Tumor-Suppressor Genes [Esteller M; N Engl J Med, 2008].

1.3. UHRF1

The UHRF1 gene (Ubiquitin-like with PHD and Ring Finger domains 1), is located in chromosome region 19p13.3 and encodes the UHRF1 protein, also known as NP95 or lCBP90. UHRF1 is a modular multi-domain protein able to mediate interactions between DNA methylation and histone modifications, by interacting directly with DNA and histones [Hideharu Hashimoto, et al.; Epigenetics; 2009].

UHRF1 gene, identified in 1998 [Fujimori et al., Mamm Genome, 1998], encodes for mainly two isoforms, consisting in 793 (isoform 1) and 806 (isoform 2) amino acid residues, with a molecular weight of about 95kDa. At the beginning, it has been characterized as a transcription factor that binds to the inverted CCAAT box of the topoisomerase 2a promoter and regulates its expression [Hopfner R, et al.; Cancer Res., 2000]. Over the last 15 years, several studies demonstrated that UHRF1 has the unique property to act as a "reader" for both DNA methylation and specific histone-modification patterns, to act as a "recruiter" for histone and DNA modifiers [Bronner C, et al.; Biochem. Pharmac., 2013] (Figure 1.14a), and more recently, to act as "ubiquitin-writer" with an E3-ligase activity [Citterio E, et al.; Mol. Cell. Biology, 2004] [Jenkins Y, et al; Mol. Biol. of Cell, 2005] toward H3K23 [Nishiyama A, et al.; Nature, 2013] and H3K18 [Qin W., et al; Cell Research, 2015].

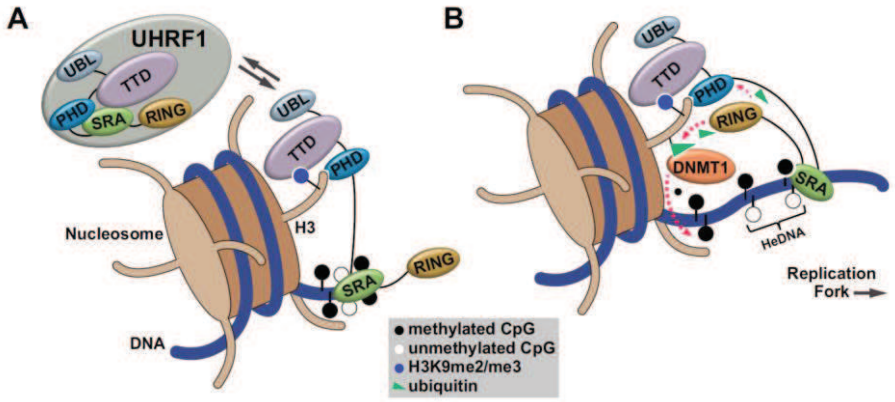


Figure 1.14. Current proposed model of multi-modular interaction of UHRF1 with both DNA and histone modifications. (A) UHRF1 recruitment on chromatin through the combined recognition of histones, by TTD-PHD domains of H3K9me2/me3 and H3R2un, and recognition of DNA, by SRA domain independently of the methylation status. (B) DNA methylation regulatory function of UHRF1, by direct enhancement of its ubiquitin ligase activity, towards N-terminal lysines of H3 tail by RING domain, upon the interaction of the SRA with HeDNA, reinforcing the stability and activity of DNMT1 in chromatin [Harrison JS, et al.; eLife, 2016].

Currently, the E3-ligase activity of UHRF1 it is acquiring greater prominence: in fact, it has been shown that UHRF1 ubiquitylates multiple lysines on the H3 tail, nearby the UHRF1 histone-binding site, with enhanced activity upon binding to hemi-methylated DNA stretches (HeDNA) (Figure 1.14b). This result provides a functional role for the ubiquitination pattern, mediated by UHRF1, as a nucleation event for DNMT1 recruitment to chromatin, in the dynamic context of replicative fork [Harrison JS, et al.; eLife, 2016].

1.3.1 Multi-modular domains and molecular functions of UHRF1

UHRF1 is a multi-modular domain protein that contains five functional domains, characterized by several molecular abilities (Figure 1.15):

- Ubiquitin-like domain (UBL), present at the N terminal, is still not elucidated but is involved, bona fide, in the proteasome degradation pathway, relying on the fact that many UBL domain-containing proteins interact with the 26S proteasome [Bronner C, et al.; Biochem. Pharmac., 2013];
- Tandem Tudor domain (TTD), recognizes specific histone marks, such as mono-, di-, and tri-methylation of lysine 9 of histone H3 [Arita K, et al.; PNAS, 2012] [Rothbart SB, et al.; Genes & Dev., 2013], in particular the TTD is able to bind stably H3K9me3, supported by PHD domain in enhanced interaction with H3 tail [Cheng J, et al.; J Biol Chem, 2013], [Xie S, et al.; J Biol Chem, 2012], [Rottach A, et al .; Nucleic Acids Res, 2010].
- Plant Homeo-Domain (PHD), it is reported that acts as a histone H3 tail-binding module recognizing the first four residues of H3 N-terminal tail, in particular H3R2 unmodified [Hu L, et al.; Cell Res., 2011] [Rajakumara E, et al.; Mol. Cell, 2011];
- SET and RING Associated (SRA) domain, specifically recognizes and binds hemi-methylated CpG dinucleotides, recruiting DNMT1 to hemi-methylated DNA to preserve the correct DNA methylation profile during DNA replication [Bostick M, et al.; Science, 2007] [Sharif J, et al.; Nature, 2007]; in particular, the SRA domain, present only in the UHRF family proteins, has the unique ability to flip the methylated cytosine out from the DNA duplex allowing the correct methylation of the un-methylated nucleotide by DNMT1 [Arita K, et al.; Nature, 2008] [Avvakumov GV, et al.; Nature, 2008];

this is the first example of not-enzymatic sequence-specific DNA-protein interaction [Hashimoto H, et al .; Nature, 2008];

- RING domain, at the C terminus, is responsible for the E3-ubiquitin-ligase activity of UHRF1; it has been shown that UHRF1 is able to ubiquitinate several histone substrates in vitro with preference for histone H3, but also histone H1 and H2B [Citterio E, et al.; Mol. Cell. Biology, 2004] [Jenkins Y, et al; Mol. Biol. of Cell, 2005]; in vivo it has been demonstrated that UHRF1 mono-ubiquitinates H3K23 [Nishiyama A, et al.; Nature, 2013] and H3K18 [Qin W., et al; Cell Research, 2015]; furthermore, are reported abilities to poly-ubiquitinated other substrates beyond histone, such as DNMT1 [Qin W, et al.; J. Cell. Biochem., 2011], recently confirmed together with the auto-ubiquitination of UHRF1 itself [Harrison JS, et al.; eLife, 2016], and RIF1, part of Double Strand Break (DSBs) repair complexes (ubiquitination that mediates Homologous Repair by promotion of BRCA1 pathway) [Zhang H. et al.; Nature Communications, 2015].

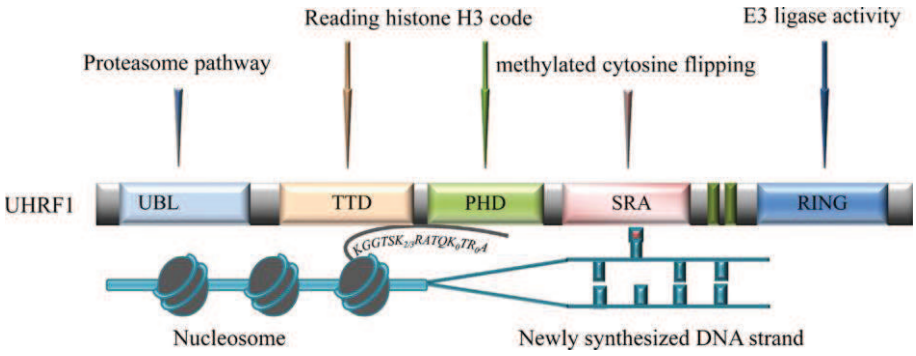


Figure 1.15. Representation of UHRF1 domains and their functional roles [Bronner C, et al.; Biochem. Pharmac., 2013].

It is clear that UHRF1, through its multi-domain abilities, is one of the main ‘‘cross-talk protein’’ involved in the maintenance and the regulation of the epigenetic information. In fact, by its SRA domain, beyond the well-reported interaction with DNMT1 and HeDNA, UHRF1 can also mediate the recruitment of histone deacetylase 1 (HDAC1) [Unoki M, et al.; Oncogene, 2004], one of the major player in chromatin condensation and gene silencing. Further studies have shown that SRA is also able to interact with the regulatory domains of the methyl-transferase de novo (DNMT3A and DNMT3B) and with the histone methyl-transferase (HMT) G9a [Meilinger D, et al.; EMBO reports, 2009]. These informations together with the knowledge about the role of the PHD and the TTD domains, leads to propose a model of the

propagation of H3K9me3 during DNA replication. UHRF1 binds to methylated H3K9 already present on the "old" histones in front of the replicative fork, then mediate the deposition of methylated H3K9 on the "new" incorporated histones, by its recruitment of the HMTs G9a and Suv-39H1 [Xie S, et al.; J. Biol. Chem., 2012]. Since UHRF1 is associated to the heterochromatin formation (e.g. in the inactivation of X-chromosome, pericentromeric regions, and/or subtelomeric) [Karagianni P, et al.; Mol. Cell. Biol., 2008] [Dillon N.; Cell Biol.; 2004], it reasonable that UHRF1-associated repressive complex (e.g. DNMT1, HDAC1, HMTs) mediates the propagation of DNA methylation and H3K9me3 in a mutually reinforced mechanism, ensuring the stability of heterochromatic states [Xie S, et al.; J. Biol. Chem., 2012].
UHFR1 knockout mice are embryonic lethal. Conversely, in UHRF1-null Embryonic Stem cells (ESC), complemented with UHRF1 mutants in SRA or TTD domains, UHRF1 is able to associate with pericentric heterochromatin, recruits DNMT1 and partially rescues DNA methylation defects [Liu X, et al.; Nat. Commun., 2013]. Interestingly, these observations were partially confirmed in our recent publication [Qin W., et al; Cell Research, 2015], showing the novel intriguing role of the RING domain and suggesting a more crucial role of PHD domain in the recruitment of UHRF1 to chromatin. In fact, RING domain ubiquitinates H3K18ub, and this histone modification is essential for the maintenance of DNA methylation, enabling methylation activity of DNMT1, through interaction with its ubiquitin interacting motif (UIM) [Qin W., et al; Cell Research, 2015]. Furthermore, recently it has been shown that ubiquitination pattern, mediated by UHRF1 on several lysines of H3 (K14, K18, K23, K27, K36), is an essential event to recruit strongly DNMT1 to hemi-methylated DNA stretches (HeDNA), in the dynamic context of replicative fork [Harrison JS, et al.; eLife, 2016].
Altogether these findings suggest that UHRF1 plays an important role as epigenetic regulator, acting in a coordinated fashion for faithful inheritance of the epigenetic code.

1.3.2 Role of UHRF1 in cancer pathogenesis

UHRF1 is involved in a large number of physiological and pathological phenomena, from embryogenesis to cell migration and cancer development and progression. In normal cells, UHRF1 is expressed exclusively in dividing cells and completely absent in terminally differentiated cells. In dividing cells, its expression is tightly regulated:

it starts to be expressed in S phase, it reaches the peak in late G1 and it is rapidly degraded in M phase upon its phosphorylation [Mousli, 2003]. Its inhibition impairs cell cycle progression and induces cell cycle arrest and apoptosis [Bonapace, 2002].
Since the pivotal role of UHRF1 in the maintenance of the epigenetic code during cell cycle, it is not surprising that UHRF1 is often deregulated in cancer cells. It is overexpressed in most of the solid cancer such as breast, lung, bladder, pancreatic, colon, prostate and cervical cancer (Bronner 2007; Unoki 2009; Unoki 2004; Unoki 2010; Jin 2010; Lorenzato 2005; Crnogorac-Jurcevic 2005; Sabatino 2012; Babbio 2012). In cancer cells, UHRF1 is not cell cycle regulated anymore and it is overexpressed throughout the all the phases of the cell cycle.
Recently it has been shown that the overexpression of UHRF1 in developing hepatocytes is sufficient to induce hepatocellular carcinoma, including UHRF1 in the list of the oncogenes [Mudbhary et al., 2014]. At molecular level UHRF1 overexpression directly induces the hypermethylation of specific tumor suppressor genes, such as BRCA1, p16INK4A, p73, peroxisome proliferator activated receptor gamma (PPARγ) and p21 [Kim 2009; Hopfner 2002; Bronner 2007; Sabatino 2012; Alhosin 2012; Daskalos 2011; Guan 2013; Jeanblanc 2005] and a genome-wide demethylation by inducing DNMT1 degradation [Mudbhari et al., 2014].
In colon cells, we reported that UHRF1 overexpression was sufficient to epigenetically silence a large panel of tumor suppressor genes, including PPARγ and CDH1 and its overexpression stimulates the migration and the invasion of these tumor cells. For cervical and adenocarcinoma of pancreas, UHRF1 has been proposed as a biomarker for diagnosis and bad prognosis [Lorenzato et al., 2005; Crnogorac-Jurcevic 2005]. Recently, it was demonstrated that the disruption between DNMT1/PCNA/UHRF1 acts as an oncogenic event and one of its signatures (i.e., the low level of methyl-transferase activity) is a molecular biomarker associated with prognosis in patients with glioblastoma [Hervouet 2010]. In non-muscle-invasive bladder cancer, UHRF1 overexpression is an independent prognostic factor for tumor recurrence [Yang 2012].
These observations are not surprising since UHRF1 is a major effector in the reading and maintenance of the epigenetic code, which itself is profoundly altered in cancer.

2. Aim of the Thesis

Colorectal cancer (CRC) is the third most common cancer, and one of the major leading cause of cancer death (fourth in men and third in women) [Ferlay J., et al.; Int. J Cancer., 2015]. CRCs, as well as other tumor types, are a complex multifactorial disease. In the last decades, an incredible amount of data demonstrates the importance of epigenetic processes at all stages of cancer development, leading to the realization that genetics and epigenetics cooperate to alter the biological processes that are fundamental to the genesis and progression of cancer [Dawson MA, et al.; Cell, 2012].

In Colon cancer, genetic and epigenetic alterations are tightly connected and together they cooperate to determine the tumor aggressiveness and patient's outcome. In fact, CRCs can be classified based on some epigenetic and molecular features, like microsatellite stability (MSS) or instability (MSI) and hypermethylation of CpG islands (CIMP), however genetic mutation (e.g. BRAF and KRAS) can deeply influence the prognosis. Their interaction is complex: for example, MSI-high tumors carry a good prognosis whereas the presence of a BRAF mutation confers a poor outcome. However, CIMP-high, that is tightly associated with MSI tumors and BRAF mutation, appears to eliminate the adverse effect of BRAF mutation, carrying better prognostic and predictive qualities [Ogino S., et al.; Gut. 2009]. Nevertheless, the effect of the interaction between genetic mutations and epigenetic changes on the patient's outcome is not completely understood [Wu C and Bekaii-Saab T; Chemotherapy Research and Practice, 2012].

UHRF1 is a key master epigenetic regulator that couples the preservation of histone-modifications through the cell cycle with maintenance of DNA methylation. UHRF1 is over-expressed in numerous tumor types. In colon cancer cells, it recruits Suv39H1 and DNMTs on specific gene promoters, such as PPARγ, E-cadherin and RARβ, mediating their epigenetic silencing [Sabatino L, Fucci A, et al.; Oncogene, 2012]. We recently published that UHRF1 ubiquitinates K18 of histone H3 and this modification is essential for the maintenance of DNA methylation, enabling methylation activity of DNMT1, through interaction with its ubiquitin interacting motif (UIM) [Qin W., et al; Cell Research, 2015].

Relying on all these observations, this PhD project arises from the idea that UHRF1 could play a crucial role in the modulation of DNA methylation changes occurring during CRC progression, cooperating with other molecular factors and influencing tumor's outcome. The overall aim of this PhD thesis was to evaluate the role of

UHRF1 in the regulation of DNA methylation both at genome-wide and at locus specific level in colon cancer, and investigate the intriguing activity of UHRF1 as histone post-translational modifications (hPTMs) "writer".

In the attempt to better understand the role of UHRF1 in colon cancer, we took under exam a large cohort of CRC samples (MSI n =47, and MSS n =88) and controls (adenoma n =25). This analysis helped us to define the importance of UHRF1 in the CRC outcome. Therefore, to deepen the molecular mechanism that underlies the clinical observations, we decided to analyze two cell lines: RKO cells, a cell line derived from MSI CRC, and HT29, derived from MSS CRC. Afterwards, we chose RKO cells (UHRF1-high, MSI CRC model, CIMP-high) in order to modulate the level of UHRF1, and evaluate the effect on DNA methylation patterns, genes expression and histone PTMs profile (in particular the changes of ubiquitination pattern in the histone tails).

According to this approach, the present work was focused to enlighten the following aspects:

- Evaluation of UHRF1 and DNA methylation levels in human specimens;
- Analysis of UHRF1 and DNA methylation levels in colorectal-cancer-derived cell lines (RKO & HT29);
- Evaluation of the effect of UHRF1 knock down on DNA methylation, gene expression and histone modifications in RKO cells;
- Genome wide analysis of the effect of UHRF1 silencing on the methylome and the transcriptome;
- Investigate the effect of UHRF1 silencing on ubiquitination of histone tails.

3. Materials and Methods

3.1. Patients and samples

Patients samples derived form a collection of 135 Formalin-Fixed Paraffin-Embedded (FFPE) CRC samples, including 83 tumors previously examined for microsatellite instability at high frequency (MSI) and for gene-specific methylation (Furlan D, 2013). The newly added cases comprise 28 MSI and 22 microsatellite stable (MSS) CRCs which were collected from Ospedale di Circolo, Varese and from Fatebenefratelli Hospital in Benevento (Italy). Twenty-five samples of histologically normal colorectal mucosa were also analyzed in this study. Of these, 12 samples derived from patients without evidence of neoplastic malignancies and 13 from morphologically normal mucosa adjacent to a CRC. All tumors were histologically reviewed according to the WHO classification of tumors of the digestive system (Hamilton SR, 2010) and the TNM staging system (Edge SB, 2010) by a histopathology physician (dott.sa Chiaravalli in Ospedale di Circolo, varese). Outcome data were collected by consulting clinical records and/or the Tumor Registry of the Lombardy region (Italy) and were available for all patients. Ninety-six patients (71%) died of disease while fifty-eight patients (29%) were alive (median follow-up time of 24 months and 69.2 months, respectively). The study was approved by the Ethics Committee of the Ospedale di Circolo, Varese, Italy.

3.2. Immunohistochemical analysis

FFPE 3-µm sections were dewaxed, rehydrated, and dipped in 3% hydrogen peroxide in water solution for 10 min. After antigen retrieval by pressure cooking in 1 mmol EDTA buffer (pH 8) for 30 seconds, sections were incubated overnight with 1:4000 anti UHRF1 monoclonal antibody. A peroxidase polymeric amplification system (HRP-Ultravision LP, LabVision, Suffolk, UK) was subsequently applied according to the manufacturer's specification and the reaction product was visualized with 3-3’ diaminobenzidine (DAB). After washing in water, sections were counterstained with hematoxylin, dehydrated, and mounted permanently. All the steps were performed manually.

The immunostaining was considered positive only when nuclear staining was present. Positivity to UHRF1 staining was determined by assessing the percentage

of positive cells. Percentage of positive cells was calculated both for the CRC samples and the control tissues.

3.3. Cell Cultures

Human colon carcinoma cell lines (RKO and HT29), human prostate carcinoma cell line (PC3), 293T and HeLa cells were purchased from the American Type Culture Collection (Manassas, VA, USA) and cultured as monolayer at 37°C and 5% CO_2.

293T and PC3 cells were grown in RPMI 1640 medium (Euroclone), supplemented with 10% of heat-inactivated foetal bovine serum (Euroclone), 1% L-glutamine (Sigma-Aldrich) and 1% Penicillin/Streptomycin (Sigma-Aldrich).

RKO, HT29 and HeLa cells were grown in DMEM medium (Euroclone), supplemented with 10% of heat-inactivated foetal bovine serum (Euroclone), 1% L-glutamine (Sigma-Aldrich) and 1% Penicillin/Streptomycin (Sigma-Aldrich).

3.4. siRNAs and transfection

siRNA against UHRF1 was purchased form Invitrogen. To the setup, two different siRNAs were used and, both of them under the same experimental conditions, gave similar results.

Transfections have been performed by using either Interferin (Polyplus) or RNAiMax (Life Technologies). Briefly 50000 cells were plated in a 6 well-plate. 24 hours after plating, the transfection has been performed as suggested in the manufacturers' instructions (two different protocols for the two reagents) by using 10nM final concentration of the siRNA. For the setup, the efficiency of the silencing has been monitored at 24, 48 and 72 hours after transfection.

By using the RNAiMax reagents and 10nM of siRNA, the 60% of the UHRF1 protein was depleted after 48 hours and at 72 hours, more that 95% of UHRF1 protein was silenced. These results were reproducible among the different biological replicates. Therefore, we decided to follow this protocol (10nM of siRNA for 72 hours) for all the experiments.

3.5. RNA extraction and Retrotranscription

Sub confluent (for the basal characterization) or 3 days transfected cells were washed by using 1X PBS (Sigma Aldrich). RNA was extracted using Trizol reagent (Life Technologies) following the manufacturer's instructions. RNA was quantified by spectrophotometer or Q-bit (Life Technologies).
0.5 to 5 µg of RNA was treated with 4U of TURBO DNAseI (Ambion) for 30 minutes at 37°C and then RNA was extracted by standard phenol:chloroform procedure.

Extracted RNA was used either for RNA-seq analysis or retrotranscribed for PCR experiments. or 0.5 to 1 µg of RNA was retrotranscribed using SSIII Superscript (Life technologies) with oligodT. cDNA was diluted at 5ng/µl and used for the PCR analysis.

3.6. DNA extraction

DNA from normal and tumor FFPE samples were obtained from manually microdissected tissues using a QIAamp® DNA FFPE Tissue kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol.
DNA from cells was extracted by using PureLink® Genomic DNA columns (Invitrogen).
DNA was quantified by Qbit and used for MS-MLPA, pyrosequencing or RRB-sequncing.

3.7. Methylation-Specific Multiple Ligation-dependent Probe Amplification (MS-MLPA) analysis

Promoter methylation of a total of 34 genes using (MS-MLPA) with SALSA MS-MLPA ME001 Tumor suppressor-1 Kit and SALSA MS-MLPA ME002 Tumor suppressor-2 Kit (MRC-Holland, Amsterdam, The Netherlands) were already available for 83 CRCs (Furlan D, 2013). We extended this analysis to the newly added CRCs as well as to CRC cell lines and the 25 samples of the normal colorectal mucosa. Transfected RKO and HT29 cells were further tested for the presence of promoter methylation of the Mismatch Repair genes by using SALSA MS-MLPA KIT ME011 Mismatch Repair genes (MRC-Holland, Amsterdam, The Netherlands).

3.8. Bisulphite pyrosequencing analysis

Bisulphite conversion was performed by using EpiTect Plus DNA Bisulfite Kit (Qiagen, Hilden, Germany). Converted DNA was analyzed by pyrosequencing in collaboration with dr. Daniela Furlan at the Ospedale di Circolo di Varese.

3.9. Polymerase Chain Reaction (PCR)

Primers were designed by using the open-access tool Primer 3 software (http://fordo.wi.mit.edu/website) choosing amplicons of approximately 75-135bp (see primers table). The selected sequences were validated in silico by using primerBLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) to confirm the specificity of the target. The melting temperature was experimentally determined for each primer pairs.

3.9.1. Semi-quantitative PCR

The semi-quantitative PCR was performed using the HotStart GoTaq polymerase (Promega Inc.) following the manufacturer's instructions. The amplicons were separated on 1.5% agarose gel and visualized by ethidium bromide staining (Sigma Aldrich) at luminometer.

3.9.2. Quantitative real-time PCR

Quantitative real-time PCR was performed in iCycler iQ (BioRad) by using 2X iQ™SYBR Green Supermix (BioRad). Data were analysed by averaging triplicates Ct. Levels of RNA expression were determined by Gene Expression Analysis for iCycler iQ Real-Time PCR Detection System v1.10 (Bio-Rad) according to the 2-ΔΔCq method. Levels of RNA expression of selected genes were normalized to the internal control reference gene (GAPDH).
Post-PCR melting curves were used to assess the quality of primer pairs.

3.10. Western Blot

Sub confluent (for the basal characterization) or 3 days transfected cells were washed using 1X PBS (Sigma Aldrich), and directly lysed by using 2X Leammli (Life Technology). Lysates were sonicated and boiled 5 minutes at 95°C; proteins were separated on a gradient gel (4–20% Mini-PROTEAN® TGX™ Precast Protein Gels, BioRad) under reducing conditions and transferred onto polyvinylidene fluoride membranes (Amersham, Biosciences, Otelfingen, CH, USA), which were then incubated with the specific antibodies.
Horseradish-peroxidase-conjugated secondary antibodies against rabbit or mouse immunoglobulin (Cell Signaling Technology; 1:5000) were then used and reactions were visualized with the enhanced chemiluminescence kit (ECL) from GE Healthcare. Lysates were normalized using GAPDH protein as reference.

3.11. Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation assay was performed according to a previously published method (Raha 2005) with some modifications. Briefly, cells were treated with 1% formaldehyde and collected in 1X PBS. Cell pellets were re-suspended in lysis buffer (5mM PIPES pH8, 85mMKCl, 0,5% NP40, 1X Proteinase inhibitors) and sonicated for 10 seconds 18 times on ice (BRANSON S250 digital sonicator, Branson, Danbury, CT, USA). Sonicated chromatin was pre-cleared for 1 hour at 4°C using 60µl of pre-washed sepharose beads (KPL). Pre-cleared chromatin was quantified using Qubit® dsDNA BR Assay Kit (Life Tecnologies). 10µg of chromatin was incubated overnight at 4°C in dilution buffer (1% SDS, 10mM EDTA pH8, 50mM Tris-HCl pH8, 1X proteinase inhibitors) with 10µg of specific antibodies. Five percent of the total non-immunoprecipitated lysate was used for input control. Antibody coupled-chromatin was incubated using 60µl pf pre-washed protein G sepharose beads (KPL) for 2 hours at 4°C. The beads were extensively washed and the DNA was extracted by the phenol:chloroform:isoamyl alcohol, ethanol precipitated and resuspended in water. Chromatin immunoprecipitation products were amplified using GoTaq Hot-Start Polymerase (Promega Inc.) and specific primers.
To better compare the ChIP results, densitometric analysis was performed using ImageJ program. First, the obtained values were normalized on input signals as follow: IP) / Densitometric value (INPUT)) X 5 (% of the INPUT)

The resulting value represents the % of the recovery of the antibody. This value indicates the amount of the specific DNA sequence recovered by a specific antibody.

3.12. RRBS and data analysis

Reduced representation bisulfite sequencing (RRBS) was obtained using 4 µg of high-quality gDNA that was restriction digested using the methyl insensitive restriction enzyme MspI, which cuts the DNA at CCGG sites. The fragments were blunt-ended, 3'-end A-tailed and ligated to indexed adaptors using the TruSeq DNA PCR-free protocol (TruSeq Stranded DNA PCR-free Sample Preparation Kit, Illumina, San Diego, CA, USA); the ligation products purified (400 ng) were bisulfite-treated with the EZ DNA Methylation kit (Zymo Research) and PCR-amplified to enrich fragments using Taq PFU CX (Agilent). Libraries were sequenced (paired-end, 2x100 cycles) at a concentration of 8pmol/L per lane on HiSeq2500 platform (Illumina) with >50 million sequence reads/sample. The raw sequence files generated (.fastq files) underwent quality control analysis using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The adapter was trimmed using TrimGalore (Version 0.4.0) (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The sequence data obtained were analyzed using Bismark (version 0.16.1) and methylation profiling were analyzed using methylKit (version 0.9.5). Annotation of CpG was performed with Homer (version 4.6).

3.13. RNA sequencing and data analysis

Indexed libraries were prepared from 1 µg/ea purified RNA by using TruSeq Stranded Total RNA Sample Prep Kit (Illumina) according to the manufacturer's instructions. Libraries were sequenced (paired-end, 2x100 cycles) at a concentration of 8pmol/L per lane on HiSeq2500 platform (Illumina) with >70 million sequence reads/sample. The raw sequence files generated (.fastq files) underwent quality control analysis using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and the quality checked reads were then aligned to the human genome (hg19 assembly) using TopHat2 (version 2.0.13), with standard parameters. A given mRNA was considered

expressed when detected by ≥10 reads. Differentially expressed mRNAs were identified using DESeq (version 1.6.3). Firstly, gene annotation was obtained for all known genes in the human genome, as provided by Ensemble (GRCh37) (https://support.illumina.com/sequencing/sequencing_software/igenome.ilmn).
Using the reads mapped to the genome, we calculated the number of reads mapping to each transcript with HTSeq-count (version 0.6.0). These raw read counts were then used as input to DESeq for calculation of normalized signal for each transcript in the samples, and differential expression was reported as Fold Change along with associated adjusted p-values (computed according to Benjamini-Hochberg). Analysis of LINE-1 were performed using RepEnrich (version 1.2).

3.14. Histone acid extraction

Cells were washed by using PBS and directly scraped in PBS. Cell pellets were resuspend in 1 ml of Hypotonic buffer (10mM Tris_HCl pH8, 1mM KCl, 1.5mM $MgCl_2$, 1mM DTT an proteinase inhibitor Cocktail) and transfer to 1.5 ml eppendorf (more or less 10^7 cells for 1 ml). Incubate for 30 min on rotator at 4°C.
Nuclei were pelleted by centrifuging at 10000 g for 10 min at 4°C. The cell nucei were resuspended in 1ml of Low salt buffer (0.6M NaCl, 10mM Tris-HCl pH 8, 20mM EDTA and proteinase Inhibitor Cocktails) and incubate for 30 min on rotator at 4°C. Chromatin was pelleted by centrifuging at 10000 g for 10 min, resuspended in 500 µl of H_2SO_4 0.2M. (0.2M sulforic acid) and incubated at least 30 min on ice. The chromatin was manually disgregated and the supernatant containing the histones was collected. Histones were precipitated by adding TCA (final concentration 33%), incubated, incubated on ice for 30 min and centrifuged at 16000 g for 5 min at 4°C. Final pellet containing histones was resuspended in 50µl of water. Histones were used for either WB or SILAC analysis.

3.15. SILAC experiments

Control and UHRF1 silenced cells were grown in SILAC media (lysine- and arginine-free DMEM/Ham's F12 (1:1), 10% dialyzed fetal bovine serum (FBS, Invitrogen)),

supplemented with 1mM non-essential amino acids (Gibco), 100 U/ml of penicillin and streptomycin (Lonza), 1mM Na-pyruvate (Gibco), 2mM Glutamine (Lonza) and 50 mM β-mercaptoethanol (Gibco)). "Heavy" and "Light" media were obtained by adding 0.146 g/L 13C6, 15N2L-Lysine and 0.84 g/L 13C615N4L-Arginine (Sigma) or the corresponding non-labeled amino acids, respectively to the SILAC media. Growth in SILAC media was carried out for eight duplications, to ensure complete protein labeling. For the experiments transfected cells were labeled with heavy medium, and control cells with light medium.

"In solution" digestion using Arg-C

Equal numbers (12×10^6) of heavy and light cells were mixed and 10μg of histones were incubated overnight at 37°C with endoproteinase Arg-C. After incubation the reaction was stopped by adding 50% volume of TFA. Product peptides were separated according to their isoelectric point by isoelectrofocusing electrophoresis using the Agilent 3,100 OFFGEL Fractionation Kit (Agilent Technologies). Samples were prepared and separated according to the manufacturer's protocol. After separation, peptides mixtures were reconstituted with 1% TFA and desalted on C18 STAGE tips.

"In-gel" digestion

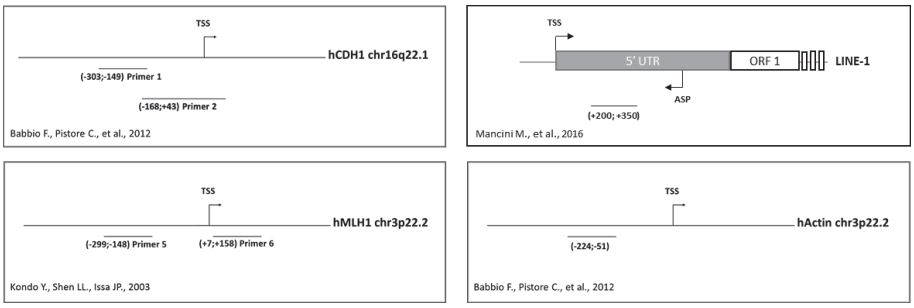
Protein samples were separated on a 8–12% gradient mini gel (Invitrogen). After Coomassie staining (Colloidal Blue Staining Kit, Invitrogen), each line was cut according to the specific bands and trypsin digested according to a previously described protocol (J. Rappsilber, et al., Nat. Protoc. 2007). Briefly, after distaining with 50% acetonitrile (ACN)/25 mM ammonium bicarbonate (NH_4HCO_3) solution, and dehydration by 100% ACN, gel pieces were incubated with 10mM dithiothreitol in 50mM NH4HCO3 for 60 min at 56°C for cysteine reduction and then alkylated with 55mM iodoacetamide in 50mM NH_4HCO_3 for 45 min at room temperature, in dark. After several rounds of washings with 50mM NH_4HCO_3 and dehydration with 100% ACN, proteins were digested with trypsin overnight, at 37 °C. The reaction was stopped by acidification with 2μl of 50% trifluoroacetic acid (TFA). Peptides were eluted with 30% ACN/3% TFA and 100% ACN. After speed-vacuum centrifugation, peptides were solubilized in 100μl of 0.1% formic acid (FA), desalted and concentrated using reverse phase C18 Stage Tips. Peptides were eluted with 80% ACN, lyophilized and re-suspended in 7μl of 0.1% formic acid for LC–MS/MS analysis.

3.16. Statistical analysis

Association analyses were performed using the Fisher exact test, ANOVA analysis, and the independent sample t-test. To define an appropriate threshold for UHRF1 immunohistochemical positivity a model based cluster algorithm (Raftery 2002) was used. Patient survival was evaluated using the Kaplan-Meier method and statistically tested with the log-rank test. Multivariate analysis was performed with the Cox Proportional hazard model using the backward method. A p-value <0.05 was considered statistically significant. These analyses were performed using MedCalc Statistical (version 11.0.1.0) and GraphPad Prism V5.0 software.

3.17. Primers and oligonucleotides

Gene name	FOR	REV	Methods
UHRF1	CTGGTACGACGCGGAGAT	CGACAGTCGTTCAGAGAATCA	PCR
MLH1	ACTCCTGGAAGTGGACTGTG	GATCAGGCAGGTTAGCAAGC	PCR
CDH1	ATGAGTGTCCCCCGGTATCTTC	ACGAGCAGAGAATCATAAGGCG	PCR
GAPDH	GAGTCAACGGATTTGGTCGT	TTGATTTTGGAGGGATCTCG	PCR
MLH1-5	CTTGCTTCTTTTGGGCGTCAT	GGCTTGTGTGCCTCTGCTGA	ChIP
MLH1-6	CCCAGCAACCCACAGAGTTGAG	CGGAAGTGCCTTCAGCCAATC	ChIP
CDH1-1	GAACTCAGCCAAGTGTAAAAGC	AGACGCGGTGACCCTCTA	ChIP
CDH1-2	TAGAGGGTCACCGCGTCTAT	GACTTCCGCAAGCTCACAG	ChIP
LINE-1	GGCCAGTGTGTGTGCGCACCG	CCAGGTGTGGGATATAGTCTCGTGG	ChIP



Genomic schemes of ChIP primers.

3.18. Antibodies

Protein name	Supplier	Dilution	Application
UHRF1	IGBMC, monoclonal	1:4000;	WB, ChIP
UHRF1	Bonapace' s Lab	1:4000	IHC
CDH1	BD Biosciences, monoclonal	1:3000	WB
MLH1	BD Biosciences, monoclonal	1:1000	WB
IgG	Santa Cruz		ChIP
H3K4me3	Active Motif, polyclonal		ChIP
H3K9me3	Active Motif, polyclonal		ChIP
H3K27me3	Active Motif, polyclonal		ChIP
DNMT1	Active Motif, monoclonal	1:250	WB, ChIP
DNMT3a	Active Motif, monoclonal	1:1000	WB, ChIP
DNMT3b	Active Motif, monoclonal	1:1000	WB, ChIP
Ub total	Cell signaling, monoclonal	1:1000	WB
H3	Abcam, polyclonal	1:3000	WB
H2BK120ub	Active Motif, monoclonal	1:1000	WB
GAPDH	Millipore, monoclonal	1:4000	WB
H3K18ub	Bonapace' s Lab, monoclonal	1:500	WB

4. Results

4.1. Evaluation of UHRF1 and DNA methylation levels in human specimens

In the attempt to better understand the role of UHRF1 in colon cancer and its possible involvement in the regulation of DNA methylation both at genome-wide and at locus-specific level, we performed IHC analysis of UHRF1 protein in a cohort of colorectal mucosa specimens, in collaboration with Dr. Daniela Furlan (Anatomic Pathology Unit, University of Insubria, Varese, Italy), Dr. Anna Maria Chiaravalli (Anatomic Pathology Unit, University of Insubria, Varese, Italy) and Prof. Vittorio Colantuoni (Dept. BGES, University of Sannio, Benevento, Italy).

4.1.1. UHRF1-high correlates with MSI colorectal cancer and better prognosis

IHC analysis of UHRF1 was performed in a cohort of colorectal mucosa samples, composed by 25 normal mucosae (divided in 13 adenomas near to CRCs, and 12 normal mucosae from healthy patients) and 135 colorectal cancer samples (divided in 47 MSI CRCs, and 88 MSS CRC).

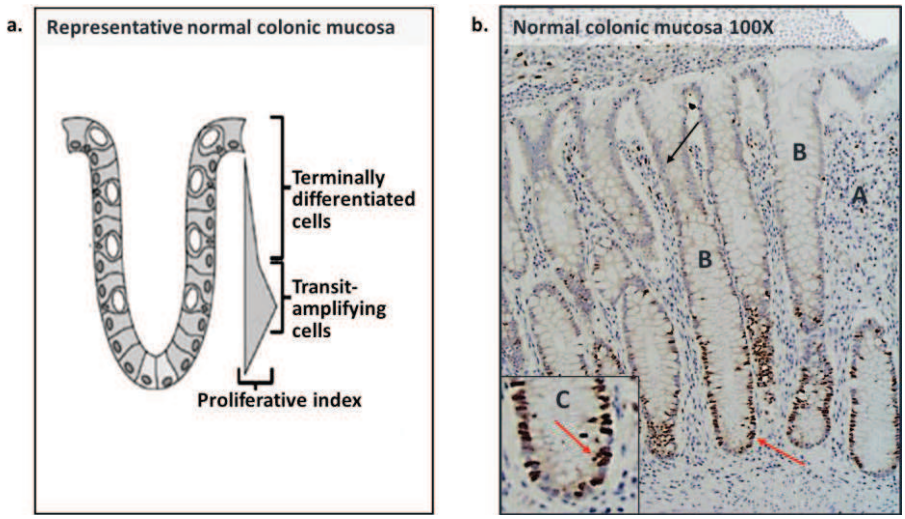


Figure 4.1. Immunohistochemical analysis of UHRF1 in normal colonic mucosa. a) Representative diagram of a normal colonic mucosa (Sabatino et al, 2012). b) IHC analysis of UHRF1: UHRF1 nuclear staining depicts the lower third of the crypt, marking the cells with the highest proliferative index, as shown in the representative diagram on the left. UHRF1 is detected with ABC-peroxidase, DAB-hematoxylin (100x).
A: Lamina propria; B: intestinal crypt; C: Detail of the lower third of the crypt.
Red arrow: cells positive to UHRF1 staining; Black arrow: cells negative to UHRF1 staining.

UHRF1 showed a nuclear positive staining in the lower third of the crypt, in the normal colonic mucosa (Figure 4.1b) where are located the cells in active proliferation (Figure 4.1b, red arrows). The enterocytes, terminally differentiated cells, in the middle-upper part of the crypt, showed a UHRF1 negative staining, in accordance with our previous reports [Sabatino et al.; Oncogene, 2012] [Bonapace et al.; J. Cell. Biol., 2002].

The normal mucosae, adjacent respectively to MSS (Figure 4.2a) and MSI (Figure 4.2b) tumors, showed a staining similar to what we observed in the normal mucosa, as shown in Figure 4.1b.

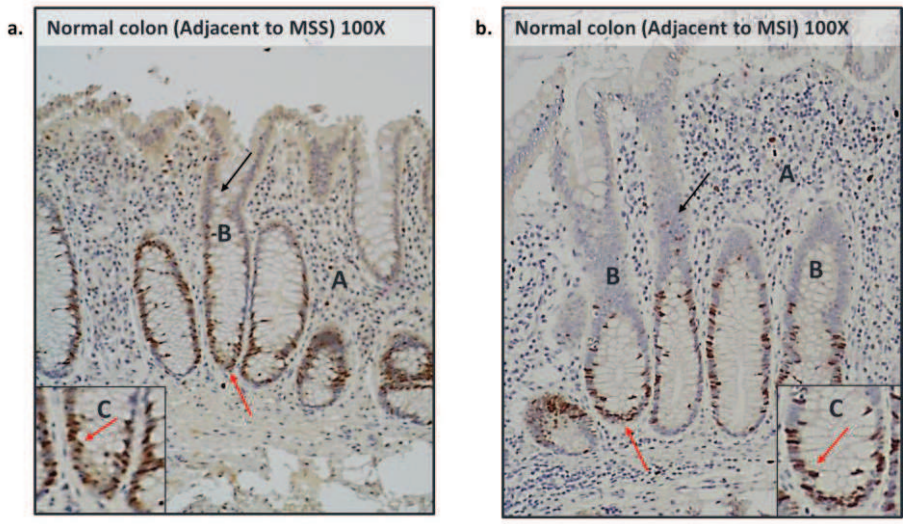


Figure 4.2. Immunohistochemical assay of UHRF1 in normal colonic mucosa adjacent to CRCs: a) Close to MSS-CRC; b) Close to MSI-CRC. In both cases, UHRF1 nuclear staining depicts the lower third of the crypt, similar to control shown in figure 4.1b. UHRF1 is detected with ABC-peroxidase, DAB-hematoxylin (100x).
A: Lamina propria; B: intestinal crypt; C: Detail of the lower third of the crypt.
Red arrow: cells positive to UHRF1 staining; Black arrow: cells negative to UHRF1 staining.

On the contrary, the tissue architecture of the tumor samples is completely subverted (Figure 4.3). Colon cancer cells, in adenocarcinomas, can extend from mucosa to muscle layer and beyond, and infiltrate in solid individual groups or aggregate in glandular and tubulo-papillary structures.

IHC analysis showed that the number of UHRF1 positive cells are significantly increased in CRC samples compared to the normal mucosa (Figure 4.3). Surprisingly, among the CRC groups, UHRF1 is significantly higher in MSI CRC samples, which have a better prognosis, compared to MSS CRC (Figure 4.4).

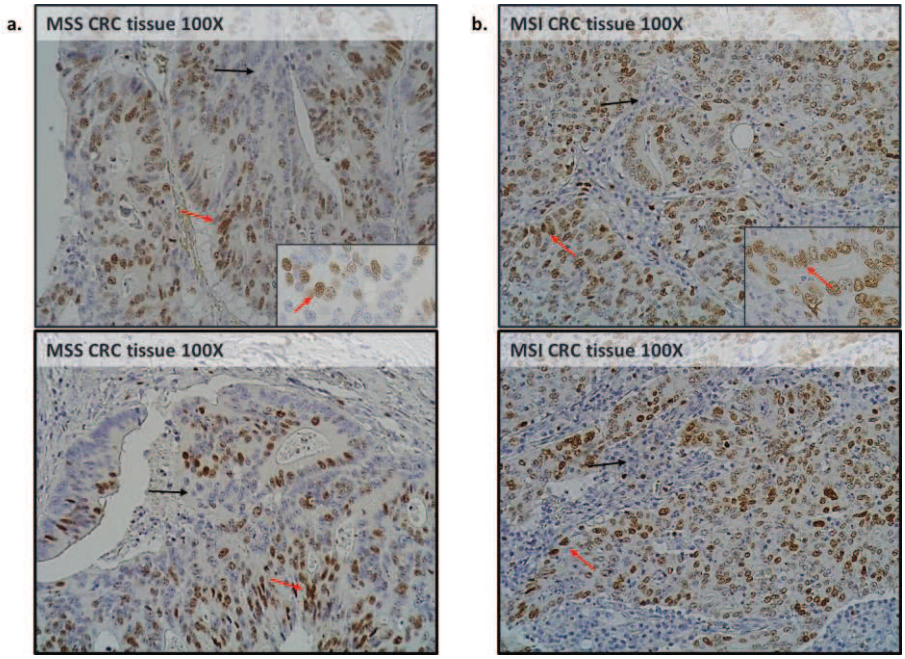


Figure 4.3. Immunohistochemical analysis of UHRF1 in CRC tissues: a) Two representative slices of human MSS-CRC; b) Two representative slices of human MSI-CRC. UHRF1 positive cells were counted in each field, for a total of 5 fields per slide. UHRF1 is detected with ABC-peroxidase, DAB-hematoxylin (100x).
Red arrow: cells positive to UHRF1 staining; Black arrow: cells negative to UHRF1 staining.

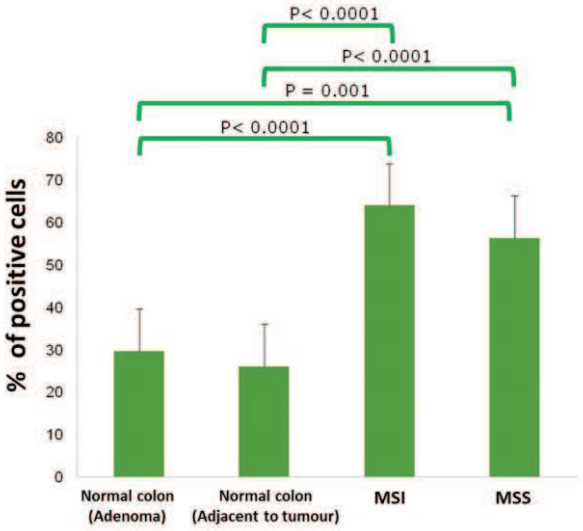


Figure 4.4. Plot of the average percentage of UHRF1 positive cells. UHRF1 positive cells were calculated with a model based cluster algorithm (Raftery 2002). The result was expressed in percentage of positive cells normalized on the total number of cells. The statistical significance was evaluated by the t-test.

In order to assess the possible correlation between UHRF1 positivity and patient's outcome, all CRCs were divided into two groups using a cut-off value of 50% of UHRF1 positive cells, that was calculated by using a model based cluster algorithm [Fraley C. and Raftery AE; Journal of the American Statistical Association, 2002]. We defined as UHRF1 low (UHRF1-L) the samples with less than 50% of positive cells per field and UHRF1 high (UHRF1-H) the samples with more or equal 50%. This threshold

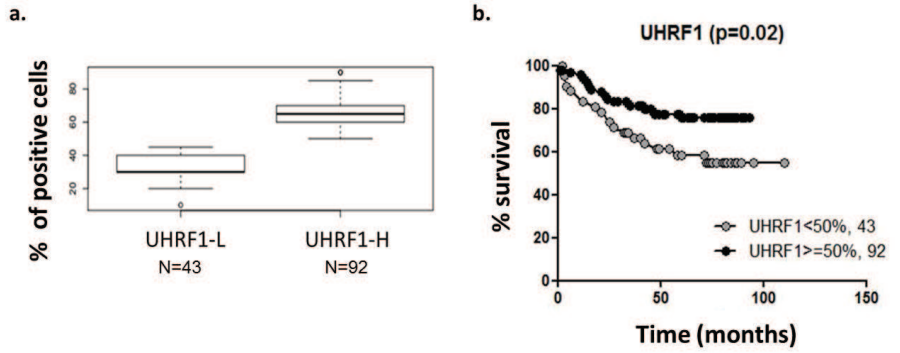


Figure 4.5. a) Stratification of the series of 135 CRCs based on the levels of UHRF1 immunohistochemical expression. The threshold value (<50% positive cells: UHRF1-L; ≥50% of positive cells: UHRF1-H) was calculated by applying the algorithm of Model Based Clustering (ref). b) Univariate survival analysis according to UHRF1 immunohistochemical expression in CRCs.

value clearly separated two subsets of 43 and 92 CRCs, exhibiting low- (UHRF1-L) or high- (UHRF1-H) positivity of the protein, respectively (Figure 4.5a). The UHRF1-H group significantly correlated with the better prognosis, in terms of overall survival. We also observed that UHRF1-H CRCs were strongly associated with MSI (p=0.0019) and with BRAF mutation (p=0.013) (data not shown). Surprisingly, univariate survival analysis (Figure 4.5b, Kaplan-Meier) showed that UHRF1-H was correlated with better prognosis compared with UHRF1-L (p=0.02).

4.1.2. Aberrant DNA methylation profiles in MSI and MSS CRCs correlates with UHRF1 positivity

UHRF1 positivity correlated with DNA methylation profiles in normal colorectal tissues as well as in MSI and MSS CRCs (Figure 4.6). We analyzed global DNA methylation, by the pyrosequencing analysis of the repetitive sequence LINE-1, considered as a surrogated of global DNA methylation. Gene-specific methylation was assessed by performing MS-MLPA assay on 34 selected Tumor Suppressor Genes (TSGs) promoters. MS-MLPA is a well-established assay used for the clinical

characterization of the tumors. We observed an overall global hypomethylation and a gene specific hypermethylation in the CRC groups (Figure 4.6, box plot, middle and right part) compared to normal mucosae (Figure 4.6, box plot, left part). This data supported the actual paradigm for solid tumors development in which genome-wide hypomethylation and gene-specific hypermethylation are constantly associated in all solid tumors including the CRCs [Esteller M; N. Engl. J. Med., 2008] [Ehrlich M; Epigenomics, 2009]. Nevertheless, considering the CRCs only (Figure 4.6, MSI vs. MSS-CRCs), high levels of UHRF1 correlated to higher degree of DNA methylation both at global and at specific sites. This detailed comparison between DNA methylation levels in MSI and MSS CRCs enlightened a new intriguing finding regarding UHRF1 overexpression in cancer. In fact, in several cancer types, UHRF1 deregulation and overexpression were associated to cancer progression and worse prognosis [Unoki M., et al.; Br J Cancer., 2010] [Babbio F, et al.; Oncogene, 2012] [Mudbhary R, et al.; Cancer Cell, 2014]. However, comparing only the MSI and MSS CRC groups, that showed as expected an overall UHRF1 overexpression compared to normal mucosae, UHRF1-high correlated with MSI model and with a better prognosis.

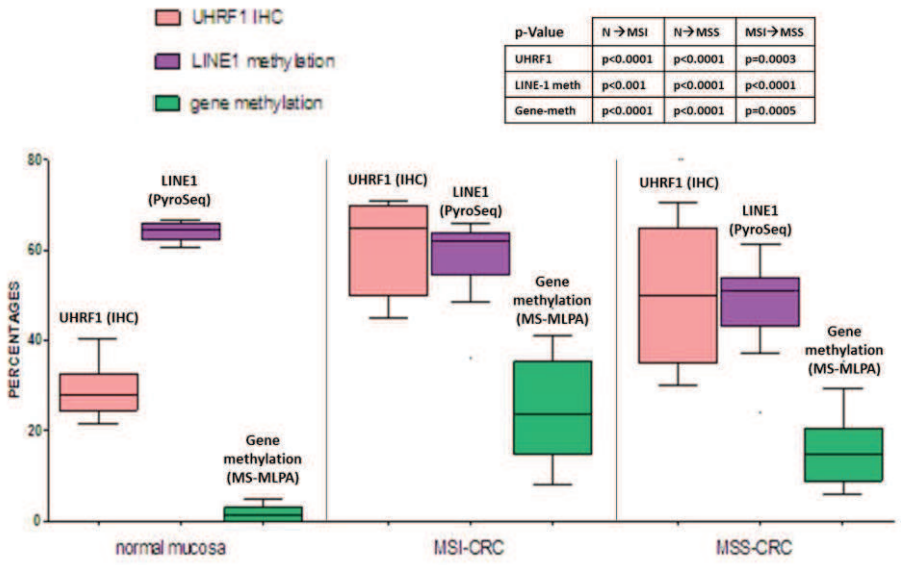


Figure 4.6. Correlation Box Plot: UHRF1 IHC analysis (pink), LINE-1 methylation measured by pyrosequencing, considered as surrogated of global methylation (purple) and gene specific methylation levels measured by MS-MLPA analysis (green). The table (top right) represented the p-value of the different correlation reported. The percentages on Y axis represented respectively IHC percentage of positive cells for UHRF1 staining and percentages of DNA methylation at global level (LINE-1 pyrosequencing) and gene promoters (MS-MLPA).

The experiments performed in this part of the thesis clearly demonstrated that the number of UHRF1 positive cells is significantly increased in CRCs compared to normal mucosa, in accordance with our previous publication [Sabatino L, Fucci A, et al.; Oncogene, 2012]. Here, we performed a more detailed analysis and we defined a strong correlation between UHRF1 positivity and MSI colorectal cancer. In fact, surprisingly, among the CRC groups, UHRF1 was significantly higher in MSI CRC samples, which have a better prognosis, compared to MSS CRCs. Furthermore, UHRF1 positivity correlated with an overall DNA hypermethylation among the CRC samples, both at genome-wide and at gene specific level. This could result in a sort of protective condition for the genome integrity, that, despite the tumor suppressor genes hyper-methylation, leads to a better prognosis [Estécio MRH, et al.; PLoS ONE, 2007] [Matsuzaki K, et al.; Clin. Cancer Res., 2005] [Holm TM, et al.; Cancer Cell, 2005].

This new intriguing data, observed in human specimens, supported us to deepen the molecular mechanism that underlies this clinical phenomenon. The project has been developed further analyzing UHRF1 and DNA methylation levels in two cell lines: RKO cells, a cell line derived from MSI CRC, and HT29, derived from MSS CRC.

4.2. Analysis of UHRF1 and DNA methylation levels in colorectal-cancer-derived cell lines (RKO & HT29)

To deepen the molecular mechanism that underlies the clinical phenomenon previously observed, we broadened the analysis to two cell lines: RKO cells, a cell line derived from MSI CRC, and HT29, derived from MSS CRC.

4.2.1. UHRF1 protein levels in RKO and HT29 are consistent with MSI and MSS-CRCs

As first step, we needed to evaluate and analyze the molecular features of the two CRCs cell lines. In WB analysis (Figure 4.7), we observed very different levels of UHRF1 protein in the two cell lines. In fact, RKO (MSI-CRC cell lines) showed higher levels of UHRF1 compared to HT29 (MSS-CRC cell lines). Then, we investigate the levels of two important tumor suppressor genes: MLH1, key gene involved in the Mismatch Repair (MMR), whose loss/mutation/silencing is crucial in the establishment of the microsatellite instability; CDH1, coding for the E-cadherin protein, crucial for the cell-cell adhesion, antagonist of cell invasiveness, and well-known to be negative regulated by UHRF1 in prostate tumors [Babbio F, et al.; Oncogene, 2012]. According to the different levels of UHRF1, we observed good levels of CDH1 and MLH1 in HT29, and no expression of both of them in RKO cells.

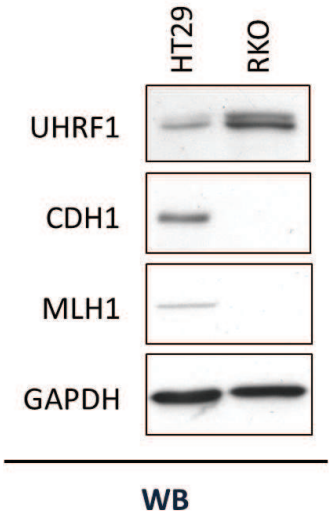


Figure 4.7. Western Blot analysis of total protein extracts of RKO and HT29 cells.

4.2.2. UHRF1 overexpression correlates with DNA hypermethylation in CRC cell lines

In order to compare the cell lines to what observed in the human specimens, we analyzed DNA methylation profile by using the same approach used for the histological specimens (pyrosequencing of LINE-1 and MS-MLPA of the same 34 selected TSGs).

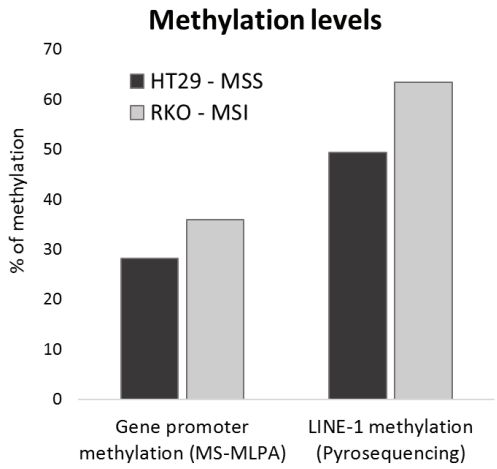


Figure 4.8. Summary DNA methylation data. Gene promoter methylation was measured by MS-MLPA analysis. LINE-1 methylation, as surrogated of global methylation, was assessed by DNA pyrosequencing of 4 CpGs located in 5'UTR of LINE-1 sequences.

DNA methylation, both at global level (LINE-1, pyrosequencing) and at specific sites (Gene promoter, MS-MLPA), is higher in RKO (MSI-derived) compared to HT29 (MSS-derived), according to the higher level of UHRF1 in this cell line (Figure 4.8). In particular, the promoter of MLH1 gene, whose silencing is crucial to establish the MSI CRC, showed a prominent differential value in the Methylation Dosage Ratio (MR) obtained from the MS-MLPA analysis. In fact, RKO showed 0.9 MR comparare to 0.1 MR of HT-29 cells, with a cut-off to consider a promoter methylated of 0.2 MR (data not shown singularly). This site specific DNA hypermethylation was consistent with MLH1 silencing in RKO cells, as shown previously in WB analysis (Figure 4.7).

Altogether these results showed that UHRF1 protein levels correlate positively to the DNA methylation levels in CRC cell lines, according to the *in vivo* observations. In fact, the two analyzed cell lines recapitulated well the features previously observed in the human samples. In order to evaluate the potential link between UHRF1 and the observed hyper-methylation we decided to manipulate the UHRF1 levels *in vitro*. Due to the interesting molecular patho-genetic features (MSI CRC model, CIMP-high, CIN-negative, BRAF-V600E, PI3KCA-H1047R) [Ahmed D., et al.; Oncogenesis, 2013], and the high levels of UHRF1 we selected the RKO cells for the further analysis.

4.3. Evaluation of the effect of UHRF1 knock down on DNA methylation, gene expression and histone modifications in RKO cells

In order to modulate the level of UHRF1, and evaluate the effect on DNA methylation patterns, genes expression and histone PTMs profile, we performed a transient silencing of UHRF1 by short interfering RNA (siRNA) in RKO cells.

4.3.1. UHRF1 KD decreases DNA methylation levels in MSI cell model (RKO)

To set up the transfection protocol, two different transfection reagents were used on RKO cells: Interferin (Polyplus) and RNAiMax (Life Techologies). Between them, the RNAiMax (together with siUHRF1 L ife Techologies, at 10nM, for 72h) gave us the best results both in term of cell viability, reproducibility and UHRF1 silencing.

We evaluated the efficiency of the UHRF1 silencing by WB analysis (Figure 4.9a). After the initial set up, we were able to obtain a reproducible know down of 95% in every experiment. Then, we investigate the impact of UHRF1 depletion on DNA methylation by pyrosequencing analysis of MLH1 and CDH1 promoters and of 5'UTR LINE-1 sequences, as surrogated of global methylation. UHRF1 knock down induced an overall decrease in DNA methylation (Figure 4.9b) both at global level (LINE-1) and at gene promoters of MLH1 and CDH1. LINE-1 showed a DNA de-methylation of 32.17% (± 0.77%), maintaining a residual methylation of 38.10% (± 0.35%). MLH1

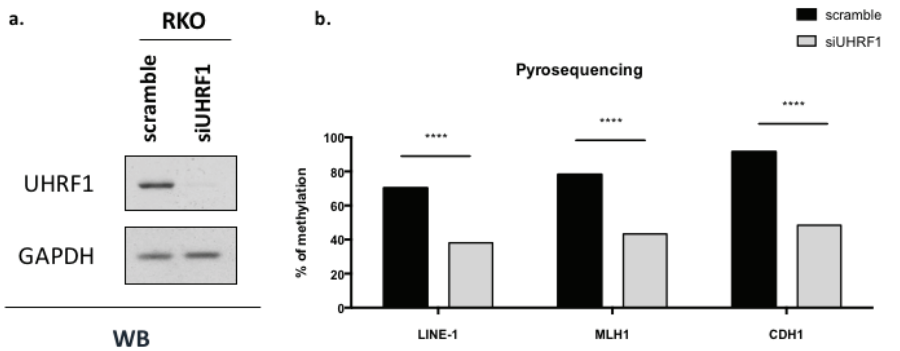
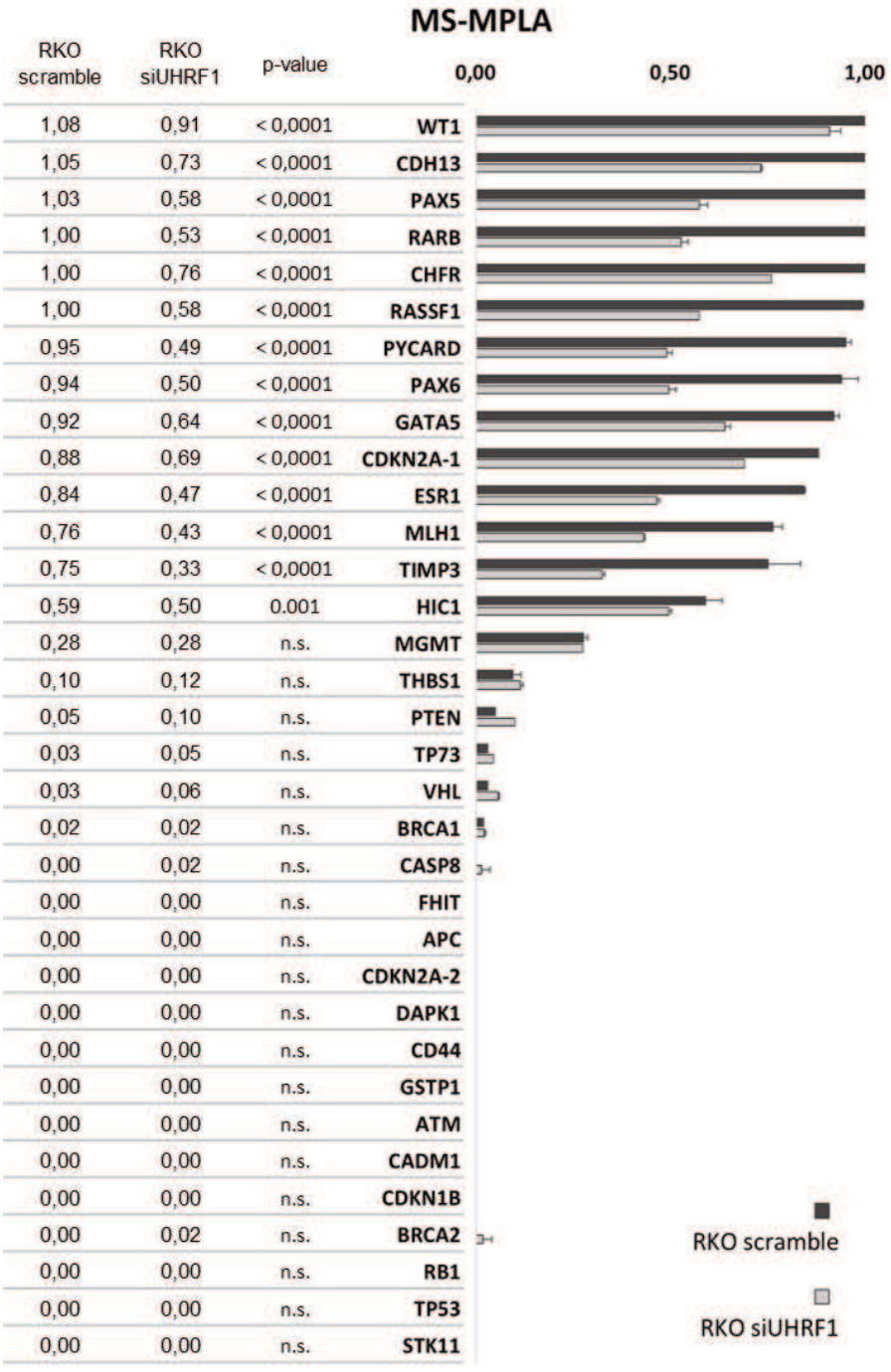


Figure 4.9. a) WB analysis of total protein extracts from scramble and siUHRF1 RKO cells, after 72h from transfection. B) Histogram of the pyrosequencing results; LINE-1 methylation, as surrogated of global methylation, was assessed by DNA pyrosequencing of 4 CpGs located in 5'UTR of LINE-1 sequences. MLH1 and CDH1 methylation was assessed by DNA pyrosequencing of respectively 5 and 6 CpGs located in their promoters. All percentages derived from an average between two independent biological replicates. Significance has been calculated using Anova. $p < 0.0001$

showed a decrement in DNA methylation of 43.08% (± 1.24%), maintaining a residual methylation of 48.47% (± 0.47%). CDH1 showed a DNA de-methylation of 34.86% (± 1.68%), maintaining a residual methylation of 43.33% (± 0.05%).

Subsequently, we analyze the DNA methylation levels in the same panel of 34 oncosuppressor genes analyzed by MS-MLPA assay in the human samples. We observed an overall decrease in DNA methylation (Figure 4.10), upon UHRF1 silencing, almost in all the analyzed promoters. Further, MS-MLPA data of MLH1 promoter confirmed the Pyrosequencing result (shown previously in Figure 4.9b), assessing the differential methylation at 43.42%. Altogether these data led us to hypothesize that in RKO cell lines, the high levels of UHRF1 could mediate the transcriptional repression of these TSGs, including CDH1 (see pyrosequencing in Figure 4.9b). In this scenario, the modulation of UHRF1 should be sufficient, not only to change the methylation levels but also the expression of the TSGs.

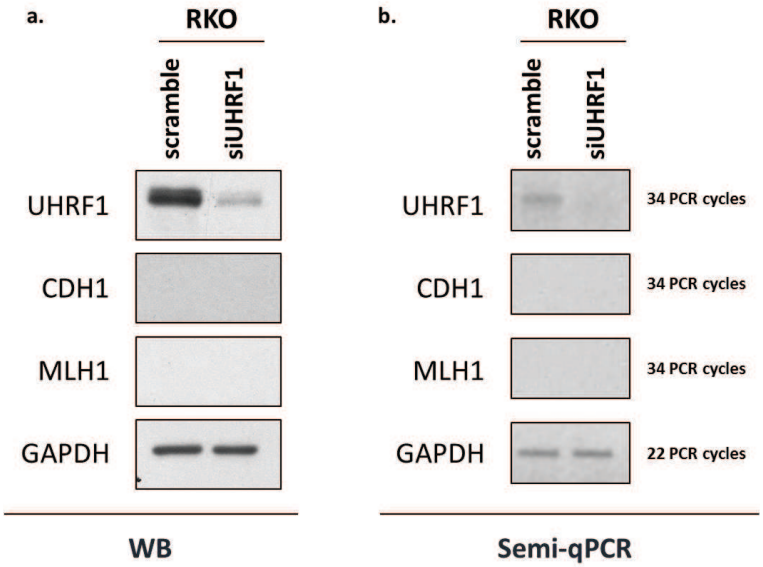


RKO scramble	RKO siUHRF1	p-value	
1,08	0,91	< 0,0001	WT1
1,05	0,73	< 0,0001	CDH13
1,03	0,58	< 0,0001	PAX5
1,00	0,53	< 0,0001	RARB
1,00	0,76	< 0,0001	CHFR
1,00	0,58	< 0,0001	RASSF1
0,95	0,49	< 0,0001	PYCARD
0,94	0,50	< 0,0001	PAX6
0,92	0,64	< 0,0001	GATA5
0,88	0,69	< 0,0001	CDKN2A-1
0,84	0,47	< 0,0001	ESR1
0,76	0,43	< 0,0001	MLH1
0,75	0,33	< 0,0001	TIMP3
0,59	0,50	0.001	HIC1
0,28	0,28	n.s.	MGMT
0,10	0,12	n.s.	THBS1
0,05	0,10	n.s.	PTEN
0,03	0,05	n.s.	TP73
0,03	0,06	n.s.	VHL
0,02	0,02	n.s.	BRCA1
0,00	0,02	n.s.	CASP8
0,00	0,00	n.s.	FHIT
0,00	0,00	n.s.	APC
0,00	0,00	n.s.	CDKN2A-2
0,00	0,00	n.s.	DAPK1
0,00	0,00	n.s.	CD44
0,00	0,00	n.s.	GSTP1
0,00	0,00	n.s.	ATM
0,00	0,00	n.s.	CADM1
0,00	0,00	n.s.	CDKN1B
0,00	0,02	n.s.	BRCA2
0,00	0,00	n.s.	RB1
0,00	0,00	n.s.	TP53
0,00	0,00	n.s.	STK11

Figure 4.10. MS-MLPA analysis performed in RKO cells after UHRF1 silencing. The table on left reports the DNA methylation values assessed for the promoters of 34 tumor suppressor genes, considered crucial for tumorigenesis of CRC. Statistical significance was calculated by Anova test.

4.3.2. UHRF1 silencing is not sufficient for the re-expression of CDH1 and MLH1 genes in RKO cells

In order to verify whether UHRF1 depletion is sufficient to induce a transcriptional re-expression of CDH1 and MLH1 genes, we performed a WB and PCR on UHRF1 KD RKO cells, investigating respectively proteins and mRNAs levels.



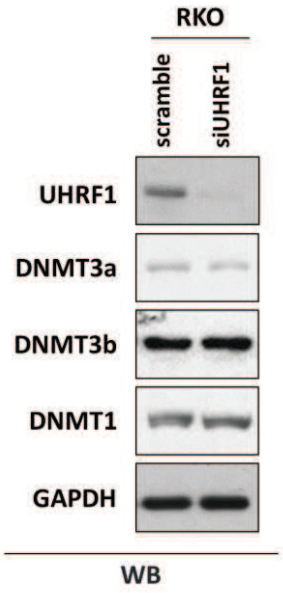
4.11. a) WB analysis of total protein extracts from scramble and siUHRF1 RKO cell lines. b) Semi-quantitative PCR analysis of mRNA (cDNA) expression levels of UHRF1, MLH1 and CDH1. GAPDH was used as reference gene.

The WB analysis (Figure 4.11a) confirmed that the UHRF1 protein level was strongly reduced in the silenced UHRF1 sample (siUHRF1) compared to the control (scramble). Nevertheless, this decrement was not sufficient to observe a re-expression of CDH1 and MLH1 proteins. In semi-quantitative PCR (Figure 4.11b), we analyzed changes in mRNA expression levels of UHRF1, CDH1 and MLH1, upon UHRF1 silencing (after 72h from transfection). UHRF1 mRNA level, as expected, was lower in siUHRF1 sample compared to control (after 34 cycles of PCR amplification). Regarding CDH1 and MLH1, there was not a re-expression either at transcriptional level (after 34 PCR cycles).

These results showed that, even if UHRF1 KD significantly reduced the DNA methylation on CDH1 and MLH1, it was not sufficient to induce their re-expression, probably due to other convergent mechanisms of transcriptional repression on these promoters (see below for more details).

4.3.3. DNA de-methylation on LINE-1, CDH1 and MLH1 promoter regions is induced by the DNMT1 delocalization

Since the relationship between UHRF1 and the DNA methyl-transferases, either of DNMT1 [Bostick M, et al.; Science, 2007] [Sharif J, et al.; Nature, 2007], or DNMT3A, DNMT3B [Meilinger D, et al.; EMBO reports, 2009] is well known, we investigated better the changes mediated by the UHRF1 depletion on these crucial epigenetic actors in RKO cells.



4.12. WB analysis of total protein extracts from scramble and siUHRF1 RKO cell lines, after 72h from transfection. GAPDH was used as reference gene.

In the attempt to elucidate the molecular basis of the strong DNA de-methylation previously observed, we performed a WB analysis on DNMTs protein levels in UHRF1 depleted RKO cells. UHRF1 KD did not induce any significant alteration in the protein levels of the DNA methyl-transferases (Figure 4.12). This result suggested that DNA de-methylation, observed in previous experiments, was not dependent on DNMTs protein degradation but probably was due to a delocalization of DNMTs on target foci.

In order to validate our hypothesis and to establish whether UHRF1 knock down was impairing the DNMTs recruitment, we performed a Chromatin Immunoprecipitation (ChIP) analysis of the DNMTs (Figure 4.13) on LINE-1 5'UTR sequences and on MLH1 and CDH1 promoter regions.

ChIP analysis showed a strong decrease of UHRF1 binding, as expected upon UHRF1 knock down. The absence of UHRF1 impaired the DNMT1 recruitment to both specific gene promoters (MLH1, CDH1) and repetitive elements (LINE-1). This data was consistent with literature about the tight interaction between UHRF1 and DNMT1 [Bostick M, et al.; Science, 2007] [Sharif J, et al.; Nature, 2007], and strongly supported our hypothesis about DNMT1 delocalization as the major cause of DNA de-methylation observed in these same regions (see Figure 4.9b). DNMT3A showed a delocalization on MLH1 and a mild decreased binding on CDH1 promoter regions, while there was no decrease on LINE-1 sequence. DNMT3B, on the other side revealed a clear increased binding to all analyzed regions, except for CDH1 promoter

(in particular amplicon 1), in which DNMT3B seemed almost stable, starting from an already higher level in the control.

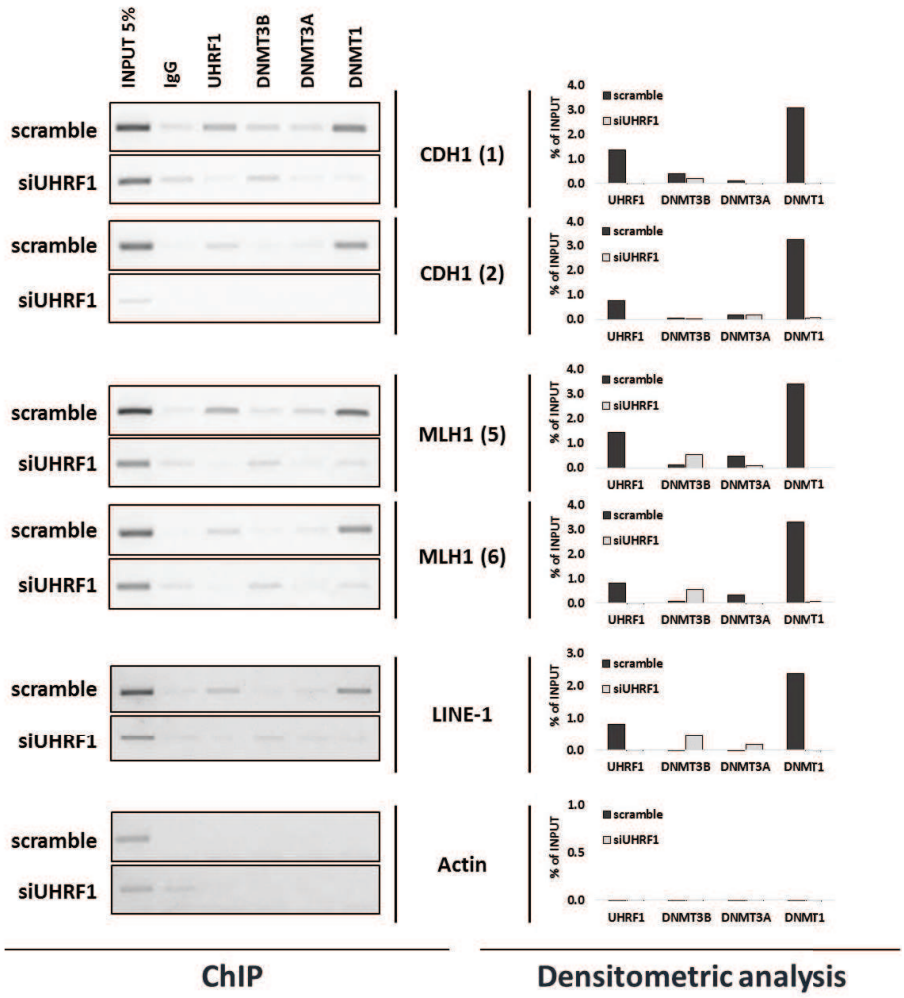


Figure 4.13. ChIP analysis on DNMTs, investigating their chromatin localization on LINE-1 5'UTR sequences and on MLH1 and CDH1 promoter regions. Images on left shows the PCR products starting from ChIP samples. Actin gene was used as negative control. Bar-plot on the right shows the relative densitometric analysis of detected PCR bands.

These observations could be due to the fact that DNMT3A/3B does not require the presence of interacting-chromatin-actors (e.g. HP1, MeCp2, EZH2, UHRF1, HDAC1, PCNA), or particular histone modification profiles, to corroborate DNMT1 in its ongoing participation in DNA methylation maintenance [Jones PA and Liang G; Nature Rev. Genet., 2009]. As expected, we denoted a complete absence of all the investigated actors on the promoter region of the Actin gene.

4.3.4. PRC2, via H3K27me3 deposition, promotes convergent mechanisms of transcriptional repression on LINE-1, CDH1 and MLH1 promoters

In the attempt to better investigate the chromatin profile of the analyzed regions and to better elucidate the involved molecular mechanisms, we performed a ChIP (on the same regions of LINE-1, MLH1 and CDH1) of H3K9me3, H3K27me3, and H3K4me3 [Kouzarides T; Cell, 2007]. H3K9me3 is a well-known repressive, generally heterochromatic, histone mark related to UHRF1, in particular it is strongly recognized by TTD domain of UHRF1 [Cheng J, et al.; 2013; J Biol Chem] [Xie S, et al.; 2012; J Biol Chem] [Rottach A, et al.; 2010; Nucleic Acids Res.]. Further, UHRF1 mediates its deposition by the direct interaction with the histone-methyl-transferases (HMTs) G9a (mono- e di-methylation of H3K9) and Suv39H1 (try-methylation of H3K9) [Kim JK, et al.; Nucleic Acids Research, 2009] [Meilinger D, et al.; EMBO reports; 2009] [Babbio F, et al.; Oncogene, 2012]. H3K27me3 is a chromatin repressive mark deposed by EZH2, part of Polycomb Repressive Complex 2 (PRC2), mainly involved in the gene silencing during the normal cell commitment, and often found mis-distributed on oncosuppressor CpG island (CGI) in cancer [Richly H, et al; Cell Death and Disease, 2011] [Schuettengruber B, and Cavalli G; Development, 2009]. We previously showed that in prostate cancer cells, EZH2 acts in an independent and convergent mechanism from UHRF1 to silence TSGs [Babbio F, et al.; Oncogene, 2012]. On the contrary, H3K4me3 is a chromatin active mark, deposed by a several redundant HMTs (e.g. PRDM9, Set9, MML1, ASH1), well-known to regulate transcriptional activation by promoting the binding of positive transcription factors [Li H, et al; Nature, 2005], and blocking negative ones [Nishioka K, et al.; Nature 2006].

Despite the de-localization of DNMT1 and the de-methylation observed previously, we did not observe a re-expression of MLH1 and CDH1 gene. These data suggested that their promoter regions could be regulated by several convergent mechanisms of transcriptional repression. We performed a ChIP analysis on hPTMs in order to verify the local chromatin state of our interesting regions (Figure 4.14).
We observed a strong alteration in chromatin profile of histone marks, upon UHRF1 silencing. In particular, we detected the strong decrease of H3K9me3 in the siUHRF1 samples compared to controls, in accordance with literature [Babbio F, et al.; Oncogene, 2012]. The co-occurrence in the control samples of H3K9me3 and a weak signal of H3K4me3 was not in contrast with the UHRF1 binding, denoted previously

in Figure 4.13. In fact, it has been reported that UHRF1-PHD domain is able to recognize the histone H3 tails also in presence of H3K4me3 (with H3K2unmodified), linking UHRF1 to the repression of euchromatin [Rajakumara, et al.; Mol. Cell, 2011].

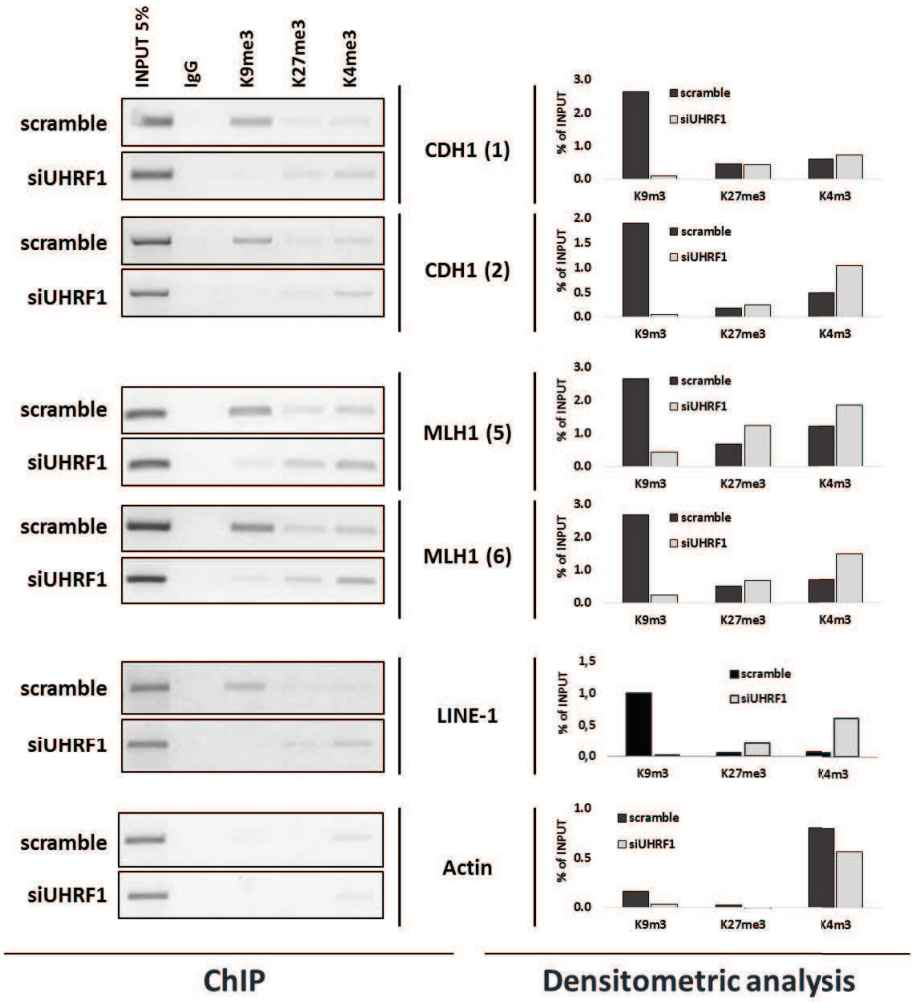


Figure 4.14. ChIP analysis of histone PTMs (H3K9me3, H3K27me3, and H3K4me3) on LINE-1 5'UTR, MLH1 and CDH1 promoter regions. Left panels: PCR of the analyzed promoters. Actin gene was used as control. Number in parentheses, close to gene names, referred to distinct primer sets. Right panels: Densitometric analysis of detected PCR bands.

We observed a mild increase of H3K4me3 on CDH1 and MLH1 promoter regions, and a considerable gain on LINE-1 sequences, probably as a consequence of the H3K9me3 loss, in accord with the activity of some H3K4 methyl-transferases, like PRDM9 [Wu H, et al.; Cell Report, 2013]. In the control sample (scramble) we

observed a weak signal of H3K4me3 and the strong H3K9me3 presence. It is well-known that other SET domain HMTs, like ASH1, part of the Polycomb Group protein (PcG), shows the ability to methylate different substrates concurrently, like H3K4, H3K9, H3K36 and H4K20 [Beisel C, et al.; Nature 2002]. ASH1, together with TRX-N-terminal, part of the transcriptional activation Trithorax protein group (TrxG), are not incompatible with repressive marks, like H3K9me3, or H3K27me3. Therefore, in our model, the presence of H3K27me3 (permanent or mild increased, regardless of UHRF1 KD), laid by EZH2, was consistent with a condition, reported in literature, called ''balanced'' state, in which PcG and TrxG proteins act simultaneously or at least in rapid alternation from active to repressive chromatin state, in concert with other positive and/or negative epigenetic regulators [Schwartz YB, et al.; PLOS Genetics, 2010].

It was worthy of note that LINE-1 sequences showed a slightly different chromatin profile. In control samples, they showed a strong presence of H3K9me3, in relation to UHRF1 and hyper-methylated DNA (very similar to normal peripheral blood cells [Baba Y, et al.; Mol. Cancer, 2010]). UHRF1 depletion caused strong alteration of their chromatin configuration: in fact, on one side there was a considerable gain of H3K4me3 that could lead to an increased transcriptional accessibility, and on the other side, the gain of mild H3K27me3 attested a sort of compensation repressive mechanism leading to a not complete re-expression of these sequences, as confirmed in qPCR analysis (data not shown) and RNA-seq data (see 4.4.2 paragraph below for details). Eventually, Actin promoter, as expected, was mainly characterized by H3K4me3 presence.

Recapitulating all these data, we observed that UHRF1 knock down induces a significantly decrease of DNA methylation both at global level (LINE-1) and at tumor suppressor gene promoters (MS-MLPA and specific Pyrosequencing of MLH1 and CDH1), without affecting the DNMTs levels (WB analysis). ChIP experiments demonstrated that UHRF1 KD impairs the DNMT1 recruitment to both specific gene promoters (MLH1, CDH1) and repetitive elements (LINE-1). Moreover, we observed that the lack of UHRF1 induces an increased binding of DNMT3B that can explain the residual DNA methylation. Regarding the hPTMs profile, UHRF1 knock down strongly decreased H3K9me3 presence on MLH1, CDH1 and LINE-1 and mild increased H3K4me3 on the same promoters. Despite all these epigenetic changes, toward a

more open transcriptional accessibility, we did not observe a re-expression of the analyzed genes, probably due to the H3K27me3 maintenance through the PCR2.

Altogether these intriguing data stimulated us to evaluate the impact of UHRF1 KD at genome wide level, both on DNA methylation and on transcriptome. To address this aim we performed a Reduced Representation Bisulfite Sequencing (RRB-seq) analysis, coupled to RNA-seq on RKO cells after siUHRF1.

4.4. Genome wide analysis of the effect of UHRF1 silencing on the methylome and the transcriptome

In order to evaluate the impact of UHRF1 KD on DNA methylation and on the transcriptome profile we performed a RRB-seq analysis, coupled to RNA-seq. This analysis, in collaboration with Prof. Alessandro Weisz (Lab of Molecular Medicine & Genomics, University of Salerno, Italy), would allow us to enlighten the involvement of UHRF1 on the loci/genes that are differential methylated and differential expressed.

4.4.1. UHRF1 KD decreases DNA methylation at genomic level in RKO cells

To evaluate the involvement of UHRF1 on the genome wide DNA methylation we prepared, from both UHRF1 silenced (siUHRF1) and control (scramble) RKO cells. 20 µg of DNA each replicate, and three biological independent replicates were used for the RRB-seq analysis.

RRB-sequencing analysis confirmed the overall decrease of the DNA methylation observed by pyrosequencing analysis upon UHRF1 silencing (Figure 4.15). In particular, we observed that the Differential Methylated CpGs (DM CpGs) are about 33.3% (1,584,975 CpGs) of the total identified CpGs, assessed around 4.8 million per each sample (with RC>10). As expected, 99.92% (1,583,862 CpGs) of the DM CpGs identified are hypo-methylated (Figure 4.15b, table upper right). Furthermore, we analyzed the percentage of the methylation detected in each cytosine in a CpG context (Figure 4.15a, and histogram in 4.15b middle-bottom). All the control samples showed a similar distribution of the total CpGs between the “fully methylated” status (an average of 26% of identified CpG) and the “not-methylated” status (an average of 27% of identified CpGs), in a sort of “bimodal distribution”. All the siUHRF1 samples showed a strong de-methylation, as already reported by raw numbers in the table of Figure 4.15b, but interestingly the hypo-methylation regarded mainly the CpGs that resulted fully methylated in the controls. Surprisingly, only less than 3% of CpGs showed a complete de-methylation to a “not-methylated” status, even if both UHRF1 and DNMT1 were de-localized in UHRF1-depleted cells (Figure 4.13). Thus, the majority of the “fully methylated” CpGs resulted not homogeneously demethylated giving rise a heterogeneous methylation status of the silenced population. This data was in accord with the pyrosequencing analysis of

LINE1, CDH1 and MLH1, in which, upon UHRF1 depletion, DNA methylation levels assessed to roughly the half of initial values. An explanation to this residual DNA methylation could reside in the enhanced localization of DNMT3B both at global level and on gene promoters, as observed in our ChIP experiment in Figure 4.13, in accordance with BRAF mutation in RKO cells [Jeong S, et al.; Mol Cell Biol., 2009] [Fang M, et al.; Mol. Cell, 2014].

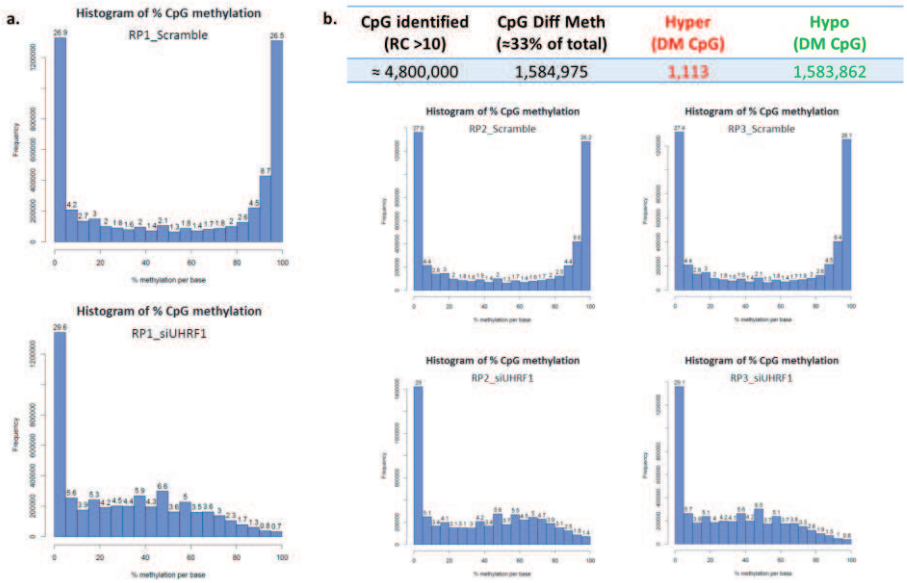


Figure 4.15. a) Histogram of the % CpG methylation. Control sample (scramble, upper plot) versus silenced UHRF1 (siUHRF1, bottom plot) shows the strong de-methylation of the "fully methylated" CpG in RKO cells. b) The table in the upper part reports raw number of CpG identified. In the middle-bottom part are reported all the other histogram of the independent biological replicates.

In order to classify better the distribution of identified CpGs and Differential Methylated CpGs along the genome we clustered them into well-known categories (Figure 4.16): on one side the distribution related to the CpGs organization (Shore, CpG island, Others), and on the other side the distribution related to annotated organization of the genes (Promoter, Exon, Intron, Intergenic). Since RRBS analysis have the peculiarity to enrich the CpG islands (CGI), we observed, as expected, that the 51% of the identified CpGs were localized at CGI regions, and the 41% of CpG are found into promoters. Nevertheless, focusing only on the Differential methylated (DM) CpGs, they are distributed equally all over the genome. In fact, even if the identified CpGs fell in CGI and promoter regions, the majority of DM CpGs are located in not-CGI (Others) sites, as intergenic or intronic regions.

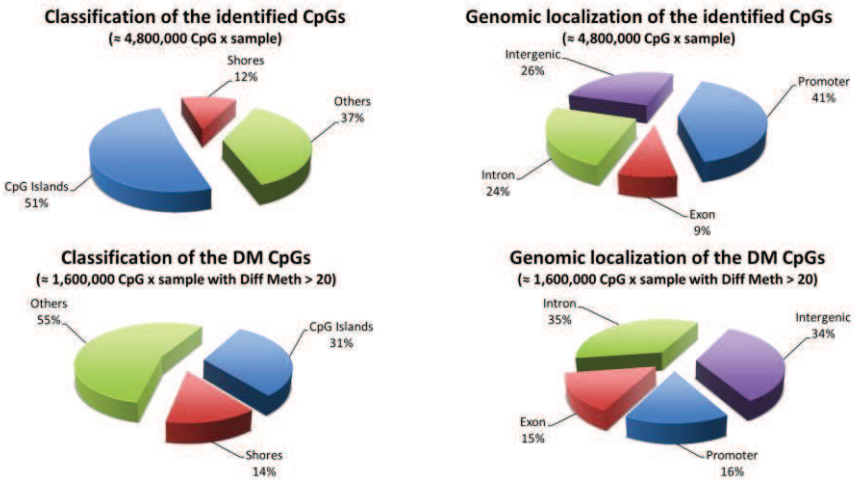


Figure 4.16. Pie Charts of classification and genomic distribution of CpGs identified in RRBS on RKO cells. Categories are related to CpG organization (Shore, CpG island, Others), and to the gene annotation (Promoter, Exon, Intron, Intergenic).

These data were in accord with the role of UHRF1, together with DNMT1, in the DNA methylation maintenance. Moreover, it has been shown that UHRF1 mainly localized at the heterochromatic regions characterized by high levels of DNA methylation. Our data showed that UHRF1 depletion reduced the DNA methylation of the "fully methylated" regions, generally representative of heterochromatic regions, such as intergenic regions.

However, we observed an important percentage of DM CpGs located in the promoter regions (16%) and in body of genes (Exon 15% + Intron 35%). Therefore, since the analyzed samples derived from a colorectal cancer cell lines, well-known to show an aberrant hyper-methylation of CGIs and promoters, as demonstrated before (by MS-MLPA and Pyrosequencing analysis), we performed a preliminary bioinformatic analysis in order to investigate the levels of the promoter methylation of CDH1 and MLH1. In particular, we assessed the methylation values of respectively 4 and 5 CpG of CDH1 and MLH1 promoters, in the same regions investigated previously by MS-MLPA and pyrosequencing analysis. The analysis of CDH1 showed a de-methylation of 39% among the 4 analyzed CpGs (Figure 4.17). MLH1 showed a milder de-methylation (29%), compared to previous analysis (Figures 4.10 and 4.9b).

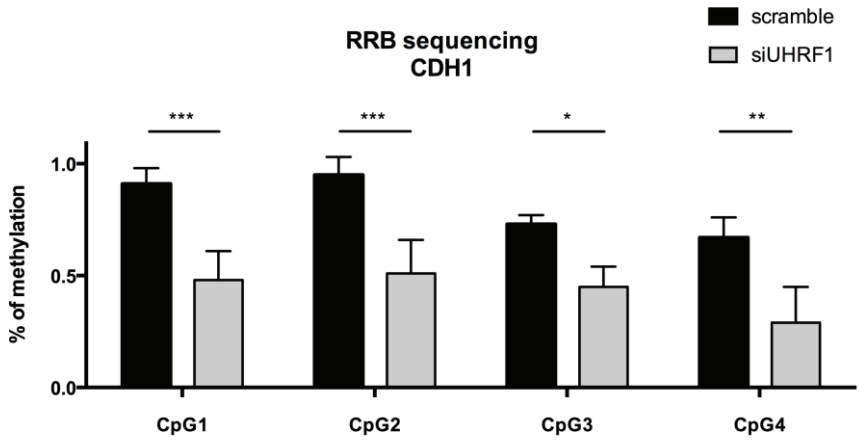


Figure 4.17. RRB-seq. Percentage of methylation of 4 CpGs on CDH1 promoter, as average of 3 independent biological replicates. The values are normalized to the respective coverage of each cytosine detected in the RRB-sequencing. Statistical significance was calculated by Anova test. (*) means p < 0.05, (**) means p < 0.002.

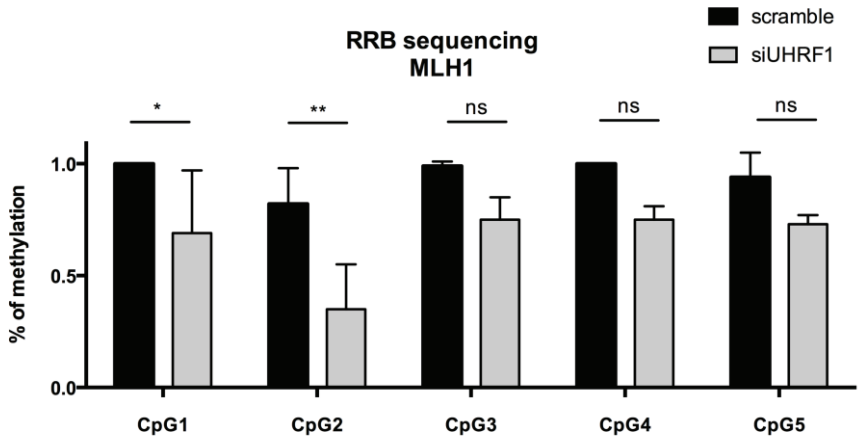


Figure 4.18. RRB-seq. Percentage of methylation of 5 CpGs on MLH1 promoter, as average of 3 independent biological replicates. The values are normalized to the respective coverage of each cytosine detected in the RRB-sequencing. Statistical significance was calculated by Anova test. (*) means p < 0.05, (**) means p < 0.005, (***) means p < 0.001.

We investigated the LINE-1 family in the genome datasets, and we observed a consistent decrease in DNA methylation (data not shown), as expected relying on the fact that 99.92% of CpGs are hypo-methylated.

Since we observed this strong effect on DNA methylation we decided to investigate better the role of this loss of DNA methylation on the transcriptome. To address this aim we performed an RNA-seq analysis in RKO cells (scramble and siUHRF1).

4.4.2. UHRF1 silencing is not sufficient for the re-expression of CDH1 and MLH1 genes, and LINE-1 sequences in RKO cells

To evaluate the impact of the UHRF1 KD on the transcriptome, we perform a RNA-seq analysis in siUHRF1 RKO cells and relative control (scramble RKO cells). The RNA-seq results were used to both determine which pathways are affected by UHRF1 KD (see paragraph 4.4.3.) and to investigate specific gene expression.
In order to explore the expression of specific genes, we compared the RPKM values. These values have been already normalized for the length of the transcript and they can be used to directly compare gene expression. Nevertheless, the DeSeq approach used for the calculation of the differentially expressed genes doesn't allow to extract information from the genes that have 0 reads in one of the 2 samples, either siUHRF1 or control.

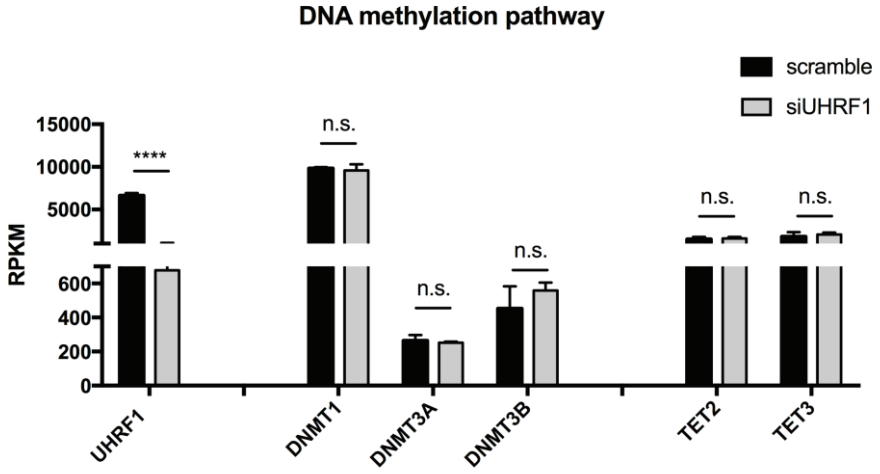


Figure 4.19a. RPKM values of the major proteins involved in the DNA methylation pathway. Statistical significance was calculated by Anova test. (****) means p < 0.0001.

First, we investigate the levels of the major enzymes involved in the DNA methylation and demethylation pathways (Figure 4.19a). The analysis showed that only UHRF1 is significantly decreased, mainly confirmed the qPCR results. Moreover, neither TET2 or TET3 were affected by UHRF1 KD. We could not analyze TET1 because in one or both of the two conditions TET1 had 0 reads. These results support the idea the changes in the DNA methylation profile are primarily induced by a reshuffling (i.e. DNMT1 delocalization and DNMT3B repositioning, as observed by ChIP assay) of the methylator complexes and not due to an overall decrease of the apparatus.

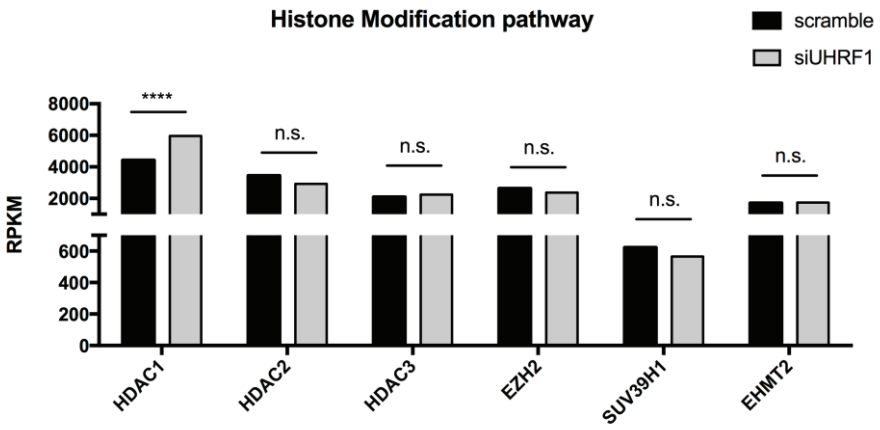


Figure 4.19b. RPKM values of the major proteins involved in the Histone modification pathway. Statistical significance was calculated by Anova test. (****) means p < 0.0001.

Then, we decided to investigate the level of the enzymes involved in the histone modification pathway (Figure 4.19b). The RNA-seq analysis showed that only the HDAC1 was significantly increased by the UHRF1 KD. Even if this data seemed to counteract the overall opening of the chromatin, none of the other analyzed factors were changed in the UHRF1 silenced cells. These results perfectly matched with the SILAC bulk analysis (see later in the Results) that showed that the UHRF1 KD did not significantly affect any histone modifications overall.
Furthermore, we checked the levels of CDH1 and MLH1. According to what we found by qPCR and WB, CDH1 did not have any reads in one or both the samples and therefore the DeSeq analysis did not show any results for it. MLH1 showed an average of 8 reads in total in both the siUHRF1 and the controls, with no significant changes between the two conditions (data not shown).

The evaluation of the differential expression of the LINE-1 sequences was more complicated. In fact, the algorithm used for the alignment automatically discharge the repetitive reads (i.e. reads that map in more than one locus on the genome). This is based on the assumption that the "real transcripts" should be uniquely mapped. This is not true for the LINE-1 since they are repetitive elements, and even if most of them resulted mutated, the level of the mutation is not enough to uniquely map them on a reference genome. This is not an easy bioinformatics problem to solve. At the moment, there are not available pipelines that can take this problem into account, and lower the threshold for the uniquely mapped reads will result in an

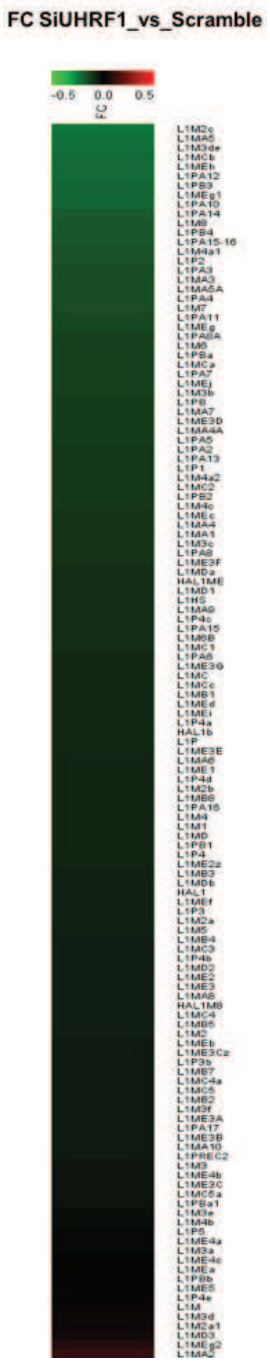


Figure 4.20. Heat map of the 131 identified LINE-1 sequences.

overall misalignment. Nevertheless, we were able to detect some of them by using a previously developed algorithm [Criscione SW, et al.; BMC Genomics., 2014]. We found 131 LINE-1 sequences with more than 1 read mapped (Figure 4.20). None of them showed a fold change greater or smaller than 0.5.

We can conclude that even if we observed great changes on both methylation and chromatin status, these changes were not sufficient to change the expression levels of these specific 131 LINE-1 sequences.

4.4.3. UHRF1 KD interferes with cell proliferation, cell growth, DNA damage and DNA repair

In order to evaluate the effect of UHRF1 depletion on the transcriptome, we performed a RNA-seq analysis on three independent biological replicates. After sequencing we follow a DESeq pipeline [Anders S, et al.; Genome Biol., 2010]. We obtained 1810 differential expressed transcripts (DE), corresponding to 512 genes up-regulated and 201 down-regulated with a Fold-Change (FC) higher or lower that 1.5 (Figure 4.21).

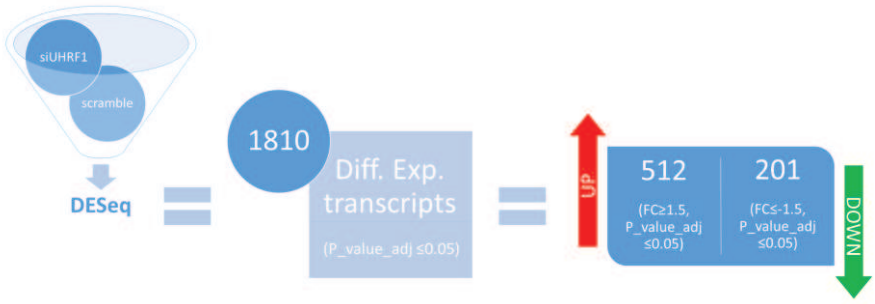


Figure 4.21. DESeq bioinformatic analysis of RNA-seq data. The number of differential expressed transcript (DE genes) are assessed with a p-value adjusted lower of 0.05. UP- and DOWN-regulated genes are found with a Fold Change (FC) cut-off of 1.5.

The fold change of the up-regulated genes spanned from 1.5 to 222.83 (corresponding to a log2FC between 0.58 and 7.80). The FC of the downregulated genes spanned from -1.5 to -3.65 FC (corresponding to a log2FC between -0.58 and -1.87). We analyzed RNA-seq data by Ingenuity Pathway Analysis (IPA), in the attempt to identify the main pathways in which UHRF1 was involved. IPA showed that UHRF1 KD interfered with several important pathways, among others cell proliferation, cell growth, DNA damage and DNA repair, cell cycle and Wnt signaling (Figure 4.22). Nevertheless, we observed a coherent induction of the whole pathway only in the TFGβ signaling pathway. Surprisingly, upon UHRF1 depletion, we observed an increase in genes belonging to the cell growth and proliferation ontological cluster.

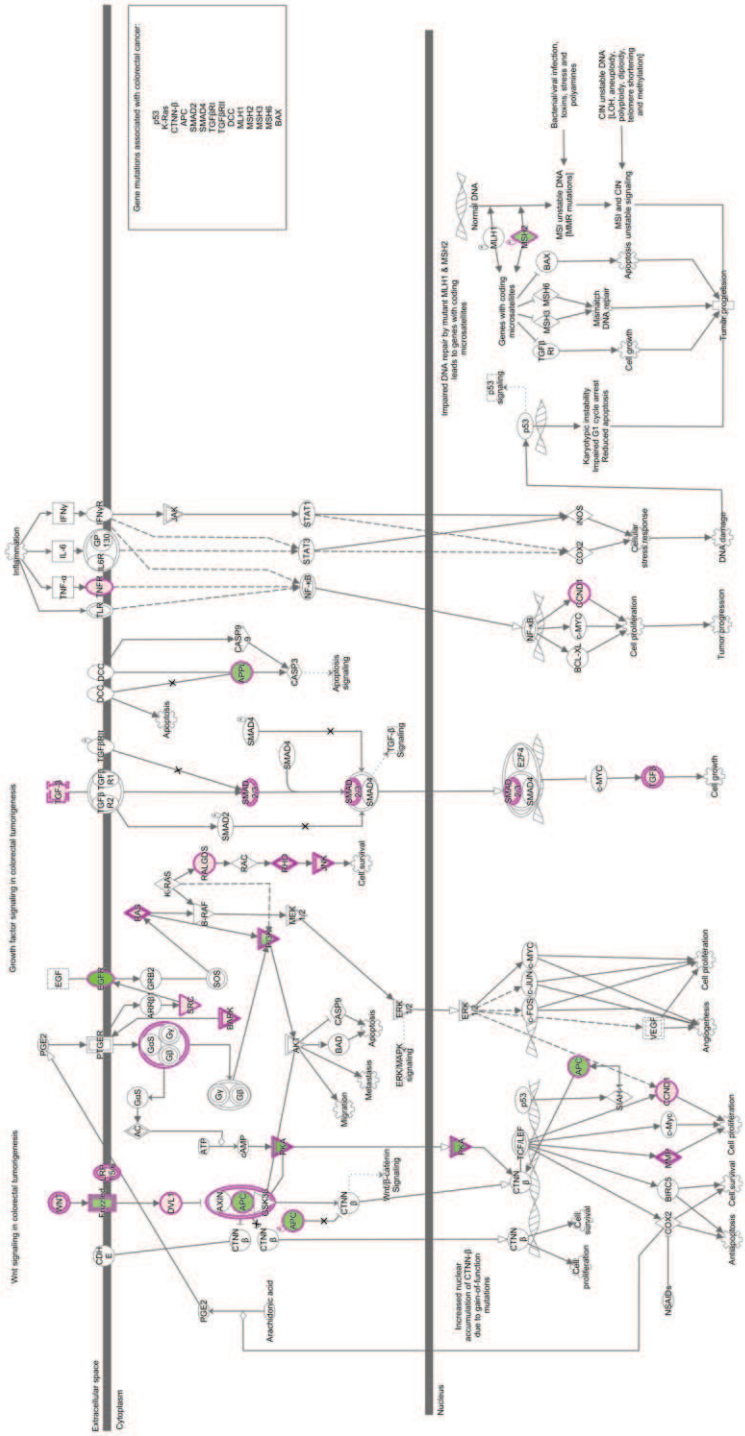


Figure 4.22. Ingenuity Pathway Analysis (IPA) of cell growth & proliferation pathways related to colorectal cancer. In read are delineated the up-regulated genes and in green the down-regulated ones, upon UHRF1 KD in RKO cells.

However, we observed a downregulation in the genes belonging to the cell cycle ontological cluster. These data agreed with our previous reports in which we showed that UHRF1 decreased cell proliferation [Babbio F, et al.; Oncogene, 2012]. Altogether these antithetical observations could be explained, at least in part, by the fact that UHRF1 is not a transcription factor or a repressor with specific target sequences. In fact, at the moment, there is no a consensus sequence that could drive the localization of UHRF1 to the target foci, rather it is able to recognize a plethora of epigenetic signatures, from DNA modifications to histone PTMs. It is plausible that its recruiting is cell-type and context dependent. Therefore, UHRF1, related to the possible epigenetic conditions, could be localized on different targets, activating or inhibiting several pathways.

Since the UHRF1 KD generated a clear strong de-methylation and the majority of the DE genes were up-regulated, we decided to investigate better the possible correlation between differential methylated and differential expressed genes.

4.4.4. DNA de-methylation correlates to transcriptional upregulation of specific loci

The RRBS and RNA-seq data prompted us to investigate the possible relationship between the genome wide methylation profile and the transcriptome. In order to enlighten the differential methylated and differential expressed (DM and DE) loci/genes, we integrated the two datasets obtained by RRBS and RNA-seq, coupling the genome wide methylation profile to the transcriptome.
In particular, we intersected the DM CpGs (with cut-off 20 in differential methylated value, and q-value < 0.05) with the DE transcripts (with cut-off 0.5 log2(FC), and p-value-adjust < 0.05). This analysis showed the presence of a cluster of differential hypo-methylated (Hypo-DM) CpGs harbored in the up-regulated (Up-DE) transcripts (Figure 4.23a, in green). The Hypo-DM CpGs, linked to Up-DE transcripts, were distributed in different genomic annotated loci (Figure 4.23b), with preference to “Intron” and “Promoter-TSS”. However even if this finding was interesting, we noted that this analysis is un-related to transcript integrity, or alternative transcript of the same gene, or antisense-transcript harbored in the same loci. This resulted in a redundant number of genes involved in each annotated locus (Figure 4.23b, number in brackets in the table), due to the possibility that one CpG has an Ensembl genomic

annotation related to a specific “Promoter-TSS” of one transcript, but it has also can be annotated in the “Intron” of a second transcript that nests the first one.
Moreover, the Hypo-DM CpGs, linked to Up-DE transcripts (47,222; Hypo_Up, green in scatter-plot and table of Figure 4.23a/b) represented less than 3% of total hypo-methylated CpG identified in RRBS (1,583,862; Figure 4.15b, upper table). Intriguingly, this observation resulted in the fact that the 97% of hypo-methylated CpGs identified are not related to transcript regulation (96.62% including the Hypo_Down, blu in scatter-plot and table of Figure 4.23a/b), in perfect accord with our previously reported finding about the hypo-methylation of CDH1, MLH1 and LINE-1 loci, that not present any important transcriptional alteration. It was not clarified if Hypo_Up loci are located to particular chromatinic cluster that allows the up-regulation of the related genes. Further investigations are necessary also in order to spanning the Hypo-DM CpGs on the Up-DE transcripts to better correlate and to statistically calculate this intersection clustering.

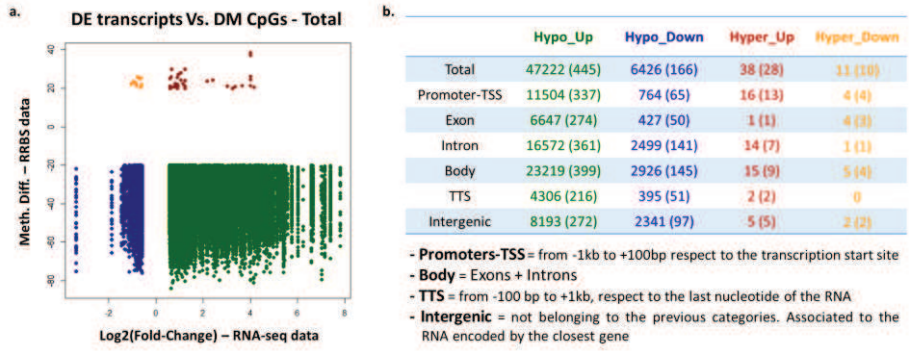


Figure 4.23. RRBS and RNA-seq intersection analysis. a) Scatter-plot shows the distribution of CpGs differentially methylated (DM), comprehended in differentially expressed (DE) transcripts. b) The table shows the number of differently methylated CpGs found in the different genomic annotation, as described bottom. In parenthesis are indicated the number of genes involved, that harbor DM CpGs in own transcripts.

The genome-wide analysis confirmed the overall DNA hypomethylation mediated by UHRF1 silencing, both on LINE-1 and on specific gene promoters. Interestingly, about 99,9% of DM CpGs were hypo-methylated, underlyng the pivotal role of UHRF1 in the DNA methylation maintenance. UHRF1 KD interfered with several important pathways, among others cell proliferation, cell growth, DNA damage and repair, cell cycle and Wnt signalling. Coupling the DM CpGs with the DE transcripts we showed that, even if there is cluster of genes that are hypo-methylated and up-regulated, these transcripts represented only the 3% of the hypo-methylated CpGs. This data

was perfectly consistent with our previous finding about CDH1, MLH1 and LINE-1 loci, that showed a strong hypo-methylation but neither important transcriptional alteration, probably due to their epigenetic chromatin profile, as observed by ChIP. However, these analyses were not sufficient to elucidate the role of the DNA methylation in the regulation of these specific loci/genes, altered upon UHRF1 silencing. Deeper analysis are now ongoing.

4.5. Investigate the effect of UHRF1 silencing on ubiquitination of histone tails

The data obtained by ChIP assay together with previously published evidences on UHRF1 ubiquitination activity [Qin W., et al; Cell Research, 2015] led us to hypothesize that the chromatin rearrangements, observed upon UHRF1 KD, could be mediated by the lack of H3K18ub seated by UHRF1 RING. In the attempt to verify this hypothesis and to better elucidate the involvement of UHRF1 in specific ubiquitination of histone tails, we performed a SILAC LC-MS/MS on RKO cells after siUHRF1. Eventually, the LC-MS/MS (SILAC) analysis would allow us to couple ubiquitination profile of hPTMs to genome wide methylation and transcriptome profile.

4.5.1. Design and production of a monoclonal antibody against the H3K18ub

As mentioned before, we recently published that UHRF1 ubiquitinates K18 of histone H3 by its RING domain, harboring a well-known E3 ubiquitin ligase activity [Citterio E, et al.; Mol. Cell. Biol., 2004] [Jenkins Y, et al.; Mol. Biol. of Cell, 2005] [Nishiyama A, et al.; Nature, 2013]. Furthermore, H3K18ub is essential for the maintenance of DNA methylation, enabling methylation activity of DNMT1, through interaction with its ubiquitin interacting motif (UIM) [Qin W., et al; Cell Research, 2015]. Altogether these evidences prompted us to design and produce a monoclonal antibody toward the monoubiquitinated lysine 18 of histone H3. This part of the project was carried out in collaboration with Dr. Mario Cinquanta (Cogentech, Campus IFOM-IEO, Milano).

We immunized 4 immuno-reactive mice with a synthetic peptide harboring the branched epitope of ubiquitinated lysine 18 on H3 histone tail (Figure 4.24a). The sera collected from mice bleeding samples, after 3 immunizations (Figure 4.24c) and 5 immunizations (Figure 4.24d) were analyzed by ELISA assay. We tested the reactivity of each serum against the same immunization peptide and, as negative control, versus an identical H3 tail peptide synthetized without the branching for K18ub. All four animals responded positively to immunization, comparing immune-reactivity of the sera against H3K18ub peptide (Figure 4.24, c and d) with the pre-immune sera collected from the same mouse (Figure 4.24b). Furthermore, we observed a clear specific antibody response to H3K18ub peptide, compared to

control H3 tail peptide, showing a good performance in the titration of the diluted sera. In the second bleeding after 5 immunizations, we noted that the mouse 1 started to show a considerable response to control H3 tail peptide not ubiquitinated (Figure 4.24d).

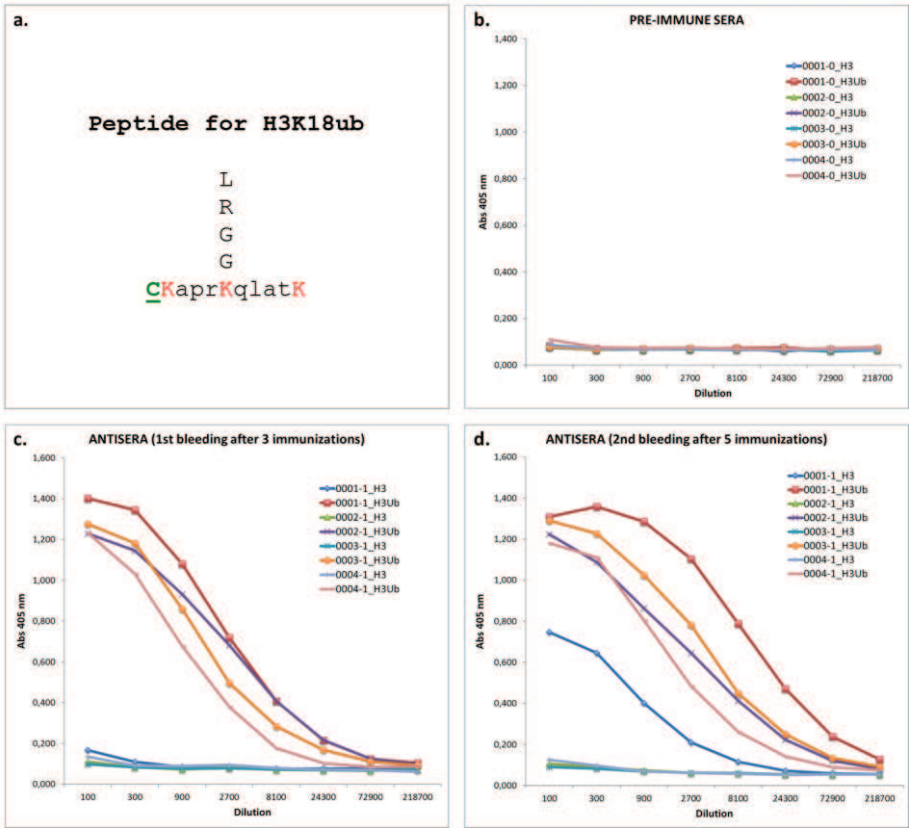


Figure 4.24. a) Designed synthetic peptide harboring the branched epitope of ubiquitinated lysine 18 on H3 histone tail. b,c,d) ELISA assay representing the titration of pre-immune sera (b), 1st bleeding sera after 3 immunizations (c) and 2nd bleeding sera after 5 immunizations (d), collected from the 4 mice immunized with H3K18ub peptide (a).

We performed a second screening on total histone acids extract, from RKO and PC3 cell lines, by Western Blot analysis. PC3 cells showed some preliminary interesting evidence in the Ubiquitination profile in total histone acids extract analyzed by classical LC-MS/MS (data not shown). WB analysis showed that the sera (re-suspended in TBS-Tween-milk 4%, with 1:50 dilution), respectively from mouse 3 and 4, and from mouse 1 and 2, are very similar in the immune response profile (Figure 4.25). In particular, the sera form mouse 2, 3 and 4 showed a band at the expected

molecular weight corresponding to mono-ubiquitination of histone H3 (around 25KDa), especially on the PC3 histone extracts.

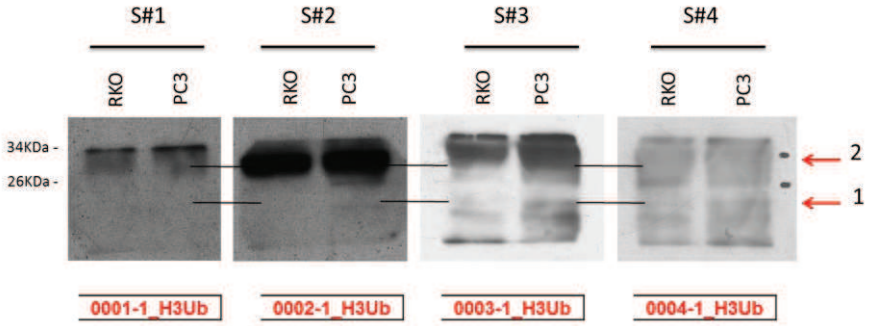


Figure 4.25. Western Blot analysis of total histone acids extract, from RKO and PC3 cells. As primary blotted antibodies we used the sera from mice 1 (S#1), 2 (S#2), 3 (S#3) and 4 (S#4), re-suspended in TBS-Tween-milk 4%, with 1:50 dilution. The expected molecular weight corresponding to mono-ubiquitination of histone H3 is around 25KDa.

After these first steps, we performed the generation of hybridoma cell lines secreting mAbs against H3K18ub, by fusion antibody-producing B cells, derived from the spleen of the immune-reactive mice (03 & 04 together), with a myeloma (B cell cancer). The hybridoma's media was screened by ELISA analysis in the attempt to detect either the best clones and eventual some trans-reaction against H3-unmodified or other ubiquitinated protein (i.e. EGFR-ub). In order to expand and to enrich cell lines secreting mAbs against H3K18ub, we performed sub-cloning process for the first four best hybridomas, identified by ELISA (data not shown). We compared the profile observed in the WB assay comparing the hybridomas signals, the expected molecular weight, and, the staining from commercial antibodies against anti-H3 and anti-Ubiquitin-Total. These analyses allowed us to identified four clones (TN13, TN12, TK7, TN0) that showed a good specificity (data not shown).

In order to obtain the best candidate for further analyses, we purified, by IgG affinity, these 4 monoclonal sera. Considering the concentration, the efficacy and the previous ELISA, the best sera was the clone TN13-7. This clone was used for subsequent experiments. In order to test the specificity, we performed a histone extraction from HEK293T cells overexpressing H3-FLAG wild type and H3_FLAG K18A, bearing a point mutation in the K18 (Figure 4.26, upper panel). The goal in this experiment was to identify an alternative profile in WB analysis comparing H3-FLAG WT with H3_FLAG K18A histones. Unfortunately, the presence of endogenous H3 (wild type) made indistinguishable the profile between the two samples (Figure 4.26,

upper panel, on right, lane 2 and 3). Nevertheless, comparing the anti-H3 and the anti-H3K18ub (subclone TN13-7), we were able to identify some corresponding protein bands, which probably represent mono-/di-/tri-ubiquitinated forms of histone H3. The serum recognized several protein bands, which probably correspond to the histone H3 variants and/or H3 poly-ubiquitinated or SUMOylated forms, containing at least one H3K18ub branching. Furthermore, we observed a strong discrepancy in the ubiquitination profile of histone acid extraction derived from the different cell lines, that showed an anti-H3K18ub (subclone TN13-7) profile much more simple in HeLa (data not shown) compared to RKO cells (see an example in SILAC experiment below).

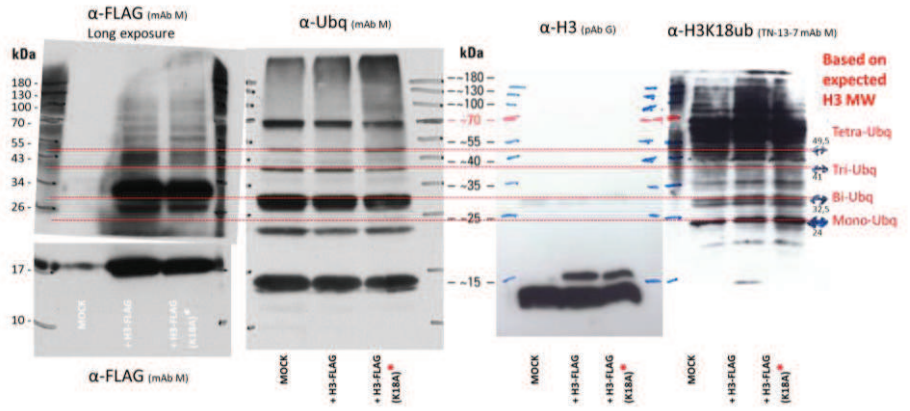


Figure 4.26. WB analyses for testing specificity of H3K18ub (clone TN13-7) on histone acid extraction. Histone extractions were performed on HEK293T cells overexpressing H3-FLAG Wild Type (lane 2) and H3_FLAG K18A (point mutation, lane 3).

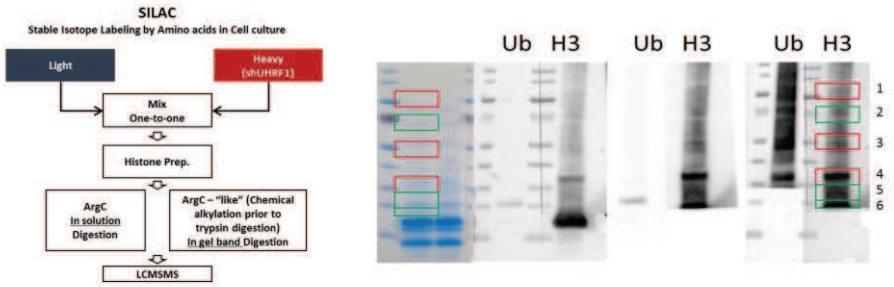
We tried to perform also an immunoprecipitation against H3-FLAG, with the specific resin Sigma, but the elution of the beads (containing α-FLAG Ab) induced the disaggregation of the light chain of the antibody that migrate, as expected, at 25KDa, as well as H3K18ub, causing the failure of this experiment (data not shown). To better characterize the specificity of our antibody α-H3K18ub and avoid the "contamination" of the endogenous H3, we are now designing the specific point mutations using the CRISPR/Cas9 technology.

4.5.2. UHRF1 KD induces alteration in H3 N-term-Tail ubiquitination

The LC-MS/MS (SILAC) analysis allowed us to analyze the proteomic profile of hPTMs related, independent or antithetical to UHRF1, with particular regard to ubiquitination of histone tails. In the attempt to observe, in a quantitative manner, the differences in histone proteomic profiles, we chose the SILAC (Stable Isotope Labeling by/with Amino acids in Cell culture) approach coupled with a LC-MS/MS, in collaboration with Dr. Tiziana Bonaldi (IEO, Campus IFOM-IEO, Milano). This technique gave us the advantage to label two samples (scramble and siUHRF1) with different and specific amino acids (lysine and/or arginine), non-radioactive stable isotopes. In this way, the two samples, homogeneously pooled in ratio 1 to 1, can be analyzed in a single mass-spectrometry analysis, obtaining a quantitative LC-MS/MS, that allowed us to discriminate the different samples comparing the shifted peaks observed for control (Light isotopes) and siUHRF1 (Heavy isotopes).

In order to prepare the biological sample for quantitative SILAC mass-spectrometry analysis, we incorporated stable isotope amino-acids specific for hPTMs profile analysis, by growing RKO cells in SILAC medium (K-0; R-10). We evaluate the complete incorporation through LC-MS/MS analysis on total extracts from RKO cells at the eighth population doublings in SILAC medium. The ratio between Heavy and Light peptides analyzed suggests that the incorporation obtained is at least 97,4% (data not shown). We proceeded in the preparation of sample for quantitative SILAC mass-spectrometry analysis, performing a histone acid extract relying on the yield of previous experiments. In the attempt to obtain at least 300µg of histone proteins for all the subsequent experiments, we performed an acid extraction on a pool of 140 millions (70 millions Heavy & 70 millions Light cultured in SILAC medium) of RKO cells mixed with ratio 1:1 (Figure 4.27, upper left flowchart). We obtain 750µg of histone proteins. For the first experimental analysis, we perform a SDS-PAGE on a gradient precast acrylamide gel to ensure the better separation of protein between 15 and 35 KDa. We loaded 15µg of histones per lane in order to identify the protein bands that coincide each other (analyzed with antibodies against H3 and Ubq-total) (Figure 4.27, upper right, WB and Coomassie). Subsequently each selected band (cut out in the Coomassie stained gel) was analyzed in LC-MS/MS following the well-established procedure of our collaborators [Soldi M., et al; Int. J. Mol. Sci., 2013]. Using the same protocol, we prepared an “in solution digestion”, and an “in gel bands digestion” using 10µg of total histones, to identify the overall hPTMs profile of RKO cells, after

siUHRF1 (Figure 4.27, upper left flowchart). Regarding the "in gel bands digestion" analysis we observed a strong alteration of H3 tails ubiquitination. In particular, we identified two ubiquitinated residues, H3K18ub and H3K23ub, in accordance with published data [Nishiyama A, et al.; Nature, 2013] [Qin W., et al; Cell Research, 2015]. Interestingly, in the band #6, correspondent to the molecular weight of the mono-ubiquitinated H3 (25KDa), we identified H3K18ub and H3K23ub that are present alternatively in the two different peptides. Furthermore, in the band #4, correspondent to bi-ubiquitinated H3 (35KDa), we observed only H3K18ub that realistically would be bi-ubiquitinated as a branched poli-ubiquitination (Figure 4.27, middle table). The SILAC quantitative analysis of identified peptides allowed us to observe a RATIO between the same peptides derived from scramble (Light) and siUHRF1 (Heavy) samples. Surprisingly this quantitative analysis showed no changes (RATIO H/L = 1.004) in H3K18ub in the mono-ubiquitinated peptides (band #6). On the other hand, in the bi-ubiquitinated peptides (band #4) we observed a strong increasing in the H3K18ub (RATIO H/L = 2.496), upon UHRF1 silencing. Meanwhile we discovered a decreasing (RATIO H/L = 0.760) in H3K23ub in the mono-ubiquitinated peptides (band #6) (Figure 4.27, bottom table).



Identified Ubiquitinated-peptides								
Sequence	Modified sequence	Modified Sites	Protein Names	Charge	m/z	Retention time	Score	Bands
KQLATKAAR	K(GG)QLATKAAR	K18	Histone H3	3	382.565805	20.595	48.091	#6
KQLATKAAR	KQLATK(GG)AAR	K23	Histone H3	2	573.345069	20.924	35.541	
KQLATKAAR	K(GG)QLATKAAR	K18	Histone H3	2	573.345069	20.671	36.424	#4

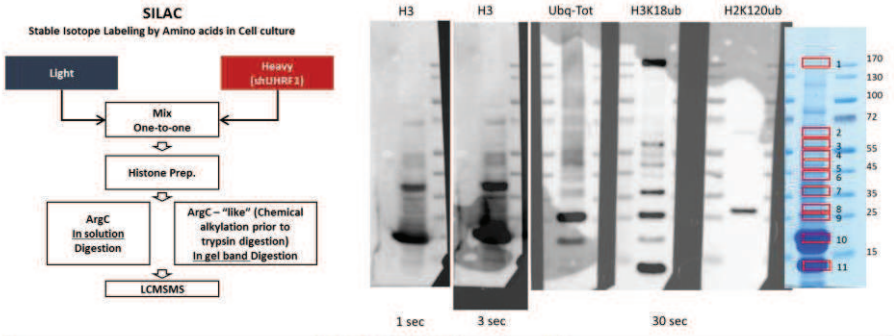
Quantitative analysis of identified peptides						
Sequence	Modified sequence	Modified Sites	Protein Names	Score	Ratio H/L	Bands
KQLATKAAR	_K(GG)QLATKAAR_	K18	Histone H3	48.091	1.004	#6
KQLATKAAR	_KQLATK(GG)AAR_	K23	Histone H3	35.541	0.760	
KQLATKAAR	_K(GG)QLATKAAR_	K18	Histone H3	36.424	2.496	#4

Figure 4.27. Quantitative mass-spectrometry analysis LC-MS/MS (SILAC), performed on RKO cell (scramble/siUHRF1), showed strong alteration in the ubiquitinated profile of H3 histone tails, especially in band #6 and #4, respectively correspondent to the molecular weight of the mono- and the bi-ubiquitinated H3.

In the “in solution digestion” analysis, we observed, as expected, mild or not changes in the global hPTMs profile (e.g H3K4me3, H3K9me3, H3K27me3), and we did not found ubiquitinated peptides detectable at global level. However, it is not possible to compare the results coming from the “in gel bands analysis” with the results coming from the “in solution”. In fact, the first has a higher sensibility compared to the second that is mainly used to compare great changes in the bulk amount of histones. The direct analysis of the bands allows the detect even low represented and rare histone modifications that can have a role in the control of the local organization of the chromatin. In our specific case the analysis of the bands allowed us to detect changes in the H3K18ub, while the bulk was not able to detect ubiquitinations at all.

In order to better investigate these data, we performed a second analysis, loading 30µg of histone per lane. The specific goal of this second experimental analysis was to verify the specificity of our new antibodies against H3K18ub, and to find other ubiquitinated residues that could be low represented in the sample (e.g. H2BK120ub) (Figure 4.28). This second mass-spectrometry analysis LC-MS/MS (SILAC), confirmed the strong alteration in the ubiquitinated profile of H3 N-terminal tails. Moreover, the identified bands by our new antibody against H3K18ub (clone TN13-7) showed a high-fidelity recognition of the H3 ubiquitinated tails (bands 7 and 8-9, corresponding to mono and bi-ubiquitination, at 25KDa & 35KDa) (Figure 4.28, middle table). In fact, these bands, analyzed by mass-spect, showed to contain mainly the K18 mono or bi-ubiquitinated. In the band #8 e #9 was also detected the residues K23ub, belonging to a different pool of histones H3, mono-ubiquitinated in K23, that migrate, as expected, in the same molecular weight (25KDa) during electrophoresis. Among the different bands analyzed (i.e. #1, #3, #4, #5, #6, #7, #9, #10, #11), corresponding to a detected signal in the blot with H3K18ub antibody, only the bands #4 and #11 showed a trans-reaction with histone H1 ubiquitinated. The band #10 probably showed a mild trans-reaction with unmodified-H3, and the band #1 with poly-ubiquitinated histone H3. Nevertheless, all of them were negligible, due to the totally different molecular weight. Last but not the least, the bands #2 and #8-9, corresponding to a detected signal in the blot with the commercial available H2BK120ub antibody (Active Motif), did not contain a detectable amount of ubiquitinated peptide of H2B, although the bands were in the expected molecular weight. In fact, it was detected the poly ubiquitination of

H2BK24ub/K25ub in band #2, and mono-ubq of H2A.ZK126 in band #9, that migrate in the same molecular weight during electrophoresis. These findings seemed to argue against the specificity of the H2BK120ub antibody, but these results could reflect high magnification of the signal due the WB. In fact, we observed a similar result also for the H3K18ub antibody.



Identified Ubiquitinated-peptides

aa Sequence	Modified sequence	Modified site	Gene Names	Protein Names	Band	Charge	m/z	Retention time
VAKSPAK	_VAK(ac)SPAK(gl)_	K182(ac)/K186(Ub)	HIST1H1A	Histone H1.1	2	2	428.748	40.63
AVTKVQKK	AVTK(d3)VQK(gl)K(gl)_	K24(Ub)/K25(Ub)	HIST2H2BF	Histone H2B		2	587.8527	41.28
KSSKNPK	K(ac)SSKNPK(gl)_	K164(ac)/K170(Ub)	HIST1H1A	Histone H1.1	4	2	472.7616	17.18
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3	7	2	573.3451	23.77
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3		3	382.5658	23.77
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3	8	3	382.5658	23.89
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3		2	573.3451	23.89
KQLATKAAR	_K(gl)QLATK(ac)AAR_	K18(Ub)	HIST2H3PS2	Histone H3		3	381.5595	23.92
KQLATKAAR	_K(d3)QLATK(gl)AAR_	K23(Ub)	HIST2H3PS2	Histone H3		2	573.3451	24.24
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3	9	3	382.5658	23.82
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3		2	573.3451	23.82
KQLATKAAR	_K(d3)QLATK(gl)AAR_	K23(Ub)	HIST2H3PS2	Histone H3		2	573.3451	24.17
KGQQKTV	_K(d3)GQQK(gl)TV_	K126(Ub)	H2AFZ	Histone H2A.Z		2	474.271	20.38
GLTLAALKK	_GLTLAALK(gl)K(di)_	K80(Ub)K81(me2)	H1FNT	Testis-specific H1 histone	11	2	528.8424	48.08

Quantitative analysis of identified peptides

aa Sequence	Modified sequence	Modified site	Gene Names	Protein Names	Band	Charge	m/z	Retention time	Score	Ratio H/L
VAKSPAK	_VAK(ac)SPAK(gl)_	K182(ac)/K186(Ub)	HIST1H1A	Histone H1.1	2	2	428.748	40.63	40.103	0.991
AVTKVQKK	_AVTK(d3)VQK(gl)K(gl)_	K24(Ub)/K25(Ub)	HIST2H2BF	Histone H2B		2	587.8527	41.28	38.792	1.087
KSSKNPK	_K(ac)SSKNPK(gl)_	K164(ac)/K170(Ub)	HIST1H1A	Histone H1.1	4	2	472.7616	17.18	40.103	1.001
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3	7	3	382.5658	23.77	58.433	2.539
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3		2	573.3451	23.89	62.463	0.962
KQLATKAAR	_K(gl)QLATK(ac)AAR_	K18(Ub)	HIST2H3PS2	Histone H3	8	3	381.5595	23.92	52.579	1.164
KQLATKAAR	_K(d3)QLATK(gl)AAR_	K23(Ub)	HIST2H3PS2	Histone H3		2	573.3451	24.24	44.318	0.772
KQLATKAAR	_K(gl)QLATK(d3)AAR_	K18(Ub)	HIST2H3PS2	Histone H3		2	573.3451	23.82	59.908	0.884
KQLATKAAR	_K(d3)QLATK(gl)AAR_	K23(Ub)	HIST2H3PS2	Histone H3	9	2	573.3451	24.17	37.556	0.630
KGQQKTV	_K(d3)GQQK(gl)TV_	K126(Ub)	H2AFZ	Histone H2A.Z		2	474.271	20.38	38.972	0.911
GLTLAALKK	_GLTLAALK(gl)K(di)_	K80(Ub)K81(me2)	H1FNT	Testis-specific H1 histone	11	2	528.8424	48.08	46.129	1.011

Figure 4.28. Quantitative mass-spectrometry analysis LC-MS/MS (SILAC), performed on RKO cell (scramble/siUHRF1), showed strong alteration in the ubiquitinated profile of H3 N-terminal tails. Moreover, the identified bands by our new antibody against H3K18ub (clone TN13-7) showed a high-fidelity recognition of the H3 ubiquitinated tails (bands 7 and 8-9, corresponding to mono and bi-ubiquitination at 25KDa & 35KDa).

The RATIO between the peptides (Figure 4.28, bottom table), derived from scramble (Light) and siUHRF1 (Heavy), showed the same trend of the first analysis (Figure 4.26). In fact, H3K18ub in the bi-ubiquitinated peptides (band #7) showed a strong increase (RATIO H/L = 2.539) and no, or mild changes in the mono-ubiquitinated peptides detected in the band #8 (RATIO H/L = 0.962, or 1.164 in the presence of

K23ac) and #9 (RATIO H/L = 0.884). H3K23ub in the mono-ubiquitinated peptides, band #8 and #9, showed respectively a consistent decrement, upon UHRF1 silencing (RATIO H/L = 0.772 and 0.630).

For the first time, we analyzed the role of UHRF1 in the modulation of histone ubiquitination in a tumor cell line. Differently to what observed in Embryonic Stem Cells [Qin W., et al; Cell Research, 2015], in RKO cells, the UHRF1 silencing induced a more complex pool of histone ubiquitination. All together these results shed a new light about the role of UHRF1 in the chromatin remodeling, in fact it is able to indirectly modulate the histone methylation profile by recruiting the histone methyltransferases on the specific loci, but also it directly alters the ubiquitination patterns affecting DNA methylation.

5. Discussion

Colorectal cancer (CRC) represents one of the major leading cause of cancer death [Ferlay J., et al.; Int. J Cancer., 2015]. In CRC, genetic and epigenetic alterations are tightly connected and not mutual exclusive events, as well as in other tumor types, cooperating in its development and progression. Interestingly, in colorectal cancer aberrant methylation events are more common than point mutations [Lao VV and Grady WM; Nat. Rev. Gastroenterol. Hepatol., 2011]. This is the case of a subclass of sporadic CRC tumors, that is characterized by peculiar molecular features, the microsatellite instability (MSI) and the hypermethylation of CpG islands (CIMP). In this subclass, the hypermethylated of the promoter of MHL1 gene led to the inactivation of the mismatch repair (MMR) mechanism [Weisenberger DJ, et al.; Nat. Genet., 2006]. Even if some mechanisms have been elucidated, most of the effects of the interaction between genetic mutations and epigenetic alterations on the patient's outcome are not clearly understood, due to their complexity [Wu C and Bekaii-Saab T.; Chem. Res. Pract., 2012].
In this context, this PhD project wants to shed light on the molecular involvement of UHRF1, as pivotal epigenetic actor in DNA methylation maintenance, in the modulation of the aberrant DNA methylation profile characterizing this subclass of CRCs. In fact, UHRF1 is a multi-domain "cross-talk" protein involved in epigenetic regulation [Bronner C, et al.; Biochem. Pharmac., 2013], coupling the preservation of histone-modification through the cell cycle with the maintenance of DNA methylation, orchestrating DNMT1 activity by deposition of H3K18ub [Qin W., et al; Cell Research, 2015] and ubiquitination of several lysines of H3 tail [Harrison JS, et al.; eLife, 2016]. UHRF1 is found over-expressed in numerous tumor types, and it is involved in gene silencing of several tumor suppressor genes (e.g. p16INK4A, MLH1, BRCA1, RB1, APC, CDH1, RARβ1, RARβ2, PSP94), mediating the heterochromatic switch through the recruitment of the several repressive proteins (DNMT1, HDAC1, Suv39H1, G9a) [Alhosin M, et al.; J. Exper. Clin. Canc. Res., 2011] [Babbio F, et al.; Oncogene, 2012] [Unoki, et al; Oncogene, 2004] [Xie S, et al.; J. Biol. Chem., 2012]. In colon cancer cells, we demonstrated that UHRF1 recruits Suv39H1 and DNMTs on specific gene promoters (e.g. PPARγ, CDH1 and RARβ), mediating their epigenetic silencing [Sabatino L, Fucci A, et al.; Oncogene, 2012].

Relying on all these observations, our studies focused first of all in the characterization of UHRF1 immunohistochemical levels, in relation to DNA methylation levels, in a cohort of colorectal mucosa samples, composed by normal

mucosae (N =25), MSI CRCs (N =47), and MSS CRC samples (N =88). Beyond the well-known association between UHRF1 high levels and CRC cancer progression [Sabatino L, Fucci A, et al.; Oncogene, 2012], our finding demonstrated, for the first time, a surprisingly correlation, among microsatellite CRCs, between UHRF1-high levels and the better patient's outcome. In fact, UHRF1 was significantly higher in MSI-CRC samples, compared to MSS-CRCs, correlating with an overall DNA hypermethylation in MSI-CRCs both at genome-wide and at gene specific level, in accordance with the role of UHRF1 in DNA methylation [Bostick M, et al.; Science, 2007] [Sharif J, et al.; Nature, 2007].

This intriguing result prompted us to deepen the molecular mechanism that underlies this clinical phenomenon. Thus we decided to use two cell lines derived from MSI-CRC (RKO) and MSS-CRC (HT29), demonstrating that UHRF1 protein levels correlate positively to the DNA methylation levels, in accord to the in vivo observations. In order to investigate the potential link between UHRF1 and the observed hyper-methylation we decided to manipulate the UHRF1 levels in RKO cells, since they showed the coexistence of UHRF1 high level and of the interesting molecular pathogenic features (MSI CRC model, CIMP-high, CIN-negative, BRAF-V600E, PI3KCA-H1047R) [Ahmed D., et al.; Oncogenesis, 2013].

Our studies, either with molecular methods (i.e. Pyrosequencing and MS-MLPA) and next generation sequencing (NGS) approaches (i.e. RRBS), demonstrated that UHRF1 knock down induces a significant overall DNA hypomethylation in RKO cells. In fact, almost the totality of identified DM CpGs were hypo-methylated in RRBS analysis, affecting mostly the not-CGI sites (e.g. "fully methylated" intergenic or intronic regions), underlying the pivotal role of UHRF1 in the DNA methylation maintenance [Bostick M, et al.; Science, 2007] [Sharif J, et al.; Nature, 2007] [Arita K, et al.; Nature, 2008] [Avvakumov GV, et al.; Nature, 2008] and in the heterochromatin integrity [Karagianni P, et al.; Mol. Cell. Biol., 2008] [Papait R, et al.; Mol. Biol. Cell., 2008]. Further, UHRF1 silencing induced a consistent de-methylation either in the specific promoter of MLH1 and CDH1 and in 5'UTR sequences of the repetitive element LINE-1, considered as a surrogate of global methylation due to the high representation of these sequences into the genome [Lao VV and Grady WM; Nat. Rev. Gastroenterol. Hepatol., 2011].

Deeper analysis of all these crucial CRC-related regions, by ChIP assay, helped us to demonstrate that this strong de-methylation was induced by DNMT1 delocalization from promoter regions. In fact, the levels of all DNMTs are not affected by UHRF1 KD, as observed either by WB, semi-qPCR and NGS approaches (RNA-seq, RPKM counting). Interestingly, focusing on DNMT3A/3B, already reported to interact with UHRF1 [Meilinger D, et al.; EMBO reports, 2009], we observed that only DNMT3A binding was impaired upon UHRF1 KD. On the contrary, DNMT3B revealed a clear increased binding on the promoter of CDH1, MLH1 and 5'UTR of LINE-1. However, this data is perfectly consistent with previous studies in which DNMT3B was positively correlated with CIMP CRCs [Nosho K, et al.; Clin Cancer Res, 2009], and was the final DNA-methylator-actor of a transcriptional repressor complex driven by phosphorylated MAFG in RKO cells [Fang M, et al.; Mol. Cell, 2014]. In fact, the model proposed by Fang et al. perfectly matched with our finding: BRAF mutation (V600E), present in RKO cells, promotes the transcriptional silencing of several TSGs promoters (including MLH1), due to increased BRAF/MEK/ERK signaling that induces an ERK-directed phosphorylation of MAFG at S124, which recruits the corepressor complex (BACH1, CHD8, and DNMT3B) on promoter regions. Moreover, further evidences sustain the released action of DNMT3A/3B from the UHRF1 related interacting-chromatin-actors (e.g. HP1, MeCp2, EZH2, HDAC1, PCNA, DNMT1). In fact, DNMT3A/3B displays a sort of "anchoring" with nucleosomes containing methylated CpG sites, in particular nucleosomes containing methylated SINE/LINE elements and CpG islands [Jeong S, et al.; Mol Cell Biol., 2009], without specific affinity to precise histone PTMs profiles, corroborating DNMT1 in their ongoing participation in DNA methylation maintenance, after the passage of the replication fork [Jones PA and Liang G; Nature Rev. Genet., 2009].

These observations, matched with our results, contributed to the first fundamental pivot of the molecular model that we suggest in this PhD Thesis: the lack of UHRF1 and the subsequent de-localization of DNMT1 might simulate the passage of the replication fork, reinforcing the binding of DNMT3B to the promoters that show a residual methylation in the CpGs, contained in nucleosome rich regions, providing a plausible explanation of the not complete de-methylation, observed in "local" and "genome-wide" experiments. In fact, even if the DNA de-methylation observed upon UHRF1 silencing is particularly strong, Pyrosequencing/MS-MLPA and RRBS showed a residual or persistent DNA methylation. Moreover, RRBS, through the analysis of

the percentage of the CpG methylation all over the genome, clearly demonstrated that, upon UHRF1 KD, CpG methylation percentages did not assessed in an average clustering around “hemi-methylated” or totally “not-methylated” status, as expected in a context of cycling cells without correct DNMT1 localization and functionality, reinforcing our idea about the role of DNMT3B in this cell line.

Since mis-regulation and over-expression of UHRF1 showed a well-known involvement in the disruption of euchromatic regions [Rajakumara E, et al.; Mol. Cell, 2011], affecting especially key tumor suppressor genes in CRCs [Sabatino L, Fucci A, et al.; Oncogene, 2012], we investigated whether, in RKO, UHRF1 silencing might lead to a re-expression of TSGs. Unexpectedly, our studies on CDH1 and MLH1 genes demonstrated that, even if UHRF1 KD significantly reduced the DNA methylation on their promoters, this is not sufficient to induce their re-expression, either at transcriptional (semi-qPCR and RNA-seq, RPKM counting) and translational levels (WB analysis). Interestingly, despite the de-localization of DNMT1 and the strong demethylation, we do not observe the re-expression of LINE-1 sequences .
These data suggested that several mechanisms of transcriptional repression could be convergent on their promoter regions, as previously published in another context [Babbio F, et al.; Oncogene, 2012]. In fact, ChIP analysis on CDH1, MLH1 and LINE-1 promoter regions, demonstrated that UHRF1 KD strongly decreased H3K9me3 and mild increased H3K4me3 on the promoters. H3K9me3 is a histone mark usually strictly associated with UHRF1, which is able to recruit the specific H3K9 histone methyltransferase (e.g. Suv39H1, G9a) [Meilinger D, et al.; EMBO reports; 2009] [Babbio F, et al.; Oncogene, 2012], and to mediate the propagation and stability of heterochromatic profiles, in concert with DNA methylation [Xie S, et al.; J. Biol. Chem., 2012]. H3K4me3 is a chromatin active histone mark, deposed by a several redundant HMTs (e.g. PRDM9, Set9, MML1, ASH1), well-related to transcriptional activity [Li H, et al; Nature, 2005] [Nishioka K, et al.; Nature 2006]. Noteworthy, we observed the co-occurrence in the RKO cells (in control sample, scramble) of H3K9me3 and a weak signal of H3K4me3 in CDH1 and MLH1 regions. This data is not in contrast with UHRF1 binding, that is able to recognize, through its PHD domain, the histone H3 tails also in presence of H3K4me3 (with H3K2unmodified), linking UHRF1 to the mis-regulation of euchromatic regions [Rajakumara, et al.; Mol. Cell, 2011].

Conversely UHRF1 silencing did not affect the presence of H3K27me3, deposed by EZH2, part of PRC2 (a transcriptional repressive Polycomb protein group, PcG), in the analyzed regions of CDH1 and MLH1, confirming the independence of this hPTM from UHRF1-related complexes (e.g. DNMT1, HDAC1, HMTs), as previously observed [Babbio F, et al.; Oncogene, 2012]. H3K27me3, is considered as the main histone mark of facultative heterochromatin [Dillon N.; Cell Biol.; 2004], involved in the gene silencing during the normal cell commitment and often found mis-distributed on the CGI of tumor suppressor genes in cancer [Richly H, et al; Cell Death and Disease, 2011] [Schuettengruber B, and Cavalli G; Development, 2009]. It was worthy of note that in the RKO cells (in control sample, scramble) we identified the coexistence of strong H3K9me3, H3K27me3, and mild H3K4me3 in CDH1 and MLH1 gene promoters. In fact, this data was consistent with the chromatin profile called ''balanced'' state, in which Trithorax protein group (TrxG) and Polycomb Group protein (PcG) act simultaneously or in rapid alternation from active to repressive chromatin state, in concert with other positive and/or negative epigenetic regulators [Schwartz YB, et al.; PLOS Genetics, 2010]. In these multi-protein complexes are counted some HMTs with peculiar ambiguous functionality: for instance, the HMTs SET-domain-containing ASH1 shows the ability to methylate different substrates concurrently, like H3K4, H3K9, H3K36 and H4K20 [Beisel C, et al.; Nature 2002], and is not incompatible with repressive marks, like H3K9me3 or H3K27me3, taking part of both PcG and TrxG protein complexes [Schwartz YB, et al.; PLOS Genetics, 2010]. Interestingly, LINE-1 sequences showed a slightly different chromatin profile. In RKO (control samples, scramble) they showed a strong presence of H3K9me3, in relation to UHRF1 binding and hyper-methylated DNA profile (very similar to normal peripheral blood cells [Baba Y, et al.; Mol. Cancer, 2010]), as expected due to their implication in constitutive heterochromatin. UHRF1 depletion caused strong alteration of their chromatin configuration, leading to a considerable gain of H3K4me3, with a mild novel onset of H3K27me3, resulting in the so-called bi-valent chromatin state [Richly H, et al.; Cell Death and Disease, 2011].

This novel onset of H3K27me3 in LINE-1 sequences, together with stable or increased presence of this hPTM in the promoter regions of CDH1 and MLH1, led us to formulate the second fundamental pivot of the molecular model delineated in this PhD Thesis: even if UHRF1 depletion results in a severe chromatin rearrangement from a strongly repressed chromatin state (H3K9me3 rich) toward a more open

transcriptional accessibility (H3K4me3 enrichment), the stable or mild gain of H3K27me3 established a sort of compensative-repressive mechanism leading to the observed not re-expression of CDH1, MLH1 and LINE-1 loci. Interestingly, in contrast with previous report about UHRF1 depletion and TSGs re-expression [Alhosin M, et al.; J. Exper. Clin. Canc. Res., 2011] [Babbio F, et al.; Oncogene, 2012], we demonstrated that in RKO, characterized by MSI-high and CIMP-positive pathogenetic features, PRC2 plays a key role in the maintenance of a chromatin repressive state, probably driven by the epigenetic instability underlying this CRC subclass [Grady WM, and Carethers JM; Gastroenterology, 2008], acting in concert with DNMT3B that cooperate to maintain a persistent DNA methylation on these loci, even after UHRF1 depletion and subsequent de-localization of DNMT1.

Our finding in ChIP analysis, and the previously published evidences about the crucial importance of UHRF1 ubiquitination activity in orchestrating DNMT1 functionality [Qin W., et al; Cell Research, 2015], led us to hypothesize that the observed chromatin rearrangements could be mediated by the lack of H3K18ub seated by UHRF1 RING domain. In the absence of a commercial specific antibodies against H3K18ub that allowed us to perform a ChIP in our regions of interest, we designed and produced a monoclonal antibody against the H3K18ub. Unfortunately, its validation is still ongoing due to the complexity of ubiquitination profile of histone H3. In the meanwhile, we performed a SILAC LC-MS/MS on RKO cells after siUHRF1, in order to better elucidate the involvement of UHRF1 in specific ubiquitination of histone tails. In fact, this technique gave us the advantage to analyze in a quantitative single mass-spectrometry analysis, a pool of histone peptide, mixed in ration 1 to 1, labeled with different non-radioactive stable isotopes, allowing to discriminate the different samples loaded in mass-spectrometry (i.e. control = Light isotopes; siUHRF1 = Heavy isotopes). In this experiment, we analyzed, for the first time, the role of UHRF1 in the modulation of histone ubiquitinations in a tumor cell line. Intriguingly, our finding demonstrated that UHRF1 KD induced a decrease in the mono-ubiquitination of H3K23 and a surprisingly strong increase in H3K18 bi-ubiquitination (bona fide relying on the molecular weight of the analyzed peptides that did not shown any other detectable ubiquitination). Interestingly we identified two different pool of peptides, mono-ubiquitinated in K18 or mono-ubiquitinated in K23, belonging to a different pool of histones H3 that migrate, as expected, in the same molecular weight (25KDa) during electrophoresis, and that present reciprocal

unmodified residues, respectively K23un in K18ub peptides and K18un in K23ub peptides. These two peptides showed an important decrement in ubiquitination, upon UHRF1 silencing, mostly in K23ub. On the contrary we identified another mono-ubiquitination in K18 peptides, belonging to a third different pool of histones H3, that present acetylation in K23 and an increased ubiquitination after UHRF1 depletion. Even if these data seem to not agree to the previous publications in ESC [Qin W., et al; Cell Research, 2015] these data are the first one to be performed in tumor cell line setting, without the overexpression of any form of UHRF1 or histones. Further, our finding in RKO cells, delineated a more complex pool of histone H3 ubiquitination related to UHRF1 E3-ligase activity, in accordance with other recent publication [Harrison JS, et al.; eLife, 2016] [Nishiyama A, et al.; Nature, 2013], assigning increased importance to RING domain functionality in the correct maintenance of DNA methylation patterns.

Relying on these finding and on previously reported role of RING domain in heterochromatic formation [Karagianni P, et al.; Mol. Cell. Biol., 2008], we formulated the third fundamental pivot of the molecular model that comes out from this PhD Thesis: the mono-ubiquitinations of histone H3 are fundamental for a correct localization and functionality of DNMT1, leading to a strong de-methylation, upon UHRF1 KD, as described before, and are also crucial in the maintenance of a repressed heterochromatic state of chromatin (H3K9me3 rich), leading to a severe chromatin rearrangement, after UHRF1 depletion, toward a more open transcriptional accessibility (H3K4me3 enrichment). In fact, it is well-know that acetylation in the same residues of histone H3 (e.g. K18, K23) are key marks of euchromatic active state of chromatin [Wang Z, et al.; Nat Genet., 2008]. Thus, we speculated that the decrement observed in these ubiquitinated residues after UHRF1 KD could be toward acetylation of the same residues. Unfortunately, this remains intricate to detect either in the band corresponding to the molecular weight of total H3 (15KDa) due to the coexistence of several modification (e.g. K4me3, K9me3, other acetylations, etc.) resulting in the impossibility to identify uniquely the corresponding peptides, and even in bulk analysis where the amount of gained acetylated K18 of K23 are undetectable and indistinguishable among the other more represented acetylations. On the contrary, the intriguing finding of a strong increase of H3K18 bi-ubiquitination, upon UHRF1 KD, prompted us to speculate the incoming of other E3-ubiquitin ligases that polyubiquitinate histone H3. It is well-know that

UHRF1 stability is regulated through a fine-tuning of poly-ubiquitination and de-ubiquitination processes, mediated respectively by proteins belonging to SCF complex (i.e. β-TrCP1 and β-TrCP2) [Chen H et al.; Mol. Cell. Biol.; 2013], and de-ubiquitinase USP7 [Qin W., et al; J. Cell. Biochem., 2011] [Zhang ZM, et al.; Cell Reports, 2015]. USP7 is also known to be part of PRC1 together with other several RING domain proteins (e.g. RING1A/B) [Aranda S, et al.; Sci. Adv., 2015]. PRC2, firmly present in the analyzed promoter regions as described before, is known to interact with CUL4–DDB1 ubiquitin ligase [Higa LA, et al.; Nature Cell Biol., 2006], reported to ubiquitinated histone H3 and H4 [Wang H, et al.; Molecular Cell, 2006]. All these intriguing observations prompted us to speculate that after UHRF1 depletion occurs a disequilibrium among all these ubiquitinase/de-ubiquitinase complexes leading to the increased bi-ubiquitinated H3K18ub observed.

Ultimately, we investigated the genome-wide relationship between the strong DNA de-methylation observed and the possible alteration of the transcriptome. Interestingly, the Ingenuity Pathway Analysis (IPA) showed that UHRF1 silencing interferes with several important pathways, among others cell proliferation, cell growth, DNA damage and DNA repair, cell cycle and Wnt signaling. In particular, we observed a coherent induction of the crucial actors of TFGβ signaling pathway (e.g. TFGβ, SMAD2/3). Noteworthy, the overexpression of these factors is generally related to CIN-CRCs and is associated with loss of heterozygosity (LOH) at chromosome locus 18q [Grady WM, and Carethers JM; Gastroenterology, 2008]. However, we observed a downregulation in some genes belonging to the cell cycle ontological cluster, in agreement with our previous reports [Babbio F, et al.; Oncogene, 2012]. Then, since we observed an modest percentage of DM CpGs located in the promoter regions (16%) and in body of genes (Exon 15% + Intron 35%), we coupled the DM CpGs with the DE transcripts. Our studies demonstrated that, even if there is cluster of genes that are hypo-methylated and up-regulated, these transcripts represented only the 3% of the hypo-methylated CpGs identified. This finding, perfectly consistent with our previous observation about not re-expression of CDH1, MLH1 and LINE-1 loci, suggested that, in RKO cells, only few specific loci are directly regulated by DNA methylation status and UHRF1 overexpression. We might speculate that these regions present a different chromatin profile, less regulated by PRC2, mostly controlled by UHRF1/DNMT1 and therefore more prone to the re-

expression after UHRF1 KD. However, further studies are necessary to elucidate the role of UHRF1 and DNA methylation in the regulation of these specific loci/genes.

Recapitulating the three main pivots that sustain the molecular model proposed in this PhD Thesis: (i) UHRF1 depletion and subsequent de-localization of DNMT1 induces a strong decrement in DNA methylation, however that might reinforce the binging of DNMT3B, due to its peculiar ability to recognize the residual methylation in CpGs contained in nucleosome rich regions, leading to a not complete de-methylation; (ii) UHRF1 depletion results in a severe chromatin rearrangement, inducing a decrease of H3K9me3 and an increase of H3K4me3, however the presence of PRC2 (H3K27me3) drives a compensative-repressive mechanism, that, acting in concert with DNMT3B, leads to the not re-expression of CDH1, MLH1 and LINE-1 loci; (iii) the H3 mono-ubiquitinations by UHRF1 RING domain are fundamental for the correct localization/functionality of DNMT1, leading to a strong de-methylation, upon UHRF1 KD, moreover, lack of UHRF1 lead to a severe chromatin rearrangement toward a more open transcriptional accessibility, probably due to the disruption of the axis UHRF1-H3ub-DNMT1-HMTs.

More broadly, we demonstrated that UHRF1 is deeply involved in the regulation of DNA methylation both in CRC human samples and CRC-derived cell lines. Moreover, UHRF1 plays a significant role in the establishment of heterochromatic epigenetic signatures modulating directly (ubiquitination) or indirectly (HMTs, HDAC1 recruitment) the histone PTMs profile.
In conclusion, since we observed in RKO MSI-CRC-derived cells that UHRF1 correlates positively with an overall DNA hypermethylation and a precise heterochromatin stability, in accord with literature [Karagianni P, et al.; Mol. Cell. Biol., 2008] [Papait R, et al.; Mol. Biol. Cell., 2008], we speculate that the better prognosis correlated with MSI-CRC model, could reside in the UHRF1-high levels that result in a sort of protective condition for the genome integrity, maintaining the global DNA methylation level closer to the normal mucosae, and probably counteracting the hyper-methylation of TSGs. Further studies, upon UHRF1 depletion, delineating genomic stability, by loss of heterozygosity (LOH) analysis [Matsuzaki K, et al.; Clin. Cancer Res., 2005], and epigenomic stability, by loss of imprinting (LOI) analysis [Holm TM, et al.; Cancer Cell, 2005], will clarify the proposed role of UHRF1 as a gate-keeper gene in MSI-CRCs.

6. Annex

6.1. References

Ahmed D, et al., **Epigenetic and genetic features of 24 colon cancer cell lines.** Oncogenesis, 2013

Alhosin M, et al., **Down-regulation of UHRF1, associated with re-expression of tumor suppressor genes, is a common feature of natural compounds exhibiting anti-cancer properties.** Journal of Experimental & Clinical Cancer Research; 2011.

Allis CD, et al. **New nomenclature for chromatin-modifying enzymes.** Cell 2007; 131(4):633-6

Anders S and Huber W. **Differential expression analysis for sequence count data.** Genome Biol. 2010

Antequera F and Bird A. **Number of CpG islands and genes in human and mouse.** Proc. Natl. Acad. Sci. U.S.A. 1993 90:11995-11999.

Aranda S, et al., **Regulation of gene transcription by Polycomb proteins**. Sci. Adv. 2015

Arita K., et al. **Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism.** Nature 2008; 455:818–21.

Avvakumov GV., et al. **Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1.** Nature 2008; 455:822–5.

Baba Y, et al., **Epigenomic diversity of colorectal cancer indicated by line-1 methylation in a database of 869 tumors.** Mol Cancer 2010;9:125.

Babbio F, et al., **The SRA protein UHRF1 promotes epigenetic crosstalks and is involved in prostate cancer progression.** Oncogene 2012.

Baylin S and Jones PA, **decade of exploring the cancer epigenome - biological and translational implications.** Nat Rev Cancer 2011; 11(10):726-34.

Beisel C, et al. **Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1.** Nature 2002

Berger SL. **The complex language of chromatin regulation during transcription.** Nature 2007. 447:407-412

Bird A. DNA **methylation patterns and epigenetic memory.** Genes & Development 2002.16:6-21.

Boland CR., et al., **Microsatellite instability in colorectal cancer.** Gastroenterology 2010, 138, 2073–2087

Bonapace IM, et al., **Np95 is regulated by E1A during mitotic reactivation of terminally differentiated cells and is essential for S phase entry.** J. Cell Biol. 2002 157:909–914

Bostick M, et al., **UHRF1 plays a role in maintaining DNA methylation in mammalian cells.** Science 2007 317:1760–1764.

Bronner C, et al., **Increasing role of UHRF1 in the reading and inheritance of the epigenetic code as well as in tumorogenesis.** Biochemical Pharmacology (2013), 86 1643–1649.

Bronner C, et al., **The UHRF family: oncogenes that are drugable targets for cancer therapy in the near future?** Pharmacol Ther. 2007 Sep;115(3):419-34. Review.

Capocaccia R, et al., **Estimation and projections of colorectal cancer trends in Italy.** Int J Epidemiol 1997; 26 (5): 924-932.

Chen et al., **Epigenetic regulation and cancer (Review).** Oncology Reports (2014), 31: 523-532

Chen H, et al., **DNA damage regulates UHRF1 stability via the SCF(β-TrCP) E3 ligase.** Mol Cell Biol 2013. 33(6):1139–1148.

Cheng J, et al., **Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger 1) domains protein.** J Biol Chem 2013.

Citterio E, et al., **Np95 is a histone-binding protein endowed with ubiquitin ligase activity.** Mol Cell Biol. 2004 Mar;24(6):2526-35.

Clements A, et al., **Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase.** Mol. Cell 12003. 2, 461–473.

Criscione SW et al., **Transcriptional landscape of repetitive elements in normal and cancer human cells.** BMC Genomics 2014

Crnogorac-Jurcevic T et al., **Proteomic analysis of chronic pancreatitis and pancreatic adenocarcinoma.** Gastroenterology. 2005 Nov;129(5):1454-63.

Cuthbert GL, et al., **Histone deimination antagonizes arginine methylation.** Cell. 2004 Sep 3;118(5):545-53.

Daskalos A, et al., **UHRF1-mediated tumor suppressor gene inactivation in non small cell lung cancer.** Cancer. 2011 Mar 1;117(5):1027-37.

Dawson M and Kouzarides T, **Cancer Epigenetics: From Mechanism to Therapy.** Cell 2012; 150(1):12-27

Deaton AM and Bird A**., CpG islands and the regulation of transcription.**

Dillon N. **Heterochromatin structure and function**. Biol Cell. 2004 Oct;96(8):631-7.

Eden A, et al., **Chromosomal instability and tumors promoted by DNA hypomethylation.** Science. 2003 Apr 18;300(5618):455.

Ehrlich M. **DNA hypomethylation in cancer cells.** Epigenomics. 2009 Dec;1(2):239-59.

Estécio MR, et al., **LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability.** PLoS One. 2007 May 2;2(5):e399.

Esteller M. **Epigenetics in cancer.** N Engl J Med. 2008 Mar 13;358(11):1148-59.

Fan T, et al., **DNA hypomethylation caused by Lsh deletion promotes erythroleukemia development.** Epigenetics. 2008 May-Jun;3(3):134-42.

Fang M, et al., **The BRAF oncoprotein functions through the transcriptional repressor MAFG to mediate the CpG Island Methylator phenotype.** Mol Cell. 2014 Sep 18;55(6):904-15.

Fearon, et al., **A genetic model for colorectal tumorigenesis.** Cell 1990, 61, 759–767

Feinberg AP and Vogelstein B, **Hypomethylation distinguishes genes of some human cancers from their normal counterparts.** Nature. 1983 Jan 6;301(5895):89-92.

Ferlay J, et al., **Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012**. Int J Cancer. 2015 Mar 1;136(5):E359-86.

Fischle W, et al., **Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation.** Nature 2005 438, 1116–1122

Fraley C and Raftery ER. **Model-based clustering, discriminant analysis, and density estimation.** Journal of the American Statistical Association 2002;97:611-631.

Fujimori et al., **Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation.** Mamm Genome 1998, (12):1032-5.

Furlan D, et al., **Diagnostic utility of MS-MLPA in DNA methylation profiling of adenocarcinomas and neuroendocrine carcinomas of the colon-rectum.** Virchows Arch. 2013 Jan;462(1):47-56.

Giorgi Rossi P, et., **Impact of Screening Program on Incidence of Colorectal Cancer: A Cohort Study in Italy .** The American Journal of Gastroenterology 2015 110, 1359-1366

Goelz SE, et al., **Hypomethylation of DNA from benign and malignant human colon neoplasms.** Science. 1985 Apr 12;228(4696):187-90

Grady WM, et al., **Genomic and epigenetic instability in colorectal cancer pathogenesis.** Gastroenterology 2008, 135, 1079–1099.

Guan D, et al., **The epigenetic regulator UHRF1 promotes ubiquitination-mediated degradation of the tumor-suppressor protein promyelocytic leukemia protein** Oncogene. 2013 Aug 15;32(33):3819-28.

Guo JU, et al., **Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain**. Cell 2011. 145:423-434.

Hampel H, et al., **Feasibility of screening for Lynch syndrome among patients with colorectal cancer.** J Clin Oncol. 2008; 26:5783–5788.

Harrison J, et al., **Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1**. eLife 2016

Hashimoto H, et al., **The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix.** Nature 2008.

Hashimoto H, et al., **UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications**. Epigenetics 2009.

Hendrich B and Tweedie S, **The methyl-CpG binding domain and the evolving role of DNA methylation in animals.** Trends Genet. 2003 May;19(5):269-77.

Hervouet E, et al., **Disruption of Dnmt1/PCNA/UHRF1 interactions promotes tumorigenesis from human and mice glial cells.** PLoS One. 2010 Jun 29;5(6):e11333.

Higa LA, et al., **CUL4–DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation** Nature Cell Biology 2006, 8, 1277 - 1283

Holm TM, et al., **Global loss of imprinting leads to widespread tumorigenesis in adult mice.** Cancer Cell 2005

Hopfner R, et al., **ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase II alpha expression.** Cancer Res. 2000 Jan 1;60(1):121-8

Hopfner R, et al., **Overexpression of ICBP90, a novel CCAAT-binding protein, overcomes cell contact inhibition by forcing topoisomerase II alpha expression.** Anticancer Res. 2002 Nov-Dec;22(6A):3165-70.

Hu L., et al. **Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2.** Cell Research 2011; 21:1374–8.

Illingwort RS and Bird., **CpG islands–'a rough guide'.** FEBS Lett 2009. 583:1713-1720.

Jacobs SA and Khorasanizadeh S, **Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail.** Science. 2002 Mar 15;295(5562):2080-3.

Jeanblanc M, et al., **The retinoblastoma gene and its product are targeted by ICBP90: a key mechanism in the G1/S transition during the cell cycle.** Oncogene. 2005 Nov 10;24(49):7337-45.

Jenkins Y. et al., **Critical Role of the Ubiquitin Ligase Activity of UHRF1, a Nuclear RING Finger Protein, in Tumor Cell Growth.** Molecular Biology of the Cell Vol. 16, 5621–5629, December 2005

Jeong S, et al. **Selective anchoring of DNA methyltransferases 3A/3B to nucleosomes containing methylated DNA**. Mol Cell Biol. 2009

Jin W, et al., **UHRF1 is associated with epigenetic silencing of BRCA1 in sporadic breast cancer.** Breast Cancer Res Treat. 2010.

Jones PA and Rethinking L. **How DNA methylation patterns are maintained**. Nature Rev. Genet. 10, 805–811 (2009).

Karagianni P, et al., **ICBP90, a Novel Methyl K9 H3 Binding Protein Linking Protein Ubiquitination with Heterochromatin Formation.** Molecular and Cellular Biology, 2008.

Karpf AR and Matsui S, **Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells.** Cancer Res. 2005 Oct 1;65(19):8635-9.

Kim, et al., **UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells.** Nucleic Acids Research, 2009, Vol. 37, No. 2 493–505

Kondo Y, Shen L, Issa JP, **Critical role of histone methylation in tumor suppressor gene silencing in colorectal cancer.** Mol Cell Biol. 2003.

Kouzarides T. **Chromatin Modifications and Their Function**. Cell; 2007.

Krishnamoorthy T, et al., **Phosphorylation of histone H4 Ser1 regulates sporulation in yeast and is conserved in fly and mouse spermatogenesis.** Genes Dev. 2006 Sep 15;20(18):2580-92.

Lao VV, et al., **Epigenetics and colorectal cancer.** Nat. Rev. Gastroenterol. Hepatol. 2011, 8, 686–700.

Li B, et al., **The Role of Chromatin during Transcription**. Cell 2007; 128(4):707-19

Li E, et al. **Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.** Cell 1992; 69:915–926.

Li H et al., **Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF.** Nature. 2005 Dec 22;438(7071):1181-5.

Liu X, et al., **UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9.** Nat Commun. 2013;4:1563.

Lorenzato M et al., **Cell cycle and/or proliferation markers: what is the best method to discriminate cervical high-grade lesions?** Hum Pathol. 2005 Oct;36(10):1101-7.

Lynch HT, et al., **Hereditary colorectal cancer.** N. Engl. J. Med. 2003, 348, 919–932

Malik S and Roeder RG. **Dynamic regulation of pol II transcription by the mammalian Mediator complex.** Trends Biochem Schwartz. Sci. 2005 30:256-263.

Mancini M, et al., **In vitro hydroquinone-induced instauration of histone bivalent mark on human retroelements (LINE-1) in HL60 cells.** Toxicol In Vitro. 2016

Mármol I, et al., **Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer.** Int. J. Mol. Sci. 2017, 18, 197

Matsuzaki K, et al., **The Relationship between Global Methylation Level, Loss of Heterozygosity, and Microsatellite Instability in Sporadic Colorectal.** Cancer Clin Cancer Res 2005

Meilinger D, et al., **Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells.** EMBO reports (2009) 10, 1259–1264.

Meissner A., **Guiding DNA Methylation.** Cell Stem Cell 2011. 9:388-390.

Mousli M, et al., **ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells.** Br J Cancer. 2003 Jul 7;89(1):120-7.

Mudbhary et al., **UHRF1 Overexpression Drives DNA Hypomethylation and Hepatocellular Carcinoma,** Cancer Cell 2014

Nan X, et al., **Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex.** Nature. 1998 May 28;393(6683):386-9.

Nathan D et al., **Histone modifications: Now summoning sumoylation** Proc Natl Acad Sci U S A. 2003 Nov 11;100(23):13118-20.

Nelson CJ, et., **Proline isomerization of histone H3 regulates lysine methylation and gene expression.** Cell 2006. 126, 905–916

Nishioka K et al., **Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation.** Nature. 2006 Jul 6;442(7098):91-5.

Nishiyama A, et al., **Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication.** Nature. 2013 Oct 10;502(7470):249-53.

Nosho K, et al., **DNMT3B expression might contribute to CpG island methylator phenotype in colorectal cancer.** Clin Cancer Res. 2009 Jun 1;15(11):3663-71.

Nowak SJ and Corces VG, **Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation.** Trends Genet. 2004 Apr;20(4):214-20.

Ogino S, et al., **Molecular classification and correlates in colorectal cancer.** J Mol Diagn. 2008 Jan;10(1):13-27.

Okano M, et al. **DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.** Cell 1999; 99:247–257

Papait R, et al., **The PHD domain of Np95 (mUHRF1) is involved in large-scale reorganization of pericentromeric heterochromatin.** Mol Biol Cell. 2008 Aug;19(8):3554-63.

Pino MS, et al., **The chromosomal instability pathway in colon cancer.** Gastroenterology 2010, 138, 2059–2072.

Probst AV, et al., **Epigenetic inheritance during the cell cycle**. Nat Rev Mol Cell Biol 2009; 10(3):192-206.

Qin W et al., **DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination.** Cell Res. 2015 Aug;25(8):911-29.

Qin W, et al., **Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1.** J Cell Biochem 2011; 112:439-444.

Raftery et al., **Model-based clustering, discriminant analysis, and density estimation** Journal of the American Statistical Association; Jun 2002; 97, 458.

Raha T., et al. **HIV-1 Tat stimulates transcription complex assembly through recruitment of TBP in the absence of TAFs**. PLoS Biol 2005; 3.

Rajakumara E et al., **PHD Finger Recognition of Unmodified Histone H3R2 Links UHRF1 to Regulation of Euchromatic Gene Expression.** Molecular Cell 2011

Rappsilber J, et al., **Analysis of the topology of protein complexes using cross-linking and mass spectrometry.** CSH Protoc. 2007 Feb 1;2007

Reik W. **Stability and flexibility of epigenetic gene regulation in mammalian development.** Nature 2007. 447:425-432.

Richly H, et al**., Roles of the Polycomb group proteins in stem cells and cancer.**

Riggs AD, et al. **Regulation in Epigenetic Mechanisms of Gene.** Cold Spring Harbor, New York, 1996.

Rothbart SB, et al., **Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation.** Genes & Development 2013 27:1288–1298.

Rottach A, et al., **The multi-domain protein Np95 connects DNA methylation and histone modification.** Nucleic Acids Res 2010

Sabatino L, et al., **UHRF1 coordinates peroxisome proliferator activated receptor gamma (PPARG) epigenetic silencing and mediates colorectal cancer progression.** Oncogene. 2012 Dec 6;31(49):5061-72.

Schuettengruber and Cavalli. **Recruitment of Polycomb group complexes and their role in the dynamic regulation of cell fate choice.** Development 136, 3531-3542 (2009)

Schulz WA, et al**., Methylation of endogenous human retroelements in health and disease.** Curr Top Microbiol Immunol. 2006;310:211-50. Review.

Schwartz YB, et al., **Alternative Epigenetic Chromatin States of Polycomb Target Genes.** Plos genetics 2010

Schwarz G. **Estimating the dimension of a model.** The Annals of Statistics 1978;6:461-464.

Sharif J, et al., **The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA.** Nature 20017 450:908–912.

Shilatifard A. **Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression**. Annu. Rev. Biochem. 2006. 75, 243–269

Smale S. **Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes.** Biochimica et Biophysica Acta 1351 (1997). 73–88

Soldi M, et al., **Mass spectrometry-based proteomics for the analysis of chromatin structure and dynamics**. J Mol Sci. 2013 Mar 6;14(3):5402-31.

Sterner DE and Berger SL, **Acetylation of histones and transcription-related factors.** Microbiol Mol Biol Rev. 2000 Jun;64(2):435-59. Review.

Stoffel EM, et al., **Familial colorectal cancer, beyond Lynch syndrome.** Clin Gastroenterol Hepatol. 2014 Jul;12(7):1059-68.

Straussman R, et al., **Developmental programming of CpG island methylation profiles in the human genome.** Nat Struct Mol Biol. 2009;16:564–571

Tan M, et al., **Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification.** Cell. 2011 Sep 16;146(6):1016-28.

Treviño LS, et al., **Phosphorylation of epigenetic "readers, writers and erasers": Implications for developmental reprogramming and the epigenetic basis for health and disease.** Progress in Biophysics and Molecular Biology (2015), 118 8-13.

Umar A, et al., **Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (lynch syndrome) and microsatellite instability.** J. Natl. Cancer Inst. 2004, 96, 261–268.

Unoki M, et al., **ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain.** Oncogene. 2004 Oct 7;23(46):7601-10.

Unoki M, et al., **UHRF1 is a novel molecular marker for diagnosis and the prognosis of bladder cancer.** Br J Cancer. 2009 Jul 7;101(1):98-105.

Unoki, et al., **UHRF1 is a novel diagnostic marker of lung cancer.** Br J Cancer. 2010 July 13; 103(2): 217–222.

Wang H, et al., **Histone H3 and H4 Ubiquitylation by the CUL4-DDB-ROC1 Ubiquitin Ligase Facilitates Cellular Response to DNA Damage Molecular** Cell 22, 383–394, May 5, 2006

Wang Z, et al., **Combinatorial patterns of histone acetylations and methylations in the human genome.** Nat Genet. 2008 July; 40(7): 897–903.

Weisenberger DJ, et al., **CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer.** Nat. Genet. 2006, 38, 787–793.

Wu C, et al., **CpG Island Methylation, Microsatellite Instability, and BRAF Mutations and Their Clinical Application in the Treatment of Colon Cancer.** Chemotherapy Research and Practice 2012

Wu H, et al., **Molecular basis for the regulation of the H3K4 methyltransferase activity of PRDM9.** Cell Rep. 2013 Oct 17;5(1):13-20.

Xie S, et al., **UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail.** J Biol Chem 2012.

Yang GL, et al., **UHRF1 is associated with tumor recurrence in non-muscle-invasive bladder cancer.** Med Oncol. 2012 Jun;29(2):842-7.

Zhang Y and Reinberg D, **Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.** Genes Dev 2001. 15, 2343-60.

Zhang ZM, et al., **An Allosteric Interaction Links USP7 to Deubiquitination and Chromatin Targeting of UHRF1.** Cell Rep. 2015 Sep 1;12(9):1400-6.

6.2. Publications & Co-Curricular Activities

Publications

Mancini M, Mandruzzato M, Garzia AC, Sahnane N, Magnani E, **Macchi F**, Oulad-Abdelghani M, Oudet P, Bollati V, Fustinoni S, Furlan D, Bonapace IM "*In vitro hydroquinone-induced instauration of histone bivalent mark on human retroelements (LINE-1) in HL60 cells*". *Toxicol In Vitro. 2016 Dec 13;40:1-10. doi: 10.1016/j.tiv.2016.12.007*

Abstracts

KAUST-UCI international Symposium 2017; Thuwal, Saudi Arabia
Macchi F, Magnani M, Sanhane N, Rizzo R, Soldi M, Zhang C, Weisz A, Sadler Edepli K, Furlan D, Bonaldi T, Bonapace IM "*UHRF1 coordinates DNA methylation, H3K9me3, H3K4me3 and H3K23ub in colon cancer cells*".

Joint National Ph.D. Meeting, 2016; Salerno, Italy
F. Macchi, E. Magnani, N. Sanhane, F. Rizzo, A. Weisz, D. Furlan, IM. Bonapace "*UHRF1 interplay between DNA methylation and histone PTMs may lead to a severe chromatin modifications rearrangement in colon cancer*".

KAUST-UCI international Symposium 2015; Thuwal, Saudi Arabia
Macchi F, Magnani E, La Mastra F, Sanhane N, Rizzo F, Leonhardt H, Weisz A, Furlan D, Bonapace IM "*DNA hypo- and hyper-methylation in colon cancer: may UHRF1 decrypt the epigenetic basis of the paradox?*".

Seminar/Congress SIBBM 2015; Torino, Italy
F. Macchi, E. Magnani, M. Legnani, N. Sanhane, F. Rizzo, A. Weisz, D. Furlan, IM. Bonapace "*DNA hypo- and hyper-methylation in colon cancer: may UHRF1 decrypt the epigenetic basis of the paradox?*".

Seminar/Congress SIBBM 2014; Trento, Italy
F. Macchi, E. Magnani, M. Mancini, C. Pistore, F Rizzo, A. Weisz, C. Catapano, IM. Bonapace "*The role of UHRF1 in CDH1 epigenetic regulation via promoter associated non-coding RNAs*".

Joint National Ph.D. Meeting, 2013; Pesaro, Italy
E. Magnani, **F. Macchi**, C. Pistore, F. Babbio, L. Curti, G. Carbone, C. Catapano, IM. Bonapace "*The role of UHRF1 in E-cadherin epigenetic regulation: a modulation via promoter associated non-coding RNAs*".

Seminar/Congress SIBBM 2013; Pavia, Italy
E. Magnani, F. Babbio, **F. Macchi**, C. Pistore, M. Pitaro, S. Tagliaferri, M. Mancini, C. Catapano, IM. Bonapace "*The role of UHRF1 in E-cadherin epigenetic regulation: a modulation via promoter associated non-coding RNAs*".

EMBO Conference: Chromatin and Epigenetics; 2013 EMBL Heidelberg, Germany
E. Magnani, C. Pistore, **F. Macchi**, F. Babbio, L. Curti, G. Carbone, C. Catapano, IM. Bonapace "*The role of UHRF1 in E-cadherin epigenetic regulation: a modulation via promoter associated non-coding RNAs*".

Conferences and Seminars

EPIGEN Seminar/Congress 2016; Roma, Italy
Annual Meeting, May 2016, Ergife Palace Hotel, Roma, IT.

Joint National Ph.D. Meeting, 2016; Salerno, Italy
Annual Meeting, April 2016, Grand Hotel Salerno, Salerno, IT.

Seminar/Congress SIBBM 2015; Torino, Italy
Seminar "*From Genomes to Functions*", July 2015, Molecular Biotechnology Center, University of Turin, Torino, IT.

Joint National Ph.D. Meeting, 2014; Pesaro, Italy
Annual Meeting, October 2014, Baia Flaminia Resort, Pesaro, IT.

Seminar/Congress SIBBM 2014; Trento, Italy
Seminar *"Emerging Arenas in Molecular Biology: from basic mechanisms to personalized medicine",* June 2014, University of Trento, Trento, IT.

EPIGEN Seminar/Congress 2014; Roma, Italy
Annual Meeting, February 2014, Ergife Palace Hotel, Roma, IT.

Seminar/Congress SIBBM 2013; Pavia, Italy
Seminar *"Revisiting the Central Dogma: Emerging New Concepts in Replication, Transcription and Translation",* June 2013, University of Pavia, Pavia, IT.

Professional Memberships

SIBBM - Italian Society of Biophysics and Molecular Biology

Further Courses & Lectures

March 2015, Busto Arsizio
Course of Next Generation Sequencing

Sept 2014, Busto Arsizio
Course of Systems Biology

6.3. PhD experience abroad

During my PhD experience, I had the opportunity to spend two long periods, both of them in the Laboratory of Prof. Kirsten Sadler Edepli. The first one (December 2015) was spent at Mont Sinai Medical School of Medicine (New York City, USA) and the second one (August 2016 – December 2016) was spent at New York University Abu Dhabi (Abu Dhabi, UAE). The goal of these experiences was to better evaluate the role of the phosphorylation of S661 in altering RING-mediated E3 ubiquitination activity. In particular, we were interested to better investigate the effects of this phosphorylation on UHRF1 localization, and on the ability of UHRF1 to mono-ubiquitinate the H3K18 and poly-ubiquitinate DNMT1. Regarding the role of the allosteric regulations and/or PTMs of UHRF1, several publications report that different factors/PTMs are able to interact with the polybasic region (PBR) of UHRF1. It has been reported that the phosphatidylinositol 5-phosphate (PI5P) [*Gelato K. A. et al.;, Molecular Cell, 2014*] and the deubiquitinase USP7 [*Qin W., et al; Journal of Cellular Biochemistry, 2011*] [*Zhang et al.;, Cell Reports, 2015*] interact with PBR (641-657 AA of UHRF1 isoform 1 or 656-669 AA of isoform 2). Furthermore, it has been shown that UHRF1 can interact with a novel lncRNA, named UPAT, over-expressed in cancer, that protect UHRF1 from ubiquitination on K650 (K663 of isoform 2) by β-TrCP1 and β-TrCP2, and its degradation via proteasome [*Taniue K. et al.; PNAS, 2015*].

Concerning PTMs, UHRF1 can be phosphorylated on different serines, nearby the polybasic region (PBR). Serine 639 (in UHRF1 isoform 1, S652 in isoform 2) was reported to be phosphorylated by CDK1/cyclin B in M phase, regulating UHRF1 stability via proteasomal degradation, and impairing the physical association with the deubiquitinase USP7 [*Ma H. et al.; PNAS, 2012*]. It has been shown that Serine 661 (S674 on UHRF1 isoform 2) is phosphorylated by the cyclin A/CDK2 during S phase, both in vitro and in vivo. This phosphorylation is essential for embryonic development, and alters the localization of UHRF1 under some conditions [*J. Chu et al.; Molecular Biology of the Cell, 2012*]. Moreover, it has been observed that the UHRF1-S661A (non-phosphorylatable residue) is less stable and constitutively nuclear in contrast with wild-type UHRF1, which is more stable and localizes both in the nucleus and the cytoplasm. Furthermore, it has been shown that UHRF1^{p661} interacts with UHRF1 binding protein 1 (UHRF1BP1) in the cytoplasm [*Oranu A. et al.; unpublished*].

Recently it has been reported a novel molecular function of UHRF1^{pS661}, that seems to be alternative or additional to its well-known role in heterochromatin formation and epigenetics. In fact, this phosphorylation confers the ability to be recruited by BRAC1 (via BRCT domain) to Double Strand Breaks (DSBs). Thus, UHRF1^{pS661}, by the RING domain, mediates K63-linked poly-ubiquitination of RIF1, promoting its dissociation from DSBs, inactivating RIF1 focus formation, and facilitating the switch of DSBs repair pathway choice from NHEJ to HR (mediated by the BRCA1 pathway) [*Zhang H. et al.; Nature Communications, 2015*].

In this recent publication, the authors claimed that this important novel function in DSBs repair, related to UHRF1 and its E3 ligase activity, is separate or alternative from its role in mono-ubiquitination of H3. Nevertheless, the authors showed only a WB on H3 unmodified, evaluating the presence of signal corresponding to expected molecular weight of the mono-ubiquitination of H3 (23KDa).

Basing on this knowledge, the first experiments were performed in order to investigate if phospho-S661 could impair/alter the RING-domain activity. We performed a histone acid extract in HeLa cells overexpressing alternatively UHRF1, UHRF1-S661E, DNMT1 and UHRF1-S661A (Figure F4, bottom panel, on left). Unfortunately, it seems that the overexpression of the plasmids had some toxicity in HeLa cell (morphological alteration, data not shown), altering the profile of H3K18ub in strange ways. Furthermore, the presence of endogenous UHRF1 (Wild Type), although present in low amount in HeLa cells could be able to act normally, causing the failure of this experiment.

For all these reasons, we planned to design the specific point mutation (on serine 661) using the genome editing tool CRISP/Cas9 in RKO cells. At the moment, I have performed the first steps in this direction, during my second experience abroad, and I will continue in the next months within the newly established collaboration between our two Labs.