Executive Summary:

Besides featuring classic genetic mutations, cancer cells present a profoundly distorted epigenetic landscape. The cancer epigenome is characterized by a massive overall loss of DNA methylation, but at the same time by the acquisition of specific patterns of hypermethylation at CpG islands of certain promoters. These changes in DNA methylation landscape compromise the stability of our genome and the correct gene expression pattern.

CpG island promoter hypermethylation has a tumor-type specific pattern that can be used as predictor of cancer diagnosis and prognosis. Thus, DNA methylation is a potential biomarker for early detection and monitoring of cancer and for the individualization of cancer treatment. However, DNA methylation biomarkers need to be defined, checked for its sensitivity and sensibility, and validated.

Therefore, in January 2008, the CancerDIP project emerged as the joint effort of leading European groups in the field of epigenetics to help define the epigenomic profiles, focusing in two types of human cancers: colon tumors and leukemias. To achieve this goal, the CancerDIP project has used state-of-the-art techniques to obtain DNA methylation profiles. First approaches were based on Methylated DNA Immunoprecipitation (MeDIP) technique. However, as time went by and the project progressed, CancerDIP decided to take benefit of the new technologies available: MethylCap and DNA methylation arrays. Actually MethylCap protocol was developed, tested and validated within the consortium.

The CancerDIP consortium has provided the cancer research community with:

- New techniques (MethylCap) and molecular tools (antibodies).
- New insights on molecular mechanisms of DNA methylation regulation.
- Several DNA methylation data from colon, breast tumors and leukemias.
- New software packages (EpiGraph, Methmarker, BiQ Analyzer HT).

Although further research is needed, undoubtedly the input from CancerDIP will benefit not only the research community but also be of great help to improve the health condition within the European community.

Project Context and Objectives:

In cancer the whole epigenetic settings of the cell appears to be disrupted. Alterations in the DNA methylation and histone modifications patterns led to genomic instability and abnormal gene expression. Cancer cells present overall DNA hypomethylation, except for some promoters located in CpG islands that tend to be hypermethylated. The profile of CpG island hypermethylation is tumor-type specific, fact that opens the door to its use as biomarkers.

CancerDIP consortium was designed aiming to obtain methylomes that allow identifying differentially methylated regions (DMRs) in matched healthy and cancer samples.

The genome wide DNA methylation analysis of patient samples has provided information on methylation patterns that can be used for diagnostic and prognostic analyses, as well as for future identification of novel therapeutic targets.

The CancerDIP project has been structured around five clear objectives:

Objective 1. Optimization and Validation of the MeDIP technology

The first objective aims to provide the consortium with a reliable protocol to immunoprecipitate methylated DNA that can be further applied to high throughput techniques such as tiling arrays or deep sequencing. In this regard, the CancerDIP goals have been:

- Optimization and validation of MeDIP protocol.
- Establishment of MethylCap: affinity purification of methylated DNA fragments using a protein resin consisting of GST-MBD-domain of MeCP2.
- MethylCap profiling of hematopoietic cell lines and leukemia samples: combination of MethylCap and deep sequencing.
-MethylCap profiling of colon cancer cell lines and tumors: combination of MethylCap and deep sequencing.

To optimize and validate MeDIP and MethylCap two groups of cell lines have been used: colorectal cell lines (e.g. HCT116,) and leukemia cell lines (e.g. NB4). These two goals are reported in Deliverable D1.1 and Deliverable D1.2 within WP1. The data obtained from MethylCap profiling of hematopoietic and colon tumors is also reported in Deliverables D1.3 and D1.4, respectively.

Objective 2. Analysis of Mechanisms of Epigenetic Deregulation in Human Cancer.

The second objective aims to shed some light into the molecular processes that guide epigenetic deregulation in cancer. In this regard, the CancerDIP goals have been:

- Analysis of the molecular mechanisms implicated in targeting DNA methylation to genes that become silenced in cancer. Role of Polycomb group proteins
- Study of the mechanisms that trigger Polycomb-mediated abnormal DNA methylation in cancer.
- Generation and characterization of ChIP-on-chip antibodies for the analysis of Polycomb group protein members.

The mechanistic connections between Polycomb group proteins and MBD proteins have been studied and reported in Deliverable D2.1. The relationship between PcG proteins and MBS characterized at a genome-wide level was impossible to achieve due to the lack of proper antibodies. As an alternative, interesting results regarding breast cancer DNA methylation profiles have been obtained and reported in Deliverable D2.3. Moreover, a novel mechanism regulating DNA methylation in cancer has been described: the regulation of DNMTs enzymatic activity by a post-translational phosphorilation, reported in Deliverable D2.2. In order to achieve the previous goals, the generation of new ChIP grade antibodies

of the epigenetic machinery was a key step. Although such great quality antibodies are not easy to obtain the attempts and successes are reported in Deliverable D2.4.

Objective 3. Epigenetic profiling in leukemias and colon cancer: establishment of clinical correlations.

A general objective common to WP3 and WP4 has been to generate DNA methylation profiles of both leukemias and colorectal cancer samples. Specific goals have been:

- Mapping of the DNA methylome in leukemias and colorectal samples.
- Generation of DNA methylation signature in samples from acute myeloid leukemia patients (WP3) and colorectal cancer (WP4).
- Correlation analysis between DNA hypermethylation and clinical parameters.
- Identification of histone modification patterns which may have an impact on the epigenetic deregulation of the tumor cell (WP4).

The DNA methylation patterns for leukemias have been obtained from samples containing different chromosomal translocations, normal CD34+ cells and AML blasts; and reported in Deliverables D3.1, D3.2 and D3.3. To obtain colorectal cancer DNA methylation profiles, samples from cell lines and paired normal versus tumor samples have been used, and reported in Deliverable D4.2 and D4.4. Clinical correlation with DNA methylation results has been assessed and reported in Deliverables D3.5 and D4.3. In colon tumors, correlation between DNA methylation and histone modifications has been studied and reported at Deliverable D4.5.

Objective 4. Development of a Cancer MeDIP Kit and Validation.

The fourth objective aimed to produce a MeDIP kit for further studies of DNA methylation. Novel kits should facilitate research and contribute to improve cancer diagnosis and prognosis, leading to better clinical management. CancerDIP goals in this regard have been:

- Development of a kit format for the MeDIP assay.
- Kit validation on clinical samples.
- Production and quality control of kits.

A new kit for MeDIP assay has been generated and nowadays it is commercially available (reported in Deliverable D5.1). Not only a MeDIP kit has been produced with great success in the market, but the kit has also been improved giving rise to the MagMeDIP kit, which uses magnetic beads in the protocol. Both kits have been used for costumers with clinical samples and inputs so far are of great content. Moreover, the MagMeDIP kit has been adapted for it to be used on Diagenode SX-8G IP-Star and SX-8G IP-Star Compact automated platforms, saving researchers time and increasing the reproducibility of the obtained results.

Objective 5. Generation of Bioinformatic Tools for analysis of MeDIP data.

Epigenomic studies generate great quantities of data, which cannot be correctly processed without the adequate bioinformatic tools. Thus, the bioinformatic support has been crucial through all CancerDIP evolution. The main bioinformatic goals have been:

- Computational discovery and validation of hypotheses regarding functional interactions within the DNA methylation machinery, Polycomb proteins and a number of epigenetic (histone modifications) features and chromosomal translocations (for leukemias).
- Computational ranking of all cancer-specific differentially methylated regions.
- Optimization and evaluation of the predictive power and robustness of biomarker candidates.

The interplay between several chromatin-modifying proteins (MeCP2, LSD1, DNMTs, Polycomb proteins) and DNA methylation has been analyzed, using bioinformatic pipelines specially developed for a number of experimental methods including: ChIP-on-chip with NimbleGen and Agilent tiling microarrays, MeDIP on the same platforms, ChIP-seq for histone modifications, and Infinium DNA methylation analysis (reported on Deliverable D6.2). A software package has been developed for identifying and ranking candidate biomarkers from DNA methylation profiles of clinical samples. The utility of this software package has been checked in a benchmarking study comparing the DNA methylation profiles of four samples (including a colon tumor and matched normal colon tissue) obtained by four different methods for DNA methylation profiling (reported in Deliverable D6.3). Part of the work focused on the development and validation of biomarkers that accurately predict the tissue type of "cancers of unknown primary origin" (CUPs). These newly discovered biomarkers have the potential to significantly improve the clinical treatment of CUPs (reported in Deliverable D6.4). New versions of other bioinformatic software have also been developed: the EpiGraph (reported on Deliverable D6.1) and the MethMarker (reported on Deliverable D6.5).

Project Results:

Work Package 1.

TASK 1.1. Optimization of MeDIP protocol

D1.1. Generation of a standard protocol for MeDIP.
D1.5. Establishment and optimization of MethylCap-seq.

In the first part of the project we optimized protocols for MeDIP and the subsequent hybridisation of the immunoprecipitated DNA onto oligo tiling arrays (deliverable 1.1). Optimization included antibody titration and the use of internal controls (figure 1). However, later we decided to switch to next-generation sequencing, because this allows for higher throughput, resolution and accuracy then a microarray-based technique. We tried to combine MeDIP with the sequencing, but encountered several problems. Eventually, we also changed the method for the capture of the methylated DNA. We started using a methyl-binding domain (MBD) instead of an antibody. We carefully worked out this approach, which we called MethylCap, and it was converted to an automatic process with the Diagenode SX-8G IP-Star™ system (deliverable 1.5).

TASK 1.2. Validation of MeDIP

D1.2. Validation by MSP and BS of the MeDIP protocol.
D1.6. Establishment of MethylCap-BS-seq for validation of MethylCap-seq.

We performed MeDIP-chip on hematopoietic CD34+ precursor cells from a healthy individual, acute myeloid leukemia (AML) patient blasts, and the acute promyelocytic (APL) cell line NB4. To validate that high MeDIP signals correlate with high absolute methylation values, classical bisulphite sequencing was used to analyse several such sites in NB4 cells. This analysis confirmed high methylation within all enriched sites tested. Furthermore, 20 DMRs between CD34+ and NB4 (both decreases and increases in methylation) were selected for targeted MeDIP analysis using quantitative PCR. 14 out of the 20 selected DMRs showed the expected difference in methylation (deliverable 1.2).

For MethylCap-seq we did tests with cell lines including a DNMT knockout, fully methylated DNA, and DNA from several different human tissues, and did bisulphite sequencing of selected regions for validation. The MethylCap-seq protocol was published in a methods paper in 2010 (Brinkman A.B. et al. 2010. Methods 52: 232-23). In addition, MethylCap-seq profiles of a matched pair of colon tumor and adjacent normal colon tissue and of two human embryonic stem cells were incorporated in cross-comparison study of four methods to map DNA methylation on a genomic scale (Bock C. et al. 2010. Nature Biotechnology 28: 1106-14). Partner 5 of the CANCERDIP consortium carried out this study.

To further validate the methylated state of the captured fragments we developed MethylCap-BS-seq. This consists of bisulphite conversion of MethylCap-captured DNA followed by sequencing. This showed that the captured fragments are highly methylated, and that unmethylated DNA is practically absent (deliverable 1.6).

The carefully developed and validated protocols of both the MeDIP and MethylCap where used by partner 6 for the development of commercial kits.

TASK 1.3. Microarrays for the epigenetic signature in leukemias and colon cancer

D1.3. Generation of a list of methylated genes in hematopoietic cell lines.
D1.4. Generation of a list of methylated genes in colorectal cancer cell lines.

In the second period of the CANCERDIP project we performed the bulk of the MethylCap-seq experiments on colorectal tumors, matched normal colon tissues, and AMLs necessary to identify

differentially methylated regions (DMRs) for these important tumors. The samples were obtained from partner 1 and 4. In total, we have MethylCap-seq data for 18 leukemia and 3 CD34+ samples and 30 matched pairs of colon tumor and normal tissue.

Bio-informatic analysis of the large amount of sequencing data was the next step in the project. Genome-wide profiles where generated for all samples and compared using hierarchical clustering and calling of DMRs (deliverable 1.3 and 1.4). Bio-informatic analysis was done in close contact with partner 5.

For the colon samples, hierarchical clustering of the full methylation profiles clearly separates most tumors from normal tissue. There were many differentially methylated regions (DMRs) per sample pair analysed. We selected 3338 DMRs: 490 high confidence hypo- and 2848 hypermethylated regions. Several of these DMRs where confirmed by conventional bisulphite sequencing and we determined which genes where associated to the DMRs. There are many genes with aberrant methylation that were also previously identified by others and many new genes. For example we detected highly frequent hypermethylation of the gene CSPG2/VCAN. Only recently, partner 1 also identified this gene as being hypermethylated in colon cancer, but not in other cancers (Fernandez A.F. et al 2011. Genome Research). This gene encodes a major component of the extracellular matrix.

For the leukemia samples we did similar analysis of the MethylCap-seq data. When we analysed where the DMRs are in the genome, we see that frequent hypermethylation in leukemia occurs mainly at CGI, whereas hypomethylation occurs mainly at intergenic regions. We selected regions that are a DMR in more than 3 leukemia samples (n=3675), and will continue our analysis with these.

On-going analyses on the DMRs of both colon cancer and leukemia are focused on splitting up the DMRs into different types and ultimately select particular interesting candidates for further validation in another sample set.

Work Package 2.

TASK 2.1. Analysis of the molecular mechanism implicated in targeting DNA methylation to genes that become silenced in cancer. Role of Polycomb group proteins.

D2.1. Report on the mechanistic connections between Polycomb group proteins and methyl-CpG binding domain proteins.

The essential epigenetic systems involved in heritable repression of gene activity are Polycomb group (PcG) proteins and DNA methylation, silencing pathways shown to be mechanistically linked. Collaborative work between partners Dr. Fuks and Dr. Esteller suggests that the pre-marking of certain genes leads to hypermethylation due to crosstalk between PRC2 components and DNA methyltransferases (DNMTs). In this way, normal repression of these genes in precursor cells can be aberrantly converted to a permanently silenced state, which prevents full differentiation and predisposes cells to become cancerous. DNMTs are likely to be recruited by PRC2 pre-marking in an early event en route to oncogenesis, when cancer cells still show a stem-cell-like epigenetic signature.

Recently, partners Dr. Fuks and Dr. Esteller have reported a previously unrecognized direct connection between Polycomb proteins and DNA methylation. They found that EZH2, in the context of the PRC2/3 complexes, controls DNA methylation through direct association with DNA methyltransferases. Binding of DNMTs to several EZH2-repressed genes depends on the presence of EZH2 and moreover EZH2 is required for DNA methylation of EZH2-target promoters. Thus, EZH2 serves as a recruitment platform for DNA methyltransferases, fact that highlights a direct connection between the two key epigenetic repression systems.

On our more recent work, we demonstrate that EZH2 interacts physically with the methyl-CpG-binding protein MeCP2 in vivo. In cells with reduced MeCP2 or EZH2 levels, we observe reactivation of several

target genes. In addition, we show that EZH2 promoter binding coincides with the presence of MeCP2 and that the latter is required for promoter occupancy of EZH2 and for H3K27 trimethylation. Further, we find that binding of EZH2 to Bdnf, a natural MeCP2-target gene, is significantly impaired in depolarized primary cortical neurons as well as in brain tissue from a Mecp2-knockout mouse. Our results indicate that MeCP2 might act as a mechanistic bridge between DNA methylation and PcG proteins, and could thereby reinforce the repressive function of these two distinct epigenetic systems.

To finish this work, ChIP in cancer cells (U2OS) by partner Dr.Fuks and in wt ES cells or MeCP2 knock-out cells by Dr. Esteller are being done. In addition, Dr. Bock has helped to perform a bioinformatic comparison of known MeCP2 targets and H3K27me3. Bioinformatic analysis of existing MeCP2 ChIP-chip (Chahrour et al., 2008) and H3K27me3 ChIPSeq (Barski et al., 2007) suggests that human promoter regions of genes whose mouse ortholog is negatively regulated / repressed by MeCP2 (according to Chahrour et al., 2008) are significantly enriched for H3K27me3 and - to lesser extent - for H3K9me9, while being depleted for PolII binding

TASK 2.2. Study of the mechanisms that trigger Polycomb-mediated abnormal DNA methylation in cancer.

D2.2. Identification of novel mechanisms that regulate DNA methylation and their impacts on cancerogenesis.

In order to identify novel mechanisms that could regulate DNA methylation and that might impact on cancerogenesis, we tackled following issues:

1. Regulation of DNA methylation patterns by CK2-mediated phosphorylation of Dnmt3a.

An article describing the results below is in revision: Deplus R, Blanchon L, Rajavelu A, Boukaba H, Defrance M, Luciani J, Dedeurwaerder S, Brinkman AB, Simmer F, Müller F, Berdasco M, Putmans P, Calonne E, Litchfield DW, de Launoit Y, Jurkowski TP, Stunnenberg HG, Bock C, Fraga MF, Esteller M, Jeltsch A, Fuks F. Regulation of DNA methylation patterns by CK2-mediated phosphorylation of DNMT3A. Nature Cell Biology. (In revision).

The de novo DNA methyltransferases are responsible for generating genomic methylation patterns, but the underlying mechanisms are still poorly understood. We show that Dnmt3a phosphorylation by the CK2 protein kinase regulates the establishment of DNA methylation patterns. We find that Dnmt3a is phosphorylated by CK2 at two key residues located near its PWWP domain.

In collaboration with Dr. Esteller, we have performed quantitative analysis of global CpG methylation in CK2 Tet-Off cells. We show that CK2-mediated decrease of global DNA methylation is dependent on DNMT3a as well as on DNMT3b.

Together with Dr. Stunnenberg and Dr. Bock, by genome-wide DNA methylation analysis in CK2-depleted cells, we reveal that CK2 primarily affects CpG methylation of several heterochromatin repeats as well as Alu elements.

We also find that CK2-mediated phosphorylation is required for proper localization of Dnmt3a to heterochromatin. By revealing phosphorylation as a new mode of regulation of de novo DNA methyltransferase function and by uncovering a previously unrecognized mechanism for the regulation of methylation at repetitive elements, our results shed new light on the origin of DNA methylation patterns.

2. The interplay between DNA methyltransferases and Lysine specific demethylase 1 (LSD1) in cancer cells is cell cycle dependent.

An article describing the results below is in preparation and includes the following CancerDIP PIs: Hendrik G. Stunnenberg, Christoph Bock, Manel Esteller, and François Fuks.

DNA methylation and histone modifications are key epigenetic regulators of gene expression and they are functionally linked. DNA methyltransferases (DNMTs) and histone demethylases are significantly upregulated in cancer cells. Here we explored interactions/investigated the mechanistic link between lysine specific demethylase LSD1, (also known as KDM1A, AOF2) and DNA methyltransferase activity in cancer cells. We show that LSD1 interacts with both the main maintenance DNA methyltransferase DNMT1 and the de novo DNA methyltransferase DNMT3B, in several cancer cell lines.

DNMT1 residues encompassing the Zn motif, the Cys-rich domain as well as the catalytic unit preferentially interacted with LSD1; while the conserved PWWP domain of DNMT3B had a stronger association with LSD1. Notably, LSD1 and DNMT1 co-localized in mid-late S-phase around heterochromatin and using the proximity ligation in situ assay (P-LISA), we were able to monitor and quantify the endogenous DNMT1/LSD1 interaction in glioma cells revealing a substantial increase during S phase.

Intriguingly, depleting LSD1 in cancer cells did not trigger any global methylation changes nor any reduction in the protein levels of DNMT1 and DNMT3B, unlike in ES cells.

LSD1 target genes were however upregulated when cancer cells were treated with a DNA demethylating agent.

Our findings suggest that the LSD1/DNMT1 interaction observed in cancer cells has an important role in the replication of the epigenome during mitosis by regulating (the stability of) DNMT1. These results further reflect the importance of targeting the epigenetic machinery for more effective therapy in cancer.

3. Expression of DNMT3b is post-transcriptionally regulated by HuR.

Dr. Esteller's group has demonstrated that expression of DNMT3b is post-transcriptionally regulated by HuR, an RNAbinding protein that stabilizes and/or modulates the translation of target mRNAs. The interaction between HuR and DNMT3b mRNA has been studied by immunoprecipitation of endogenous HuR ribonucleoprotein complexes followed by RT-qPCR detection of DNMT3b mRNA, and by in vitro pulldown of biotinylated DNMT3b RNAs followed by western blotting detection of HuR. These studies reveal that binding of HuR stabilizes the DNMT3b mRNA and increases DNMT3b expression. Unexpectedly, cisplatin treatment triggers the dissociation of the [HuR-DNMT3b mRNA] complex, in turn promoting DNMT3b mRNA decay, decreasing DNMT3b abundance, and lowering the methylation of repeated sequences and global DNA methylation (López de Silanes et al., 2009).

TASK 2.3. Analysis of the Polycomb implication in DNA methylation at a genome-wide level: ChIP-on-chip analysis.

D2.3. Genome-wide scale profiling and overlap with PcG proteins and proteins of the DNA methylation machinery.

As mentioned above, our results suggest that MeCP2 might act as a mechanistic bridge between DNA methylation and PcG proteins, and could thereby reinforce the repressive function of these two distinct epigenetic systems.

We wished to investigate whether the connection between DNMTs/MBDs and PcG proteins occurs at a genome-wide level. We also performed numerous ChIP-on-chip experiments, in collaboration with Dr. Esteller, using antibodies against MeCP2, EZH2 as well as anti-trimethyl K27 H3 antibodies in several cellular models. Unfortunately, despite extensive efforts, no conclusive data could be obtained. It is possible that better antibodies, in particular for MeCP2 and EZH2 might be needed. Another possibility

could to perform ChIPSeq using the above-mentioned and/or novel antibodies. So far, we have not been able to obtain good enough MeCP2 antibody that work in ChIPSeq.

As an alternative to this part of the project, we looked further into the DNA methylome in cancer. Indeed, having a deep interest in exploring DNA methylation changes in cancers, we subsequently used a technology developed during the course of the CancerDIP project, termed Infinium Methylation 27. DNA methylation profiling reveals a predominant immune component in breast cancers. An article describing the results below is under submission.

Previous studies have documented aberrant methylation events in breast carcinogenesis and it was notably found that specific DNA methylation patterns can be related to some of the known «expression breast cancer subtypes». However such patterns have not been precisely related to novel and specific tumor traits. Our goal was to explore the DNA methylation landscapes of phenotypically heterogeneous tumors, to relate this diversity to landscape features, and extract novel biological and clinical meaningful information.

In this study, using Illumina's' Infinium methylation 27K platform, we have uncovered novel and precise epigenetic portraits in breast cancer, highlighting a key contribution of the DNA methylome to the complexity of the disease. Further, one of the major new finding of the present study is that we showed for the first time that DNA methylation profiles can reflect the cell-type composition of the tumor microenvironment, and in particular a T lymphocyte infiltration of these tumors. Interestingly, we found immune components as good markers of breast cancer clinical outcome.

Breast cancer is a molecularly, biologically and clinically heterogeneous group of disorders. Understanding this diversity is essential to improving diagnosis and optimizing treatment. Both genetic and acquired epigenetic abnormalities participate in cancer, but the involvement of the epigenome in breast cancer and its contribution to the complexity of the disease are still poorly understood. By means of DNA methylation profiling of 248 breast tissues, we have highlighted the existence of previously unrecognized breast cancer groups that go beyond the currently known "expression subtypes" (Figure 15). Interestingly, we showed that DNA methylation profiling can reflect the cell type composition of the tumor microenvironment, and in particular a T lymphocyte infiltration of the tumors. Further, we highlighted a set of immune genes having high prognostic value in specific tumor categories. The immune component uncovered here by DNA methylation profiles provides a new perspective for the importance of the microenvironment in breast cancer, holding implications for better management of breast cancer patients.

TASK 2.4. Generation and characterization of ChIP and ChIP-on-chip antibodies for the study of Polycomb group protein members.

D2.4. Generation of antibodies against Polycomb group proteins, proteins of the DNA methylation machinery and histone modifications.

Analysis of the molecular mechanisms of gene silencing in cancer requires ChIP and ChIP-on-chip grade antibodies. Commercial antibodies available in the market often show non- specific background that is not compatible with a genome-wide analysis. Thus, the generation of good quality antibodies was of prime importance to ensure the successful accomplishment of the project. As the immunogenicity of peptides can differ considerably, several peptides per target protein have been designed and further used for immunization. Antibody generation was performed according to the developed pipeline for the generation of rabbit polyclonal antibodies. Each peptide was injected into two SPF rabbits (New-Zealand white) and each rabbit was injected 6 times. Six bleeds were taken from both rabbits: a small bleed of 2 ml at Day 0, which served as a negative control; a second small bleed at Day 38, which was used to monitor the immune response; 3 large bleeds of 20 ml at Day 66, Day 87 and Month 4; and a final bleed of 50 ml after 4.5 months. The 4 large bleeds of these rabbits were sent to partner 1 for external characterization and/or further distribution among the other interested partners. During the second reporting period a total of 264 bleeds corresponding to 33 targets were sent for characterization.

The antibodies produced within the CancerDIP project are summarized in a table within the whole final report (attached)

Work Package 3.

TASK 3.1. MeDIP analysis in acute myeloid leukemia cell lines.

D3.1. Global DNA methylation profile in NB4, U937, KG1a and HL60 cells; relevance of the fusion protein in respect to normal kariotype. (15 Month)

Dr. Altucci's group has been performing DMH assays and has defined novel DMRs in leukemia cell lines. The experiments have been carried out in NB4, U937 cells and K562 cells. These models are all different models of acute myeloid leukemias (AMLs). The DMRs obtained have been divided into 'known' and 'novel' DMRs. The most interesting have been confirmed by bisulphite sequencing.

TASK 3.2. Generation of DNA methylation signatures in acute myeloid leukemias.

D3.2. Generation of a list of methylated genes of AML patients blasts; relevance of the fusion protein in respect to normal kariotype. (15 Month)

TASK 3.3. A comparison of the patterns of DNA methylation between AML patients versus normal individuals.

D3.3. Differential profile of CD34+ vs AML cells; candidate loci causal for acute myeloid leukemias. (19 Month)

With the aim to decipher new and known methylation patterns, as well as targets in leukemias and representative cell lines, the MeDIP technology has been employed. This tool allows us the screening of a large set of samples with more probability to obtain candidates causal for acute myeloid leukemias (AMLs). We have selected a set of differentially methylated regions (DMRs) that have been confirmed by bisulfite sequencing in the same sample, as well as in novel CD34+ and AML samples. One of the strongest candidates confirmed by bisulfite sequencing is the WT1 gene, which has been recently described as differentially expressed in AMLs samples. We also have analysed the expression of 5 other candidate genes in 10 additional AML patients and CD34+ normal progenitors, detecting differences in the expression levels of the tested genes. Suggesting then that DNA methylation is responsible for their gene expression alteration and may contribute to tumorigenesis.

We decided to go further with a higher resolution analysis as the enabling technology of MethylCap that allows generation of whole-genome methylation profiles. In this study we have obtained the methylation profile of 18 patient blast samples and three CD34+ samples. A total of 17,057 DMRs were detected with high confidence, 99 of these DMRs occurred within at least 10 samples. Those 99 DMRs were considered as AML-common DMRs, they include hyper-DMRs, with higher levels of methylation; or hypo-DMRS, which have lost methylation relative to the controls. According to their genomic location, we could assign 40 hyper- and 26 hypo-DMRs that were subjected to gene-enrichment analysis, as it is shown in figure 16. Interestingly, WT1 was present within the set of 40 hyper-DMRs, providing platform-independent confirmation for its importance in AML aetiology. The most significantly enriched genes among the hyper-DMRs are the cAMP responsive element binding protein 1 (CREB1) and the binding sites for homeobox 1 (HMX1). However, among the hypo-DMRs the mostly represented were related to signalling functions; thromboxane, ADP, and thrombin signalling, and also signal amplification and G alpha signalling.

D3.4. Mechanistic study of the methylation mechanisms on methylated loci. (36 Month)

In the first period, SUNAP has performed differential methylation analyses (DMH) to compare the methylation status between AMLs and normal progenitors (CD34+) not only in cell lines but also in primary samples and in normal progenitors. These analyses have been done on a total of 25 samples as differential methylation study (DMH). We selected 15 novel candidates that have been confirmed by bisulfite sequencing.

In the second period we analysed mechanisms that account for the identified DMRs in the MethylCap profile. We analyzed which regions in the genome are most susceptible to differential methylation and we found that hyper-DMRs occur mainly within CpG islands, either overlapping or non-overlapping with transcription start sites, being strongest for AML-common DMRs. On the contrary, the genomic location of DNA methylation loss (hypo-DMRs) was less specific, and occurred most frequently in exons, introns, and intergenic regions.

TASK 3.4. Correlation of molecular data (DNA methylation profiles) with clinical parameters (outcome).

D3.5. Correlation analysis between molecular data and clinical outcome. (24 Month)

During the first period we were putting in correlation the karyotype, the morphology and the differentiation with the clinical outcome (also considering the gender, the age and possible, independent pathologies). An interesting preliminary observation came from the analysis of the methylation and expression level of the candidate genes in primary sample 112. In this case of very aggressive AML, the sample was taken before treatment (at the diagnosis of the disease). Given that the patient (ended later on) relapsed, another sample was taken at the time of the relapse before starting the novel treatment. As shown in figure 19 of this report the expression level of all candidate genes is further decreased thus suggesting a potential prognostic role of these candidates both for expression level and DNA methylation.

In the second period and after the analysis of AML patient versus normal samples by MethylCap, we have correlated the results with the presence of specific chromosomal translocations. Then, the Kariotype, immunophenotype and specific alterations/samples (such as inversions or fusion proteins) have been characterised.

Work Package 4.

TASK 4.1. Identification of new epigenetically altered genes involved in colorectal tumors.

D4.1. Standardize expression arrays for the in vitro models, collected tissues and the ex vivo models. (Month 12)
D4.2. Selection of a list of candidate genes involved in cancer after compilation of all expression arrays. (Month 16)

With the aim to analyze the DNA methylation differences associated with colorectal tumors we have applied the MeDIP approach to a 44K human proximal promoter array to evaluate the CpG hypomethylation changes in the DNMT1/DNMT3b double knockout HCT-116 cells (DKO) in relation to the wild-type HCT-116. From the global genomic perspective, we observed abundant DNA demethylation events in DKO cells in comparison with wild-type HCT-116 cells. Of the 17,917 printed promoters in the array, we observed in DKO significant hypomethylation in 126 candidate genes.

As an alternative approach to the MeDIP technique, HCT-116, DKO, SW480, SW620, HCT115, LoVo, SW48, Co115 and RKO cell lines and paired samples normal versus tumor tissue from colon cancer patients have been studied using the Illumina Array. For those candidate genes and DNA sequences with relevant differences, we have proceeded with bisulfite sequencing to further characterize their DNA methylation status.

In the second period, we have decided to use the Golden- Gate DNA methylation BeadArray (Illumina, Inc.) technique since it allows an extremely comprehensive study of a human sample population with an intermediate level of resolution of CpGs at the genomic level. Indeed, the analysis of the colon cancer samples is part of a study that has obtained the DNA methylation fingerprint of 1628 human samples. The array interrogated 1505 CpG sites, located from 1500 bp to +500 bp around the transcription start sites of 808 selected genes. This study have been developed in collaboration with Dr. Christoph Bock group, which has been responsible of the computational epigenetics and the genome-scale integrative analyses, resulting in a recent publication, "A DNA methylation fingerprint of 1628 human samples" (Fernandez et al, Genome Research 2011). We have studied the DNA methylation profile of 97 colon mucosa samples (normal), and 110 colon cancer samples. The results show a tumor-type-specific profile characterized by a progressive gain of CpG methylation within CpG-island-associated promoters, and a cumulative loss of CpG methylation outside CpG islands in the different steps of tumorigenesis. Then each type of tumour samples has its own aberrant DNA methylation fingerprint.

TASK 4.2. DNA methylation status analysis by bisulfite genomic sequencing.

D4.3. Correlation between molecular data and clinical outcome. (Month 18)
D4.4. Analysis of the methylation status of the promoters by bisulfite genomic sequencing. (Month 22)

During the first period, the expression microarrays and Illumina Arrays data were crossed, compared and analyzed. A list of candidate genes was generated. So far, 7 genes have been tested with different approaches (Real Time Quantitative PCR, Methylation Specific PCR (MSP), Bisulfite Genomic Sequencing) in order to prove their reliability as trustful biomarkers: DRD5, ABCG2, CDH11, CDH13, GATA4, CA4 and APOE. Two of the genes have already been discarded: ABCG2 as it appears to be unmethylated in HCT-116 cell line and DRD5 as it has tissue specific methylation in colon.

Once generated a list of candidate genes from the DNA methylation profiles, we have analysed their methylation status by using bisulfite genomic sequencing and pyrosequencing. Validated DNA methylation profile data will be studied in WP6 for their potential use as biomarkers with clinical applications. In the first study we have validated two genes that are hypermethylated in colon cancer samples versus normal samples, the NPY and GSTM2 genes. We also have validated two other genes that are hypermethylated in colon-brain metastases, such as CD40 and SLC5A8.

We considered of interest to perform bisulfite genomic sequencing of miRNA genes with tumor suppressor features, such as miR-9-1, miR-9-2, miR-9-3, miR-148a, miR-34b/c, miR-124a1, miR-10a, whose expression has been previously reported to be silenced, specifically in cancer. Table 4 shows a summary of the DNA methylation status of these miRNAs in different colon cancer cell lines. The results demonstrate that promoters of these miRNAs are methylated in all cancer cells but unmethylated in normal colon.

TASK 4.3. Generation of profiles of histone modifications in colon cancer cells.

D4.5. Generation of profiles of histone modifications and chromatin proteins by ChIP-on-chip. (Month 36)

Dr. Esteller's group has analyzed the methylation status of the three lysine residues on histone H3 (K4, K36 and K79), known to be associated with active transcription, in the context of tumor suppressor genes that become hypermethylated in cancer cells. It was also been studied whether these marks are reversed after treatment with DNMTs inhibitors that result in pharmacologically induced transcriptional reactivation. We have found that the active histone lysine methylation signature present at the promoter of unmethylated tumor suppressor genes, namely H3K4me3 and H3K79me2, is permanently disrupted when the gene becomes epigenetically silenced by promoter hypermethylation, because after DNA demethylation at these promoters only H3K4me3, but not H3K79me2, is re-established (Jacinto FV et al. Oncogene, 2009).

Recently, accumulating evidence demonstrates altered miRNA expression profiles in many types of cancer, linking these molecules to the process of carcinogenesis. Recent studies have shown that during malignant transformation normal cells accumulate aberrant epigenetic changes, such as hyper-methylation of the promoter-related CpG islands, which is associated with specific histone modifications, including dimethylation of histone H3 at lysine 9, deacetylation at this same residue, trimethylation of histone H3 at lysine 27 and loss of the transcriptional activating mark H3K4me3, resulting in the silencing of those genes (Ballestar, E. and Esteller, M. Carcinogenesis, 2002) (Herman, J.G. and Baylin, S.B., New Eng J Med 2003). Under this scope, we were interested in analyzing the epigenetic landscape of the promoters of miRNA genes in colon cancer versus normal cells.

The table below summarizes the histone modification patterns found in the samples analyzed. As it is expected, the active gene GADPH contains high levels of the active mark H3K4 trimethylation, and the repressed genes SAT2 and MYT-1 have high levels of the repressed marks, H3K9 and H3K27 trimethylation in normal colon and colon cancer cell lines. The type of histone modification marks present in miRNAs is however very heterogenous. The miRNAs that have been analyzed are unmethylated in normal colon but methylated in all cancer cell lines included in the study. Apparently there is not a correlation between DNA methylation and histone modification profile, and the histone modification profile is mainly depending on cell and miRNA type, probably as a consequence of the grade of tissue development.

Deviation from plans

The genome wide chromatin and histone modifications profile by using ChIP-on-chip methodology has not been carried out for different technical and scientific reasons. The antibodies that have been prepared were not good enough for ChIP analysis. In addition, the ChIP-on-chip methodology has been displaced for the ChIP-seq technology, which has a deeper genome wide resolution. Thus, we decide to perform histone modification analysis of miRNAs in order to correlate the histone marks promoter occupancy with DNA methylation and tumorigenesis.

Work Package 5.

TASK 5.1. Development of a kit format of the MeDIP assay.

At the time the CancerDIP project started, a new technique to study DNA methylation (methylated DNA immunoprecipitation (MeDIP)) had been described (Weber et al., 2005). Basically, an antibody specific for methylated cytosines is used to immunoprecipitate methylated DNA previously randomly fragmented. The resulting enrichment in the immunoprecipitated fraction can be determined by PCR or massive parallel sequencing for genome-wide analysis.

Based on this principle, and thanks to fruitful interactions between the different partners of the CancerDIP consortium, Diagenode has developed and commercialized two kits: MeDIP and MagMeDIP kits, allowing the analysis of DNA methylation. Basic features/concepts of this kit are described below. A fully detailed and updated protocol can be found at: http://www.diagenode.com/media/catalog/file/Manual_MeDIP-kit_012010V4.pdf. Partners and customers are successfully using it and always give positive feedback about their reproducibility and sensitivity.

TASK 5.2. Kit validation.

D5.1. Development of a kit format and kit validation on clinical samples.

Our antibody against 5-mehylcytidine used in the MeDIP kit ensures the high specificity of results. This kit includes positive and negative internal controls. In fact, the immunoprecipitation is performed in the presence of fully methylated (positive internal control) or unmethylated (negative internal control) BAC DNAs from Arabidopsis thaliana. Two sets of primers are used to amplify DNA from either positive or

negative internal controls. Consequently, the efficiency of immunoprecipitaion can be monitored for each experiment. On the other hand, samples used for MeDIP can be controlled by amplifying DNA sequences that are known to be either methylated or unmethylated. For this purpose, the primers specific to human GAPDH promoter and X-linked ?-satellite DNA are included in the kit. GAPDH is a housekeeping gene, ubiquitously expressed in the cell and its promoter is unmethylated. X-linked ?-satellite DNA is a highly repetitive methylated DNA from centromere regions. Thus, the simultaneous utilization of internal controls together with amplification of immunoprecipitated DNA using primers for GAPDH and X-linked ?-satellite DNA guarantee the reliable results.

Typical results obtained with our MeDIP kit are presented in Figure N. Genomic DNA from NB4 cells was immunoprecipitated with an antibody against 5-mehylcytidine. The immunoprecipitation is carried out in the presence of positive and negative internal controls (corresponding to fully methylated or unmethylated BAC DNA from Arabidopsis thaliana DNA, respectively). Both immunoprecipitated DNA and Input sample were analysed by qRT-PCR. The methylation status of GAPDH promoter and X-linked ?-satellite DNA was compared. The efficiency of immunoprecipitation is assessed by amplification of internal positive (meDNA) or negative (unDNA DNA of) controls. Two sets of primers (control 1 and 2) are used for each control. As expected, significant enrichment for methylated DNA is observed in positive control comparing to negative control. No enrichment was found for GAPDH promoter reflecting its active state. Meanwhile, X-linked ?-satellite DNA shows a high level of enrichment that corresponds to its high methylation.

Although the MeDIP kit gives reliable and specific results, a simplification of the protocol at different steps became possible over time. The development of a quicker and simplier MeDIP technique was of prime importance to facilitate and accelerate the large-scale analysis of DNA methylation pattern. Diagenode has worked on the development of a new simplified MeDIP kit using magnetic beads for immunoprecipitation. The use of magnetic beads is not only easier but it also increases the sensitivity of the method that is an important issue in the case of rare and small biopsies sample analysis from patients. From these experiments, Diagenode launched a new kit on the market (MagMedip kit) whose fully updated protocol can be found at http://www.diagenode.com/media/catalog/file/MA_MME-V1_10_08_10.pdf.

Firstly, we compared the efficiency of DNA recovery using sepharose-protein A/G coated beads versus magnetic beads. Different types of magnetic beads were tested: protein A/G or protein G coupled beads, or magnetic beads coated with Sheep anti-mouse IgG (IgG beads). So far, the higher recovery of methylated DNA was found using the IgG magnetic beads.

We next tested how beads volume used in the assay influences the methylated DNA recovery. MeDIPs with 20, 10 and 5 µl of IgG magnetic beads were run. No difference was found between 20 and 10 µl of magnetic beads. DNA recovery is slightly lower with 5 µl of beads. Thus, the volume of 10 µl could be used in the new kit format providing the same recovery rate.

We further carried out some experiments in order to simplify washing steps. Our tests indicate that only two washing buffers can be used instead of four originally included in the current MeDIP kit (data no shown). The volume of washing buffers can be reduced from 400 µl (the current MeDIP kit) to 100 µl without recovery loss (data no shown). The introduction of this modification will reduce the duration of experiment as well as a cost of kit production.

The phenol/chloroform extraction of DNA is one of the steps in the current MeDIP kit which is toxic and requires a fume-hood. Moreover, this step is time-consuming as several transfers of samples are needed. Furthermore, it might lead to a loss of DNA. Non toxic and rapid elution of DNA without sample transfer is proposed. In fact, DNA can be eluted from magnetic bead using an elution buffer with high pH. Eluted DNA is directly analyzed by PCR. Thus, a simple elution with high pH buffer seems to be efficient that might significantly improve the performance of MeDIP kit.

Thus, the following improvements have been introduced in the MagMeDIP kit:

1. The number of spike controls and corresponding primer sets were reduced. In the original kit, two spike controls were used for positive methylated DNA and negative unmethylated DNA, resulting in redundant information. The use of only one positive and one negative spike controls (and corresponding primer sets) allows sufficient monitoring of immunoprecipitation and clear results interpretation. Moreover, this improvement reduced the production cost of the kit.
2. A new primer set for some methylated region in the human genome was also introduced. A region in X-linked a-satellites DNA was initially used as positive locus. This was a repetitive sequence, thereby making the primers very sensitive to contamination. We were looking for an alternative locus in which there were no repetitive regions and which was ubiquitously methylated in the human genome. We eventually tested and selected a CpG region of the human TSH2B gene. The new primers were validated on different cell types.
3. We also introduced a purification step after immunoprecipitation using magnetic beads. A new Ipure kit was specially developed and optimized for the purification of DNA after MeDIP kit's use (http://www.diagenode.com/en/catalog/kits-2/dna-purification-50/product/ipure-kit-x100-353). The introduction of the Ipure purification benefits from many features. The method provides pure DNA for any downstream application (e.g. next generation sequencing), does not use any toxic reagents (unlike phenol/chloroform). The purification results in much higher yield than with column-based protocols and can be used for recovery of small amounts of DNA (Figure 34). Moreover, magnetic beads purification allowed the automation of the MagMeDIP kit.

Finally, during the second period, Diagenode has also adapted the MagMeDIP kit for it to be used on Diagenode SX-8G IP-Star and SX-8G IP-Star Compact automated platforms (http://www.diagenode.com/en/catalog/kits-2/dna-methylation-11/product/auto-medip-kit-7).
Diagenode's Automated Systems are automated bench-top instruments that standardize different epigenetic applications (i.e. ChIP, MeDIP or MethylCap). Diagenode designed these automation systems to make ChIP and DNA methylation studies accessible and reproducible, and ensure consistent data in every experiment. The process of transferring the MagMeDIP kit from the bench to Diagenode automated platforms has required some extensive knowledge in the software of the automated platforms in addition to an extensive process of parameters optimization. The Auto MeDIP now provides many advantages to the research community because it requires minimal operator intervention (reducing "hands-on" time) and allows researchers to run more assays. In fact, with the Auto MeDIP kit the researcher can process up to 16 samples per run in less than 10 hours. Auto MeDIP kit increases as well the reproducibility between different experiments allowing reliable comparison of data from lab to lab.

Work Package 6.

TASK 6.1. Integrative epigenomic analysis of the DNA methylation machinery.

D6.2. Summary of the integrative epigenome analysis.

The interplay between several chromatin-modifying proteins (MeCP2, LSD1, DNMTs, Polycomb proteins) and DNA methylation has been analyzed. To that end, bioinformatic pipelines have been developed for a number of important experimental methods, including ChIP-on-chip with NimbleGen and Agilent tiling microarrays, MeDIP on the same platforms, ChIP-seq for histone modifications, and Infinium DNA methylation analysis. These data analysis pipelines have been completed and are now routinely used.

The data obtained from this study supported a model of epigenetic repression that may help explain how specific promoter regions become aberrantly silenced in cancer cells. Current results indicate that the methyl-binding protein MeCP2 can, in turn, recruit Polycomb binding. Hence, these two mechanisms could give rise to a self-propagating feedback loop enforcing long-term transcriptional repression of specific genes. Hence, defects and de-regulation of the proteins involved in these two mechanisms of induced epigenetic repression are prime candidates for a causal role in aberrant DNA methylation.

TASK 6.2. Ranking of differentially methylated regions by their potential as diagnostic biomarkers.

D6.1. Report on the completed EpiGRAPH version for the CancerDIP project.
D6.5. Report on the completed MethMarker software for the CancerDIP project.

A version of EpiGRAPH has been adapted to CancerDIP project needs. The EpiGRAPH software analysis (http://epigraph.mpi-inf.mpg.de/) performs candidate biomarker prioritization by machine learning algorithms. The user uploads a set of genomic regions that exhibit DNA methylation alterations in cancer. Next, EpiGRAPH performs data mining on multiple types of genomic attributes (including DNA sequence and predicted structure, distribution of genes and presence or absence of evolutionary conserved elements) to identify properties that co-locate with cancer-specific alterations throughout the genome. Subsequently, a support vector machine is trained and its prediction score is used to prioritize candidate biomarkers for experimental follow-up (Bock, C., Halachev, K. et al., 2009; Bock, C., Kuster, G.V. et al., 2009).

Other computational algorithms and software tools for ranking biomarkers from large-scale DNA methylation data have been developed. The developed algorithms are implemented in the freely available MethMarker software package (http://methmarker.mpi-inf.mpg.de/) (Schüffler et al., 2009) The MethMarker software implements a systematic approach to the optimization and validation of DNA methylation biomarkers. The user has performs bisulfite sequencing for a number of representative tumor and control samples, which and imports them from BiQ Analyzer into MethMarker. Based on these data, MethMarker identifies experimental assays that are highly consistent with the results of bisulfite sequencing but substantially more cost-efficient and time-saving. Furthermore, MethMarker uses clinical information to build logistic regression models that translate measurement data obtained with the optimized assay into a patient-specific diagnosis.

Moreover, the first software package for graphical processing of locus-specific high-throughput DNA methylation data has been developed. The tool is named "BiQ Analyzer HT" and is freely available from http://biq-analyzer-ht.bioinf.mpi-inf.mpg.de/ (Lutsik et al., 2011).

TASK 6.3. Optimization and validation of selected candidate biomarkers.

D6.3. Summary of the candidate biomarker ranking.
D6.4. Biostatistical summary and validation report covering all biomarkers discovered by CancerDIP.

A software package developed for identifying and ranking candidate biomarkers from DNA methylation profiles of clinical samples has been tested in a benchmarking study comparing the DNA methylation profiles of four samples (including a colon tumor and matched normal colon tissue) obtained by four different methods for DNA methylation profiling (Bock et al., 2010). The software package described in the latter paper is currently used to process MethylCap data from a larger number of colon cancer samples.

A review summarizing the CancerDIP approach to epigenetic biomarker discovery has also been published (Bock, 2009).

Moreover 1,628 samples were analyzed, including the development and validation of biomarkers that accurately predict the tissue characteristics of "cancers of unknown primary origin" (CUPs) (Fernandez et al., 2011).

Bibliography:

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., Zhao, K. 2007. High-resolution profiling of histone methylations in the human genome. Cell. 129(4):823-37.

Bock, C. 2009. Epigenetic biomarker development, Epigenomics 1, 99-110.

Bock, C., Kuster, G.V., Halachev, K., Taylor, J., Nekrutenko, A. Lengauer, T. 2009. Web-based analysis of (epi-) genome data using EpiGRAPH and Galaxy. Methods in Molecular Biology.

Bock, C., Halachev, K., Büch, J., Lengauer, T. 2009. EpiGRAPH: User-friendly software for statistical analysis and prediction of (epi-) genomic data. Genome Biol. 10: R14.

Bock, C., Tomazou, E.M., Brinkman, A., Müller, F., Simmer, F., Gu, H., Jäger, N., Gnirke, A., Stunnenberg, H.G., Meissner, A. 2010. Genome-wide mapping of DNA methylation: a quantitative technology comparison. Nature Biotech. 28:1106-14.

Chahrour, M., Jung, S.Y., Shaw, C., Zhou, X., Wong, S.T., Qin, J., Zoghbi, H.Y. 2008. MeCP2, a key contributor to neurological disease, activates and represses transcription. Science. 320(5880):1224-9.

Fernandez, A. F., Assenov, Y., Martin-Subero, J., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A. C., Galm, O., Ferrer, I., Sanchez-Cespedes, M., Villanueva, A., Carmona, J., Sanchez-Mut, J. V., Berdasco, M., Moreno, V., Capella, G., Monk, D., Ballestar, E., Ropero, S., Martinez, R., Sanchez-Carbayo, M., Prosper, F., Agirre, X., Fraga, M. F., Grana, O., Perez-Jurado, L., Mora, J., Puig, S., Prat, J., Badimon, L., Puca, A. A., Meltzer, S. J., Lengauer, T., Bridgewater, J., Bock, C., and Esteller, M. 2011. A DNA methylation fingerprint of 1,628 human samples. Genome Res.

López de Silanes, I., Gorospe, M., Taniguchi, H., Abdelmohsen, K., Srikantan, S., Alaminos, M., Berdasco, M., Urdinguio, R.G., Fraga, M.F., Jacinto, F.V., Esteller, M. 2009. The RNA-binding protein HuR regulates DNA methylation through stabilization of DNMT3b mRNA. Nucleic Acids Res. 37(8):2658-71.

Lutsik, P., Feuerbach, L., Arand, J., Lengauer, T., Walter, J., Bock C. 2011. BiQ Analyzer HT: locus-specific analysis of DNA methylation by high-throughput bisulfite sequencing. Nucleic Acids Res. (Web Server issue):W551-6.

Schüffler, P., Mikeska, T., Waha, A., Lengauer, T., Bock, C. 2009. MethMarker: User-friendly design and optimization of gene-specific DNA methylation assays. Genome Biol. 10, R105.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., Schübeler, D. 2005

Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nature Genet. 37:853-62.

Potential Impact:

The main dissemination activities and exploitation of results of the CancerDIP project are mainly scientific publications, conferences and workshops. There have been already published around 40 publications in scientific journals in the frame of CancerDIP during these three years. Their results have also been exposed in about 40 conferences from all the participants, as it is shown in the attached list. With the aim to disseminate the results from the CancerDIP project, as well as to invite and share these results with the prominent researchers in DNA methylation profiling and cancer epigenetics, two public workshops have been organised in Barcelona, the first was the "DNA Methylomes in Health and Disease" Mid-Term Workshop, the 2nd of July 2009, and the last one was this year, the "DNA Methylation: from Biology to Disease" Final Workshop, in June the 29th. The quality of all publications and conferences offered by all the participants in the CancerDIP project demonstrated the added-value and positive impact of this project in the basic knowledge of cancer epigenetics. The publications and thus the conferences have been related to the main objectives of CancerDIP and all the participants have been contributed.

Regarding to the first objective about optimization and validation of the MeDIP technology, the CancerDIP project have published the methodology related to the establishment and the use of MethylCap technology, thus this new technique is now at reach of the whole scientific community and it is been extensively applied for DNA methylation profiling in many laboratories all over. The developed protocols for profiling of DNA methylation have enabled scientists to embark on comparisons of whole-genome DNA methylation profiles rather than individual genes. The MethylCap technology has already appeared in several independent publications, where other types of cancer and other biological systems are have been studied. Our data provides a catalogue of epigenetic changes linked to colon carcinoma as well as acute myeloid leukemia, two major cancer types. The catalogue will allow us to define distinct subtypes within these cancers based on novel markers. The use of such sub-type specific methylation markers is a prerequisite for development of personalized treatment strategies.

Regarding the objective of the analysis of mechanism of epigenetic deregulation in human cancer, a novel mechanism regulating DNA methylation in cancer has been described: the regulation of DNMTs enzymatic activity by a post-translational phosphorilation, whose results are compiled in a publication that is currently in revision. The relationship between PcG proteins and MBDs characterized at a genome-wide level was impossible to achieve due to the lack of proper antibodies. As an alternative, interesting results regarding breast cancer DNA methylation profiles have been obtained and are submitted for publication.

Respect to the aim about the epigenetic profiling in leukemias and colon cancer: establishment of clinical correlations, we expect that the results obtained within the timeframe of the CancerDip consortium will stand even after the end of the consortium. A part depict the scenario of DNA methylation deregulation in colon cancer and in AMLs, the data obtained might have a clinical value both at diagnostic and prognostic level. If some DMRs identified in AMLs (WP3) and in colon cancer (WP4) might be useful to corroborate the diagnosis, some of these DMRs and some of the identified gene targets might assume a prognostic significance. On one hand these data might help in the identification of the MRD (minimal residual disease) in AML. On the other, the results might be exploited to set up new protocols and kits to be used as clinical DNA methylation deregulation-tools in the follow up of patients with colon cancer or AML. Given that colon cancer is the third cause of death for cancer in the world and that more than 70% of AML relapse after standard treatment, our results might have a great socio-economic impact with enormous implications on human health.

In addition, some of the data, relative to the case of acute promyelocytic leukemia (APL, AML-M3) have been published in Cancer Cell. Interestingly, candidate DMRs between AML and CD34+ cells identified by DMH have been confirmed by MeDIP and MethylCap analyses, suggesting their potential impact as diagnostic markers and, possibly, as prognostic markers as indicated also in the sections below. Furthermore, it has been identified and characterised the impact of known and novel epigenetic modulators on DNA methylation both in leukemias and colon cancer (upon revision). Moreover, it has

been published the characterisation of DNA methylation of the PPARG (peroxisome proliferator-activated receptor gamma) promoter region in colon cancer to define its role as prognostic marker. And it is submitted that UHRF1 negatively regulates PPARG and is associated with a higher proliferative, clonogenic and migration potential. These findings strongly suggest that UHRF1 and PPARG play a crucial role in CRC (colon rectal cancer) development and might be regarded as prognostic biomarkers.

Regarding the objective generation of bioinformatic tools for analysis of MeDIP data, a software package has been developed and the computational discovery is published, BiQ Analyzer HT, the web-based analysis of (Epi-) genome data using EpiGRAPH and Galaxy. With the aim to optimize and to evaluate the predictive power and robustness of biomarker candidates, it has been published the epigenetic biomarker development article and the MethMarker: User-friendly design and optimization of gene-specific DNA methylation assays. Thus, the new software packages that have been produced in the frame on CancerDIP are now publics and can be used for the scientific community.

Respect to the objective development of a cancer MeDIP kit and validation, novel kits that facilitate research and contribute to improve cancer diagnosis and prognosis, have been produced, which will be useful for further studies of DNA methylation facilitating better clinical management. In fact, it has been generated three types of kits: the MeDIP assay, the MagMeDIP kit, which uses magnetic beads in the protocol, and the MethylCap kit. These kits are currently available and commercialized by Diagenode.

Most importantly, it has been provide the proof of principle that the DNA methylation fingerprints obtained might be useful for translational purposes by showing that it is possible to identify the tumor type origin of cancers of unknown primary origin (CUPs). Thus, it is tempting to propose that the prediction of a foster primary site for CUPs based on the DNA methylation profiles might identify a more specific treatment regimen for these patients that would improve their quality of life and survival, being an important social impact. With the aim to disseminate the wider social implications of the project, the following media have been used to communicate information about the CancerDIP project to the general public, such as press release, media briefing, articles published in the popular press and coverage in specialist press, as well as a website for the general public. The consortium has also made a video in order to disseminate the final results of the project to both experts and general public. During the opening of the "DNA Methylation: from Biology to Disease", the final project workshop, the coordinator showed the video to all participants and the media. The video has been also published in CancerDIP website as well as in other websites (IDIBELL ...) and social media (i.e. Facebook, Youtube, Twitter...).

List of Websites:
http://www.cancerdip.eu

Scientific coordinator:
Dr Manel Esteller
Email: mesteller@idibell.cat
Tel. 0034 93 260 7140 / 0034 93 260 7243

Project Manager:
Mrs Verónica Padial
Email: pebc@idibell.cat
Tel. 0034 93 260 7143