UNIVERSITA' DEGLI STUDI DI MILANO-BICOCCA
FACOLTA' DI MEDICINA E CHIRURGIA

DOTTORATO DI RICERCA IN EMATOLOGIA SPERIMENTALE
(XXIV ciclo)

Coordinatore: Prof. ENRICO MARIA POGLIANI

TESI DI DOTTORATO DI RICERCA

GENOME-WIDE DNA METHYLATION PROFILES BY HIGH-THROUGHPUT TECHNIQUES IN ALCOHOL-RELATED HEPATOCELLULAR CARCINOMA: IDENTIFICATION OF EPIGENETIC SIGNATURES IN LIVER TISSUE AND PERIPHERAL BLOOD CELLS DNA AS POTENTIALLY USEFUL BIOMARKERS OF DISEASE

Relatore:
Dott.ssa SIMONETTA FRISO
Università degli Studi di Verona

Dottorando in Ricerca:
Dott.ssa SILVIA UDALI

ANNO ACCADEMICO 2011-2012

TABLE OF CONTENTS

1. ABSTRACT

DNA methylation is the major epigenetic feature of eukaryotic cell DNA and consists in the covalent binding of a methyl group to the 5' carbon of a cytosine in a CpG dinucleotide sequence and acts regulating gene expression. Methyl-transfer reactions occur within one-carbon metabolism pathway that takes place principally in the liver: hepatic tissue that is, therefore, to be considered among the most interesting target tissues for DNA methylation analysis. Moreover alcohol, a major risk factor for hepatic cancer, is known to disturb one-carbon metabolism but the mechanisms underlying the alcohol-related liver carcinogenesis are still incompletely understood. We, therefore, designed this study to investigate DNA methylation profiles in alcohol-related hepatocellular carcinoma (HCC) by high-throughput techniques for genome-wide analysis. Main scope of the present project was to define a possible role for DNA methylation in the development of non-viral alcohol-related HCC in DNA obtained either from liver tissue and peripheral blood mononuclear cells (PBMCs) with the final goal of identifying potentially useful epigenetic biomarkers for HCC from an easily accessible DNA source, namely PBMCs.

The methylation status and the transcriptional levels of all the annotated genes were assessed on liver HCC tissue compared to homologous non-neoplastic tissue using a genome-wide, array-based approach in eight patients undergoing curative surgery. The merging of DNA methylation and gene expression data allowed identifying hypermethylated and transcriptional repressed genes among which six genes belonging to retinol metabolism (*ADH1A, ADH1B, ADH6, CYP3A43, CYP4A22* and *RDH16).* Among other hypermethylated-repressed genes, was detected a key gene of one-carbon metabolism, *SHMT1*, *ESR1*, a transcription factor with an hormone-binding domain involved in cell cycle regulation, and hepcidin, a liver peptide hormone involved in iron homeostasis were also identified as epigenetically regulated through DNA methylation

inducing transcriptional repression. Interestingly, the gene expression analysis on RNA extracted from buffy coat of HCC patients, alcoholic patients without liver cancer and healthy subjects revealed that transcriptional repression of *RDH16* was significantly associated with hepatic cancer. Moreover, the down-regulation of *RDH16, SHMT1* and *ESR1* was associated to chronic alcohol intake compared to controls. These findings suggest that expression profiles of specific genes obtained from PBMCs may be useful biomarkers for HCC.

2. ABBREVIATIONS

ADH 6	Alcohol dehydrogenase 6
BHMT	Betaine-Homocysteine methyltransferase
CBS	Cystathionine-β-synthase
Ct	Cycle threshold
DNMTs	DNA methyltransferases
ESM1	Endothelial specific molecule 1
ESR1	Estrogen receptor 1
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
HAMP	Hepcidin
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
Hcy	Homocysteine
IP	Immunoprecipitated sample
INPUT	Non-immunoprecipitated sample
MeDIP	Methylated-DNA immunoprecipitation
MS	Methionine-synthase
MTHFR	Methylenetetrahydrofolate-reductase
PBMCs	Peripheral blood mononuclear cells
RDH 16	Retinol dehydrogenase 16
ROI	Region of interest
ROS	Reactive oxygen species
SAH	S-Adenosyl-Homocysteine
SAM	S-Adenosyl-Methionine
SHMT1	Serine-hydroxymethyl-transferase 1
THF	Tetrahydrofolate

3. INTRODUCTION

3.1 DNA methylation, the best characterized epigenetic mechanism

Epigenetics is an emerging field of molecular research that investigates those heritable mechanisms able to modulate gene expression without modifying the base sequence of DNA (Bird, 2007). Epigenetic phenomena can be considered a bridge from genotype to phenotype since they are the way by which a specific cell or tissue interprets the genome information determining the phenotypical expression (Goldberg et al., 2007). Epigenetic phenomena are characterized to be potentially reversible and to be influenced by nutritional-environmental factors (Friso and Choi, 2002), so they appear promising in the field of prevention of pathologic conditions in addition to the understanding of molecular mechanisms underlying different pathologies.

Epigenetic mechanisms include post-translational histone modifications, RNA-based mechanisms and DNA methylation (Figure 1) (Udali et al.). DNA methylation is a fairly stable epigenetic modification that consists in the covalent binding of a methyl group to the 5' carbon of cytosine occurring at CpG dinucleotide sequences in the mammalian genome (Feinberg, 2007). This reaction is catalyzed by DNA methyltransferases (DNMTs) a class of enzymes that are distinguished in maintenance and *de novo* DNMTs for the differential specific functions. DNMT1 acts to maintain the methylation levels during mitotic processes by adding methyl groups to hemi-methylated DNA during DNA replication, while DNMT3A and 3B act after DNA replication and introduce new methylation sites (Boland and Christman, 2009; Jia and Cheng, 2009). CpG dinucleotide sequences, the putative methylation sites, are present in the human genome with a lower frequency than what expected and this could be due to the spontaneous deamination of methylcytosine to thymine (Illingworth and Bird, 2009).

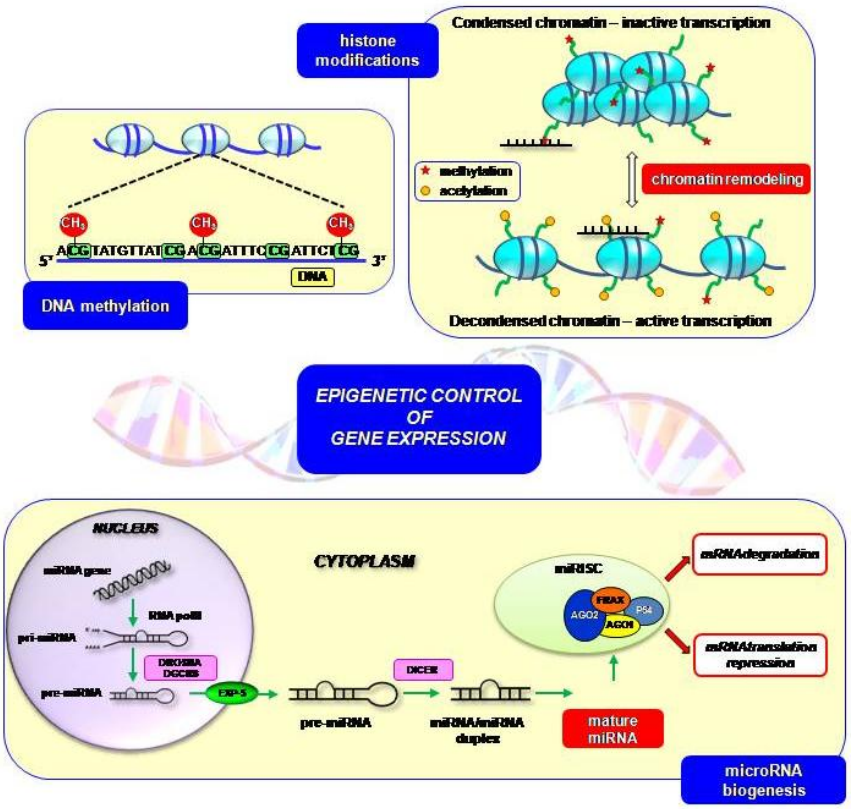

Figure 1. Principal epigenetic mechanisms of gene expression regulation: DNA methylation, histone modifications and microRNAs (authorized reproduction for scholarly purposes, Udali et al.).

CpG sequences are distributed genome-wide but they are often grouped (with a frequency ≥50%) in regions of variable length (200 bp – 2 Kb), the so called CpG islands. The regions within genome called CpG islands are mostly not-methylated and they localise in the promoter of constitutively expressed (*i.e.* housekeeping) genes. On the other hand, about 40% of tissue-specific genes contain CpG islands in the promoter region and in these cases methylation could have a role in transcriptional regulation (Illingworth and Bird, 2009).

The main function of DNA methylation is to modulate the expression of the genetic information by modifying the accessibility of DNA to the transcriptional machinery. DNA

methylation in the promoter region is classically associated to transcriptional inhibition while demethylation seems to be necessary to let the transcription start (Luczak and Jagodzinski, 2006). The transcriptional repression mediated by DNA methylation has been explained by two different mechanisms: 1. the methyl groups directly block the binding sites of specific transcription factors in the promoter region; 2. methylated CpG sites are recognized by specific binding proteins that form a spatial obstacle to the formation of transcriptional complexes (Luczak and Jagodzinski, 2006).

DNA methylation has crucial physiological roles in the cell acting in the stabilization of chromosomes, genomic imprinting, X-chromosome inactivation, mammalian embryogenesis, and it is important in the inhibition of repeat elements and transposons transcription. This epigenetic mechanism has also been studied in relation to several pathologic conditions, mostly cancer (Jones, 1986; Jones and Laird, 1999; Ehrlich, 2002; Feinberg and Tycko, 2004; Feinberg, 2007) but also other chronic diseases such as cardiovascular disease (Udali et al., ; Friso et al., 2012). A wide variety of cancers are characterized by aberrant DNA methylation and, in particular, a global DNA hypomethylation and a concurrent gene-specific hypermethylation of tumour-suppressor genes have been described (Ehrlich, 2006; Jones and Baylin, 2007).

3.2 DNA methylation and one-carbon metabolism

One-carbon pathway (or folate-dependent one-carbon metabolism) is a network of methyl-transfer reactions that are involved both in nucleic acids synthesis and in biological methylation (Figure 2) (Choi and Mason, 2002). Folate is the principal carrier of methyl

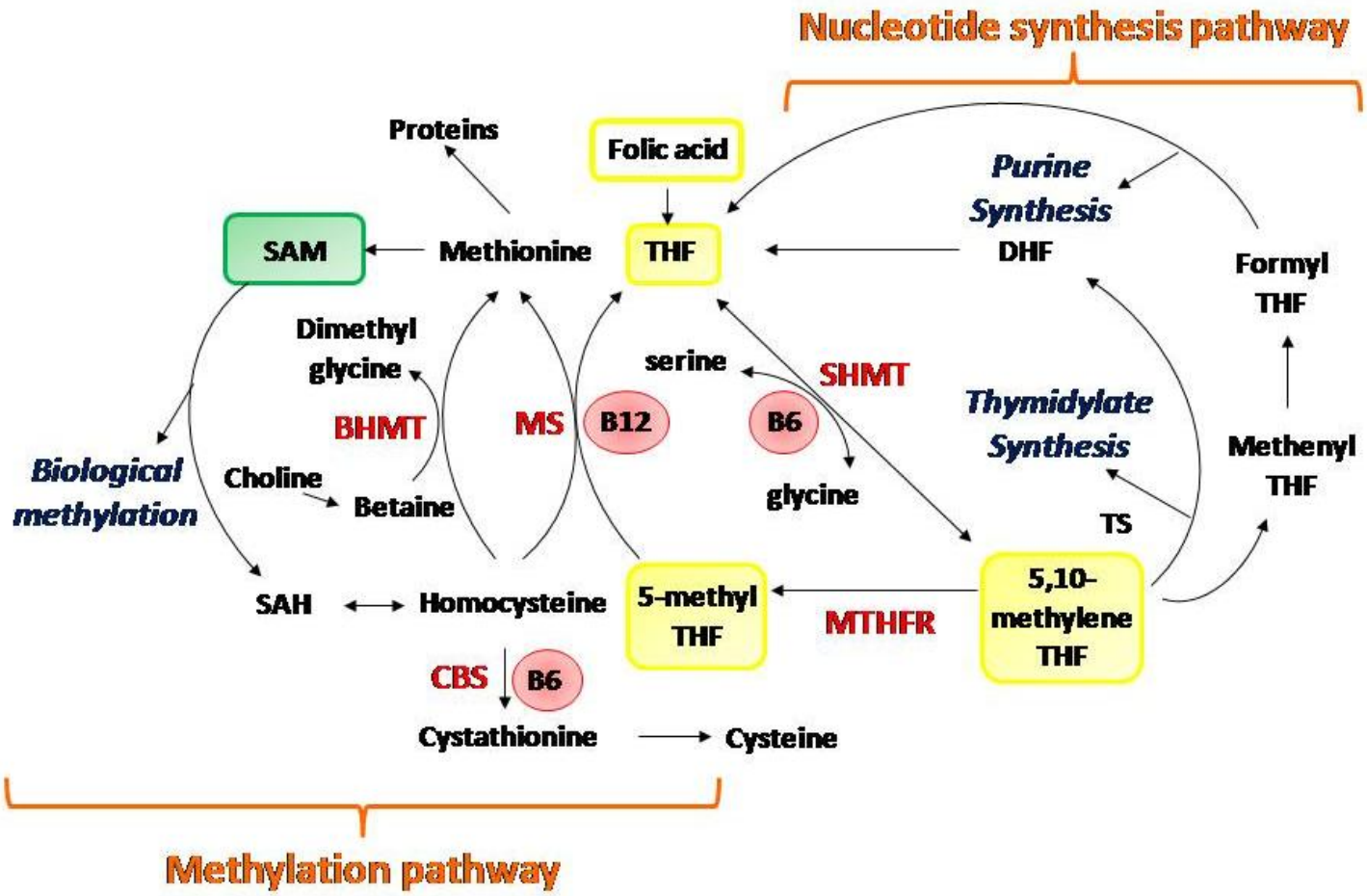

Figure 2. One-carbon metabolism. BHMT, betaine-homocysteine methyltransferase; CBS, cystathionine-β-synthase; MS, methionine synthase; MTHFR, methylenetetrahydrofolate-reductase; SAH, S-adenosyl-homocysteine; SAM, S-adenosyl-methionine; SHMT, Serine-hydroxymethyl-transferase; THF, tetrahydrofolate. Modified from Choi and Mason (Choi and Mason, 2002).

units and it enters the pathway in its reduced and active form, *i.e.* tetrahydrofolate (THF). The approach to DNA methylation implies a deeper investigation of one-carbon metabolism since, within this cluster of reactions, occurs the methyl-transfer from S-Adenosyl-Methionine (SAM) to DNA (Choi and Mason, 2002). One-carbon metabolism reactions (Figure 2) are catalyzed by several enzymes that play an essential role in the methyl-groups transfer reactions and in the balance between DNA synthesis and biological methylation. One key enzyme is SHMT (Serine-hydroxymethyl-transferase) that catalyzes the reversible conversion of serine and THF to glycine and 5,10 methylene THF by utilizing vitamin B6 as a coenzyme (Figure 3).

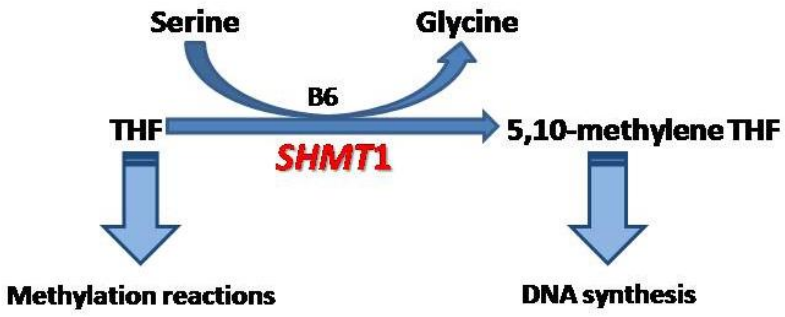

Figure 3. SHMT1: a metabolic switch between DNA synthesis and methylation reactions. SHMT1, Serine-hydroxymethyl-transferase 1; THF, tetrahydrofolate.

The 5,10 methylene THF molecule can enter the nucleotide synthesis pathway (thymidylate and purine synthesis) or can be irreversibly converted to 5 methyl THF by methylenetetrahydrofolate-reductase (MTHFR), another central enzyme of the pathway. Then methionine-synthase (MS), a vitamin B12-dependent enzyme, catalyzes the transfer of a methyl group from 5-methyl THF to homocysteine (Hcy) leading to the formation of methionine and THF, which can re-enter the cycle. Methionine is then converted into SAM, the universal methyl groups donor for biological methylation reactions, including those of DNA, RNA, proteins and lipoproteins. After the transfer of a methyl unit, SAM is

converted into S-adenosyl-Homocysteine (SAH) which is then hydrolyzed to adenosine and Hcy. Hcy can be re-methylated by MS, as already described, or can be converted to methionine through the choline and betaine pathway mediated by the enzyme betaine-homocysteine methyltransferase (BHMT). Hcy can alternatively be condensed with serine to form cystathione in an irreversible reaction catalyzed by cystathionine-β-synthase (CBS), using vitamin B6 as a coenzyme (Figure 2) (Choi and Mason, 2002).

3.3 One-carbon metabolism, liver and alcohol

One-carbon metabolism reactions take place in the liver that is the best target tissue to study DNA methylation and one-carbon metabolism. Chronic alcohol intake is known to interfere with one-carbon metabolism in different ways. Alcohol reduces folate availability both by reducing intestinal uptake and increasing renal excretion (Hamid and Kaur, 2006) and it has been described to reduce gene expression of folate carriers (*RFC*, reduced folate carrier, and *PCFT*, proton-coupled folate transporter) and to decrease RFC affinity for folate (Wani et al., 2012). Alcohol is also known to reduce the activity of methionine synthase (MS) and of methionine-adenosyl transferase (MAT) determining a decreasing in the availability of SAM, the universal methyl-donor for methylation reactions (Lu and Mato, 2005; Lu et al., 2006). Moreover chronic alcohol consumption is usually associated with a deficit of B6 and B12 vitamins (Cravo et al., 1996) which are coenzymes of several methyl-transfer reactions. Since alcohol interferes with one-carbon metabolism, it is interesting to analyze the effects on DNA methylation in liver tissue, the target tissue of this pathway.

3.4 Alcohol, hepatocellular carcinoma and DNA methylation

Chronic alcohol consumption is associated with higher risk of hepatocellular carcinoma (HCC) and cancer of the upper aerodigestive tract, colorectum and breast (Poschl and Seitz, 2004).

HCC, in particular, is the most frequent primary liver cancer accounting for 70% to 85% of the total liver cancer burden worldwide. The main risk factors for HCC are HBV and HCV infection in developing countries while high alcohol intake has a leading role in developed countries, especially the United States and several other Western countries (Jemal et al., 2011).

Most recently an updated report has been released by World Health Organization, see http://www.who.int/gho/publications/world_health_statistics/EN_WHS2012_Full.pdf.

Although ethanol is a well recognized etiological factor for HCC, the exact mechanism by which it promotes liver carcinogenesis is still not completely elucidated. A central role is attributed to acetaldehyde, the first metabolite produced during alcohol degradation, that acts as a carcinogen by interfering with DNA synthesis and repair mechanisms (Seitz and Stickel, 2007). The carcinogenic role of acetaldehyde is well recognized for the upper and lower gastrointestinal tract, but in the liver the role of this compound seems to be less important (Seitz and Stickel, 2007). In the hepatic tissue the action of alcohol is likely to be mediated by oxidative stress, due to ethanol-induced reactive oxygen species (ROS) formation. Moreover, alcohol intake is associated to aberrant methyl group transfer and this event may play a role in alcohol-mediated carcinogenesis (Seitz and Stickel, 2007).

DNA methylation is abnormally regulated in HCC (Herceg and Paliwal, 2011; Pogribny and Rusyn, 2012), as well as in a wide variety of cancers (Ehrlich, 2006), and it is of high interest to elucidate how this epigenetic mechanism is affected by alcohol, in order to shed light on alcohol-related hepatic carcinogenesis.

3.5 DNA methylation profiling by MeDIP-chip analysis

DNA methylation profiling is an emerging field in the epigenetic studies on cancer with the aim of both unravelling the carcinogenic processes and of finding candidate cancer biomarkers. Nowadays different techniques are available to perform the profiling: whole-genome bisulphite sequencing, restriction enzyme-enriched sequencing techniques and affinity-enrichment-based techniques combined either with sequencing or microarray hybridization (Heyn and Esteller, 2012). Among these, a technique that combines high resolution and affordable costs is methylated-DNA immunoprecipitation followed by microarray analysis (MeDIP-chip). This is an affinity-enrichment-based technique that enables to perform genome-wide DNA methylation analysis coupling the use of methylation-specific antibodies and microarray hybridization (Weber et al., 2005). Five-methyl cytidine antibody recognises methylated CG dinucleotides and allows highly efficient enrichment of methylated DNA fragments; immunoprecipitated samples (IP) are then analysed in comparison with total DNA (INPUT, not immunoprecipitated).The microarray analysis and the rough data elaboration enable then to calculate an absolute methylation value for all annotated genes (Down et al., 2008; Rakyan et al., 2008). These high-throughput techniques provide a high amount of data that need to be validated by a different method, usually by bisulfite sequencing-based techniques. Bisulfite sequencing is based on sodium bisulfite treatment that converts non-methylated cytosine into uracil while methylated-cytosine remains unaltered. This reaction produces, in correspondence to non-methylated cytosines, a cytosine to thymine conversion in the genome sequence that is then detected either by Sanger sequencing (Friso et al., 2012) or by pyrosequencing.

3.6 DNA methylation in peripheral blood mononuclear cells as a biomarker of cancer disease

Epigenetic features are strongly tissue-dependent so it is of particular interest to analyse the target tissue in order to stabilize a possible involvement of epigenetics in the carcinogenic process. On the other hand, it could be very interesting to identify epigenetic biomarkers in an easily accessible tissue in humans such as blood cells.

Most recently, DNA methylation has been tested in blood as a circulating tumor cell DNA marker (Zhang et al., 2007) and a number of studies evaluated the possible role of circulating white blood cells DNA methylation in different types of cancer as a potential marker to define the risk for malignancies of different tissue origin (Pufulete et al., 2003; Lim et al., 2008; Hou et al., 2010; Terry et al., 2011).

A stimulating scientific debate is ongoing to clarify the usefulness of genomic methylation status in DNA obtained from PBMCs as a suitable biomarker even for cancer tissue of different origin. Recent results that we obtained analysing global DNA methylation in PBMCs support the hypothesis that an hypomethylation is correlated to an increased risk for cancer development and that genomic PBMCs-DNA methylation may be a useful epigenetic biomarker for early detection and cancer risk estimation (Friso et al., 2013).

3.7 Aim of the study

Main scope of the present project was to define a possible role for DNA methylation in non-viral alcohol-related HCC in DNA obtained either from HCC tissue compared to non-neoplastic liver tissue and in peripheral blood mononuclear cells (PBMCs) DNA extracted from the same patient with the final goal of identifying potentially useful epigenetic biomarkers for HCC from an easily accessible DNA source in humans, specifically PBMCs.

4. MATERIAL AND METHODS

4.1 Subjects

The study was approved by the Institutional Review Board Ethical Committee of the University of Verona School of Medicine Hospital (Verona, Italy). Written informed consent was obtained from each patient after a detailed explanation of the study.

The subjects enrolled were distinguished in three different groups: HCC patients, alcoholic patients without hepatic neoplasia and healthy controls.

HCC patients. Thirty-three HCC patients were selected among those referring to the Division of Surgery, Section A of the Verona University Hospital. Key eligibility criteria included age ≥18 years and surgical resectability criteria were preserved liver function, class A Child-Pugh score, presence of a resectable single tumor or oligofocal resectable nodules (maximum three nodules), absence of extrahepatic metastases. For preoperative staging chest-abdomen computerized tomography (CT)-scan or nuclear magnetic resonance imaging (NMRI) were used. Positron Emission Tomography (PET-CT) or diagnostic laparoscopy was applied in selected cases. Resectability assessment included also tumor local stage, major vascular invasion and presence of affected lymphonodes. Exclusion criteria included a coexisting human immunodeficiency (HIV), hepatitis B (HBV) or C virus (HCV) infections; presence of relevant concurrent medical conditions such as chronic inflammatory diseases and haematological disorders, including autoimmune liver diseases and hereditary hemochromatosis; presence of acute inflammatory disease and decompensate liver cirrhosis (Child-Pugh B, C). Patients under B vitamins supplementation and/or using drugs known to interfere with folate-related one-carbon metabolism in the three months before study enrolment were excluded.

Alcoholic patients without hepatic neoplasia. Ten alcoholic patients without hepatic neoplasia were enrolled at the Division of Internal Medicine, Section B, of the Verona University Hospital. A condition of chronic alcohol consumption was evaluated by means

of AUDIT (Alcohol Use Disorder Identification Test) and CAGE questionnaires and defined as ≥ 36 g ethanol/day intake for males and ≥ 24 g ethanol/day for females. Patients under B vitamins supplementation and/or using drugs known to interfere with folate-related one-carbon metabolism in the three months before study enrolment were excluded.

Healthy subjects. Ten subjects were enrolled as healthy controls and key eligibility criteria included age ≥18 years, absence of neoplasia of any type, no history of viral infections (HIV, HBV, HCV), absence of other relevant medical conditions, no dietary supplementation of B vitamins and/or consumption of drugs interfering with folate-related one-carbon metabolism in the three months before study enrolment and alcohol intake ≤ 36 g ethanol/day for males and ≤ 24 g ethanol/day for females.

4.2 Blood analysis and biopsy specimens

Chemical clinical analysis. From each subject samples of venous blood were drawn after overnight fasting for routine laboratory analysis that included: complete blood count, aspartate transaminase (AST), alanine transaminase (ALT), gamma-glutamyl transpeptidase (GGT), cholinesterase (CHE), immunoglobulin (Ig) fractions including IgA fraction dosage, alpha-fetoprotein serum concentration, serological tests for hepatitis B and C viruses and for Epstein-Barr and Cytomegalovirus, antibodies anti-smooth muscle, anti-nuclear, anti mitochondrial, anti-liver-kidney microsomal type 1.

Buffy coat isolation. For buffy coat isolation samples of venous blood from each subject were drawn into Vacutainer® containing EDTA as anticoagulant after an overnight fast. Buffy coat was obtained from each blood sample by centrifuging at 2,500 *g* for 15 min at 4°C and collecting the white phase that stratified between plasma (upper phase) and red cells (lower phase).

Hepatic tissue collection. Liver samples were collected from the 33 HCC patients immediately after surgical resection; HCC tissue and non-neoplastic tissue, from a region

far from the tumor mass and histologically tumor-free, were obtained for each patient. The tissues were excised from the patients during the surgical procedure and examined by the surgeons by means of macroscopic intraoperative evaluation, that was subsequently confirmed by microscopic histological analysis. The pathologist who performed the histological diagnosis was unaware of the patient participation to the study.

4.3 DNA extraction from buffy coat and liver tissues

DNA extraction was performed by standard phenol/chloroform procedure in order to obtain high-quality genomic DNA. Concentration and purity were assessed by NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and only samples showing a suitable purity (260/280≥1.7, 260/230≥1.7) were used for epigenetic analysis. Extracted DNA was stored at -20°C.

Blood and tissue samples were treated with *ad hoc* protocols before the phenol/chloroform DNA extraction procedures.

Buffy coat. 300 µl of buffy coat were treated twice with 1 ml of cold sterile water and centrifuged at 3,300 *g* for 15 min at 4°C in order to remove red blood cells and preserve PBMCs. White cell membranes lysis was performed by adding 1.5 ml of Igepal CA-630 (Sigma-Aldrich, St. Louis, MO, USA) 0.1% to each aliquot, vortexing and centrifuging at 3,300 *g* for 15 min at 4°C. The supernatant was discarded and the nuclear pellet resuspended in 500 µl lysis solution (NaCl 100 mM, EDTA 25 mM, pH 8) to dissolve nuclear membranes. RNase treatment was performed by addition of 2.5 µl RNase 4 mg/ml and incubating at 37°C for 15 min. Then 10 µl Proteinase K (Promega, Fitchburg, WI, USA) 33.3 mg/ml and 30 µl 10% SDS were added and the samples were incubated overnight at 37°C before performing the "Standard Phenol-Chloroform DNA extraction".

Liver tissues. After surgical excision, tissue samples for DNA extraction were immediately sliced into aliquots of about 100 mg and snap-frozen in liquid nitrogen to be

then immediately stored at -80°C until use. 100 mg of tissue stored at -80°C were thawed and homogenized by Tissue Master 50 homogenizer (Omni International, Kennesaw, GA USA) in 2 ml of NaCl 0.9% w/v (50 mg of tissue in 1 ml of solution) and subdivided into 4 aliquots of 500 µl. 1.25 ml of Igepal CA-630 (Sigma-Aldrich, St. Louis, MO, USA) 0.1% was added to each aliquot, which was then vortexed and centrifuged at 13,400 *g* for 15 min at 4°C. The supernatant was discarded and the pellet resuspended in 250 µl lysis solution (NaCl 100 mM, EDTA 25 mM, pH 8). Two aliquots for each sample were merged (final volume 500 µl) and treated with RNase (2.5 µl RNase 4 mg/ml at 37°C for 15 min). Then 15 µl Proteinase K (Promega, Fitchburg, WI, USA) 33.3 mg/ml and 50 µl SDS 10% were added and the samples were incubated overnight at 37°C before performing the "Standard Phenol-Chloroform DNA extraction".

Standard Phenol-Chloroform DNA extraction

Phenol-chloroform extraction was performed by adding 500 µl of phenol/water/chloroform solution (Applied Biosystems, Carlsbad, CA, USA) to the samples, followed by centrifugation at 13,400 *g* for 15 min at room temperature and collection of the upper phase that was extracted once more with 500 µl of phenol/water/chloroform solution. Phenol traces were removed by chloroform/isoamyl alcohol extraction with 1 ml of chloroform/isoamyl alcohol 24:1 and centrifugation at 13,400 *g* for 15 min at room temperature. The upper phase was collected and DNA precipitated by adding 80 µl NaCl 4 M and 2 ml ice-cold absolute ethanol; then DNA pellet was dried and redissolved in 100 µl TE (TrisHCl 10 mM, EDTA 1 mM, pH8). Genomic DNA was completely resuspended by incubation for 1 h at 65°C.

4.4 RNA extraction from buffy coats and liver tissues

RNA was extracted by guanidinium thiocyanate-phenol-chloroform-based method following protocols specifically modified for buffy coat and liver tissue. RNA concentration was determined by NanoDrop and the purity and integrity of nucleic acid were assessed by 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). RNA samples were used in array-based gene expression analysis only when the RNA Integrity Number was ≥7. Extracted RNA was stored at -80°C.

Buffy coat. 150 µl of buffy coat were treated with 565 µl of TRI Reagent® BD (Sigma-Aldrich, St. Louis, MO, USA) and 15 µl of acetic acid 5N; the samples were then vortexed, incubated 5 min at room temperature and stored at -80°C until use.

For RNA extraction, 150 µl of chloroform were added to the thawed samples, vortexed, incubated 5 min at room temperature and then centrifuged at 12,000 *g* for 15 at +4°C. The centrifugation allowed the separation of different phases: the upper phase containing RNA, the middle phase DNA and the lower phase proteins. The upper phase was collected and RNA was precipitated by addition of 375 µl isopropanol and incubation for 10 min at room temperature; the samples were then centrifuged at 12,000 *g* for 8 min at 4°C. The pellet was washed with 750 µl 75% ethanol and centrifuged at 7,500 *g* for 5 min at 4°C. The pellet was dried and resuspended in 25 µl nuclease-free water followed by incubation at 60°C for 15 min.

Liver tissues. Immediately after surgical excision, 100 mg of tissue were homogenized by Tissue Master 50 homogenizer (Omni International, Kennesaw, GA USA) in 2 ml TriReagent® (Sigma-Aldrich, St. Louis, MO, USA) and stored two aliquots of 2 ml at -80°C until use. To extract RNA, 200 µl of chloroform were added to the thawed samples, vortexed, incubated 5 min at room temperature and centrifuged at 12,000 *g* for 15 at 4°C. The upper phase, containing RNA in aqueous solution, was collected and RNA was precipitated by addition of 500 µl isopropanol and incubation step of 10 min at room

temperature; the samples were then centrifuged at 12,000 *g* for 10 min at 4°C. The pellet was washed with 1.5 ml 75% ethanol and centrifuged at 7,500 *g* for 5 min at 4°C. The pellet was dried and resuspended in 50 µl nuclease-free water by incubation at 60°C for 15 min.

4.5 Methylated-DNA immunoprecipitation (MeDIP) analysis

MeDIP-chip analysis was performed on eight male patients selected among the thirty-three HCC patients on the basis of the availability of: a. all the laboratory data, b. adequate liver biopsy specimen with confirmed unequivocal HCC diagnosis and homologous tumor-free liver tissue, and c. a clear history of alcohol use habit. The analysis permitted to obtain the methylation profile of neoplastic and non-neoplastic tissue.

4.5.1 Genomic-DNA fragmentation: optimization of shearing protocol

DNA fragmentation is a crucial step in immunoprecipitation procedure and a uniform population of molecules is essential for achieving the highest performance.

As suggested by MeDIP protocol the first technique we applied was sonication, using the ultrasonic disintegrator Soniprep 150 (MSE, London, UK) on DNA samples extracted from buffy coat (100 ng/µl, in a total volume of 100 µl). The shearing conditions were optimized adopting cycles of 15 sec "ON" and 15 sec "OFF" at low power, for a total time of 5 min. However, successive experiments demonstrated a high variability among different samples, probably due both to the quality of starting DNA (degree of degradation) and to sonication procedure itself, that is difficult to standardize. Therefore we tested a protocol of nebulisation, using the GS Nebulizers Kit (Roche Applied Science, Penzberg, Germany), where tissue genomic DNA (15 µg) suspended in specific nebulisation buffer (containing glycerol, Tris-HCl 1 M and EDTA 0.5 M) was sheared by nebulisation using argon with a pressure of 3.5 bar for 1 minute. This method generates a uniform population

of molecules (ranging in size from 300 to 1,000 bp) and proved to be very reproducible. The sample was then cleaned up from nebulisation buffer components using the DNA Clean & Concentrator-25 Kit (Zymo Research, Irvine, CA, USA). Fragment size was assessed either by 2% agarose gel electrophoresis or by Agilent Bioanalyzer 2100 analysis.

4.5.2 MeDIP assay

Methylated DNA immunoprecipitation was performed using MeDIP kit™ mc-green-003 by Diagenode (Liège, Belgium) (Weber et al., 2005; Magdalena and Goval, 2009) and following the manufacturer protocol with some minor modifications.

12 µl of fragmented DNA (~ 1 µg) and 78 µl of incubation mix (containing buffer and positive methylated and negative unmethylated DNA controls) was incubated at 95°C for 7 min (instead of 3 min as indicated in the original protocol) to denature DNA and favour antibody binding. The sample was chilled on ice, spinned at 4°C and then 15 µl were drawn to constitute the INPUT (control sample not incubated with the antibody).

The remaining 75 µl were immunoprecipitated (IP) by adding 5 µl of antibody mix, containing buffers and anti-5methyl cytidine antibody (1:10 dilution) and 20 µl of meDNA-IP blocked beads (50% suspension); the immunoprecipitation was carried on at 4°C overnight on a rotating wheel. IP samples were washed 6 fold by adding 450 µl of ice-cold wash buffer, mixing by rotation for 5 min at 4°C, centrifuging at 500 *g* for 2 min at 4 °C (instead of 6,000 rpm for 1 min): at all steps the samples were kept at 4°C or on ice.

DNA (both IP and INPUT samples) was eluted from bead pellets by adding 120 µl of elution buffer and incubating at 65°C for 10 min, by vortexing every 30 sec (original protocol described incubation in a thermo-shaker for 10 min at 65°C at 1,000 to 1,300 rpm). The samples were purified by GenElute™ PCR Clean-Up Kit, Sigma (instead of QIAquick PCR purification columns, Qiagen), according to the manufacturer instructions,

and resuspended in 50 µl nuclease-free water (instead of TE buffer); the samples were then incubated at 50°C for 5 min and centrifuged at 13,400 *g* for 1 min.

4.5.3 Evaluation of immunoprecipitation efficiency

Immunoprecipitation enrichment was checked by RealTime qPCR (7500 Real-Time PCR System, Applied Biosystem, Carlsbad, CA, USA) with SYBR Green as fluorophore, both on internal and external controls; internal controls were represented by human genomic regions either methylated (X-linked α satellites, AlphaX1) or unmethylated (*GAPDH*), while external controls were DNA specimens totally methylated or totally unmethylated, that were added to the sample before immunoprecipitation.

Primers pairs (10 µM each):	internal ctrls	hum meDNA primer pair (AlphaX1)
		hum unDNA primer pair (*GAPDH*)
	external ctrls	meDNA pos control primer pair #1 and #2
		unDNAneg control primer pair #1 and #2

qPCR mix (total volume of 25 µl/reaction):

- 1 µl primer pair
- 12.5 µl SYBR PCR Green master mix
- 5 µl diluted DNA → sample dilution: 10 µl DNA and 35 µl water for all primers
  1:1,000 dilution for AlphaX1
- 6.5 µl water

qPCR temperature profile: 7' 95°C /40 cycles [95°C 15"- 60°C 60"] / 95°C 60"

The efficiency of methylated DNA immunoprecipitation of particular genomic loci was calculated from qPCR data and reported as a percentage of starting material:

$$\% \text{ (MeDNA-IP/Total input)} = 2\text{^}[(Ct^{(20\%input)} - 2.322) - Ct^{(MeDNA\text{-}IP)}] \times 100\%$$

where 2 = amplification efficiency (ΔE) and 2.322 = compensatory factor that takes into account the dilution 1:5 of the INPUT (compared to IP).

4.6 DNA-microarray analysis

Microarray analysis was performed on Human DNA Methylation 3x720K CpG Island Plus RefSeq Promoter Arrays (NimbleGen-Roche, Madison, WI, USA) in order to analyze the promoter methylation state of all annotated genes. These microarrays have high detection sensitivity since they have long (50-75 mer), isothermal oligonucleotide probes and high resolution (100 bp spacing) (Array design: Genome build HG18, promoter upstream tiling -2.44 Kb, downstream tiling +0.61 Kb, CpG Island 27,728). IP and INPUT samples were previously amplified, to obtain the DNA quantity necessary for the hybridization on arrays, the labelling and hybridization steps were then performed on NimbleGen Systems using established protocols.

4.6.1 Whole Genome Amplification

GenomePlex® Complete Genome Amplification (WGA) kit (Sigma-Aldrich, St. Louis, MO, USA) was applied following the producer's protocol with minor modification: the fragmentation step was skipped since DNA was already fragmented.

The library was prepared with 10 µl of sample (both IP and INPUT), 2 µl of Library Preparation Buffer and 1 µl of Library Stabilization Solution; after vortexing and centrifugation the samples were incubated at 95°C for 2 min. The samples were cooled on ice, 1 µl of Library Preparation Enzyme was then added and the mixture was incubated in thermal cycler as follows: 16°C for 20 min, 24 °C for 20 min, 37°C for 20 min, 75°C for 20 min, 4°C hold.

The amplification was then performed adding 7.5 µl of 10x Amplification Master Mix, 47.5 µl of Nuclease-free water and 5 µl of WGA DNA Polymerase; the thermal programme was as follows: 95°C for 3 min, 20 cycles at 94°C for 15 sec and 65°C for 5 min. The amplification products were then purified by GenElute™ PCR Clean-Up Kit (Sigma-Aldrich, St. Louis, MO, USA) and eluted in 50 µl nuclease-free water.

4.6.2 Sample labelling

Sample labelling was performed by *NimbleGen Dual-Color DNA Labeling kit*. Pairs of samples (1.5 µg IP and 1.5 µg INPUT, each in a volume of 40 µl) were labelled in parallel with 40 µl of Cy5-Random Nonamers (IP) or Cy3-Random Nonamers (INPUT) by adding 2 µl of Klenow Fragment (3'>5' exo) 50 U/µl, 10 µl of 10 mM dNTP mix and 8 µl of Nuclease-free water (final volume 100 µl). The reagents were kept on ice and the samples were carefully assembled avoiding vortexing; the reaction was performed incubating 3 hours at 37°C. The reaction was stopped by adding 10 µl of Stop Solution (0.5 M EDTA). Labelled-DNA was then precipitated by addition of 11.5 µl 5M NaCl and 110 µl isopropanol; the samples were thoroughly vortexed and incubated 10 min at room temperature protected from light; finally they were washed in ice-cold 80% ethanol, resuspended in 25 µl nuclease-free water and quantified by NanoDrop.

4.6.3 Hybridization

16.5 µg of each labelled samples (IP and INPUT) were dried in SpeedVac with low heat, protected from light, and then resuspended in 5.6 µl of different Sample Tracking Control (STC) specific for each sample, that permitted to distinguish the samples on the array. After the addition of 14.4 µl of hybridization solution (2x Hybridization Buffer, Hybridization Component A and Alignment Oligo) the samples were thoroughly vortexed, incubated at 95°C for 5 min and then placed at 42°C while preparing the hybridization chip. The chip was prepared with the aid of a mixer (HX3 mixer for 3x720K array) following the detailed instruction provided in the manual and 18 µl of each sample were loaded on the chip. The hybridization was carried on at 42°C for 17 hours on the Hybridization System 4 (NimbleGen-Roche, Madison, WI, USA).

4.6.4 Washes and Two-colours array scanning

The arrays were washed three times with vigorous, constant agitation: Wash I for 2 min, Wash II for 1 min and Wash III for exactly 15 sec (*NimbleGen Wash Buffer Syste*m). The microarray was dried by centrifugation (1 min) and immediately scanned at 2.5 μm resolution by Axon GenePix 4400A scanner (Axon Instruments Inc, Union City, CA, USA) acquiring Cy3 and Cy5 signals respectively at 532 nm and 635 nm wavelength. Fluorescence intensity raw data were obtained from scanned images of the arrays by using *Nimblescan 2.5 extraction software* (NimbleGen-Roche, Madison, WI, USA). We obtained two "pair reports" for each array (one for the Cy3 image and one for the Cy5 image), representing the raw data format for NimbleGen DNA Methylation experiments, that contain the signal intensities for each probe on the array.

4.6.5 Methylation data analysis: Batman algorithm

MeDIP-chip raw data were analyzed by Batman, Bayesian tool for methylation analysis (Down et al., 2008), a cross-platform algorithm freely available (http://td-blade.gurdon.cam.ac.uk/software/batman) under the GNU Lesser General Public License, that permits to calculate absolute methylation values. In the Batman analysis the tissue samples, distinguished in neoplastic and non-neoplastic tissues, were considered as biological replicates. The promoter region of each gene was subdivided in 500 bp-long regions of interest (ROIs) and an absolute methylation value was associated to each ROI. A ROI was considered differentially methylated when the difference between absolute methylation values of neoplastic and non-neoplastic tissue was ≥30% (Feber et al., 2011).

4.7 Validation of MeDIP-chip data by direct bisulfite sequencing

DNA methylation data were validated on three hypermethylated (*ESR1, RDH16, SHMT1*) and one hypomethylated (*ESM1*) genes by direct bisulfite sequencing.

4.7.1 Bisulfite treatment

Bisulfite treatment was performed using the EpiTect® Bisulfite Kit (QIAGEN, Germantown, MD, USA) performing the modification on membrane-bound DNA.

Bisulfite reaction (total vol 140 µl):

2 µg DNA (add RNase-free water to reach a volume of 20 µl)
85 µl Bisulfite Mix (high bisulfite salt concentration, low pH)
35 µl DNA Protect Buffer (containing a pH-indicator dye that turns from green to blue if the pH is correct)

Temp. profile: 5’ at 95°, 25’ at 60°, 5’ at 95°, 85’ at 60 °, 5’ at 95°, 175’ at 60°, hold 20°C.

The bisulfite-treated DNA was cleaned up to remove bisulfite salts and other chemicals used in the conversion process, that inhibit the sequencing procedures. 560 µl of freshly prepared Buffer BL containing 10 µg/ml carrier RNA (recommended if the DNA is fragmented) was added to the sample: this step promotes the binding of the converted single-stranded DNA to the column membrane. The entire mixture was transferred to the EpiTect spin column and then centrifuged at maximum speed for 1 min. One wash was performed by adding 500 µl Buffer BW and centrifuging at maximum speed for 1 min. Desulfonation, the final step in chemical conversion of unmethylated cytosine into uracil, was achieved by adding 500 µl of Buffer BD onto the column membrane and incubating 15 min at room temperature. The column was then centrifuged at maximum speed for 1 min and washed twice (500 µl Buffer BW and centrifugation at maximum speed for 1 min). Residual liquid was removed by placing the column in new 2 ml tubes, centrifuging at maximum speed for 1 min and then placing the column with open lid in heating block

(56°C for 5 min). Purified DNA was eluted from the column with 20 µl of Elution Buffer and centrifuging at 15,000 *g* for 1 min; to increase DNA yield the last step was repeated twice, recovering a final volume of 40 µl. The bisulfite-treated DNA was stored at -20°C.

4.7.2 Direct sequencing

The differentially methylated ROI was amplified with *ad hoc* primers, designed avoiding CpGs, adopting optimized PCR conditions.

PCR conditions. In Table 1 are reported, for each gene, the primer sequences (the underlined primers were used for the sequencing), the annealing temperature (Ta), the length (bp) of the fragment amplified and the number of CG present in the amplified region.

Table 1: primers pairs for bisulfite sequencing

Gene	forward	reverse	Ta	bp	n°CG
ESR1	GTATTGGGTATTGGGATAGAGAG	TCTTACTCAAACATAAACTCA	55°	462	4
RDH16	TAGAAAGGTTTTATTGGGTAG	CCTAATATACCATTTACTAAAACC	55°	575	5
SHMT1	GTAGGGTGGTTATTTAAAGTAGGA	CTCCTAAACTCAAACCATCTACC	55°	512	10
ESM1	TTGTTGTTATAGTGTTGAGGGTAG	AAACTCTAAAACAAAACTACACCT	58°	547	12

The general PCR conditions were 0.4 µM primers, 0.2 mM dNTPs, 1.5 mM $MgCl_2$ and 0.75 U Super AB Taq (AB Analitica, Padova, Italy); the volume of the PCR mixture was 25 µl and the bisulfite treated DNA 4 µl. The thermal cycler used was the GeneAmp® PCR System 9700 (Applied Biosystems, Carlsbad, CA, USA).

The temperature profile was: 5‘ 95°/ 40 cycles [95° 60“- Ta 60”- 72° 60”] /72° 7’

Agarose gel electrophoresis. To verify the presence of a single specific band the amplification products were run on 3% w/v agarose gel: 50% agarose (Promega, Fitchburg, Wi, USA) and 50% High Resolution agarose (Sigma-Aldrich, St. Louis, MO, USA)

dissolved in TBE (Tris-borate-EDTA buffer pH 8.0) containing Ethidium bromide (EtBr). The amount of amplified DNA was determined comparing the intensity of the band with bands at known concentration of molecular weight marker MWM VIII (Boehringer-Manheim, Germany).

PCR-products clean-up. The PCR-products were purified by GenElute^{TM} PCR Clean-UP kit (Sigma-Aldrich, St. Louis, MO, USA) to remove primers and dNTPs, known to interfere with the sequencing procedure.

Sequencing. The direct Sanger sequencing was performed on the Capillary Electrophoretic Nucleic Acid Sequencer CEQ 8800 (Beckman Coulter, Brea, CA, USA); the sequencing was performed by the Section of Pediatrics, Department of Life and Reproduction Sciences (University of Verona).

4.7.3 Methylation index evaluation

Bisulfite sequencing methylation data were obtained by calculating a methylation index for the CpG sites present in the ROI of each gene, as previously reported (Friso et al., 2012). The analysis was performed according to the following procedure: a. measurement of the height of the T peak from a CpG site (T_{CpG}), this T derives from C and represents how much this CpG site is unmethylated; b. evaluation of the height of two T peaks one before and one after each candidate CpG site, checking that these two peaks were originally T bases. The average height of these two Ts is assumed as control (T_{mean}); c. the methylation percentage at each CpG site is calculated according with the formula:

$$\% \text{ methylation at the CpG site} = 100 - [(T_{CpG} / T_{mean}) \times 100]$$

4.8 Gene expression analysis by microarrays

Gene expression analysis was performed by means of Human Gene Expression 12x135K Array (Nimblegen-Roche, Madison, WI, USA) that analyzes 45,033 target genes with 60mer probes (3 probes / target), following the producer's protocol.

The analysis was performed on RNA extracted from the tissues of the same eight patients, in which methylation was analyzed by MeDIP-chip.

4.8.1 Double-stranded cDNA synthesis

Double-stranded cDNA was synthesized by Superscript® Double-Stranded cDNA Synthesis Kit (Invitrogen, Carlsbad, CA, USA).

First strand cDNA synthesis. 10μg of total RNA, 1μl of oligo dT Primer and DEPC Water (11μl total volume) were mixed, heated to 70°C for 10 minutes, then briefly spinned and placed on ice for 5 minutes. After the addition of 4μl of 5X First Strand Buffer, 2μl of 0.1M DTT and 1μl of 10mM dNTP Mix, the samples were incubated at 42°C for 2 minutes, and then, after the addition of 2μl of SuperScript II, the samples were incubated at 42°C for 60 minutes.

Second strand cDNA synthesis mix:	Reaction mixture of the previous step	20μl
	DEPC Water	91μl
	5X Second Strand Buffer	30μl
	10mM dNTP Mix	3μl
	10U/μl DNA Ligase	1μl
	10U/μl DNA Polymerase I	4μl
	2U/μl RNase H	1μl

The samples were incubated at 16°C for 2 hours.

2μl of 5U/μl T4 DNA polymerase were added to samples and incubated at 16°C for additional 5 minutes; the reaction was stopped placing the samples on ice and adding 10μl

of 0.5M EDTA. RNase treatment (1µl of 4mg/ml RNase A solution, incubation at 37°C for 10 min) was performed and cDNA was then purified by phenol:chloroform:isoamyl alcohol extraction. cDNA was precipitated with 16µl of 7.5M ammonium acetate, 7µl of 5mg/ml glycogen, 326µl of ice-cold absolute ethanol and centrifuged at 12,000 *g* for 20 min. The pellet was washed twice with 500µl of ice-cold 80% ethanol (v/v) and then rehydrated with 20µl of water.

4.8.2 cDNA labelling: One-Color DNA Labelling Kit

1µg of cDNA was labelled with 40 µl of Cy3-Random Nonamers by adding 2 µl of 50 U/µl Klenow Fragment (3'>5' exo), 10 µl of 10 mM dNTP mix and 8 µl of Nuclease-free water (final volume 100 µl) and incubating 2 hours at 37°C. The reaction was stopped by adding 10 µl of 0.5 M EDTA.

Labelled-cDNA was then precipitated by adding 11.5 µl of 5M NaCl and 110 µl isopropanol; the samples were thoroughly vortexed and incubated 10 min at room temperature protected from light; cDNA was then washed in ice-cold 80% ethanol, resuspended in 25 µl nuclease-free water and quantified by NanoDrop.

4.8.3 Hybridization

4 µg of Cy3-labeled cDNA were dried in SpeedVac with low heat, protected from light, and then resuspended in 3.3 µl of Sample Tracking Control (STC). After the addition of 8.7 µl of hybridization solution (2x Hybridization Buffer, Hybridization Component A and Alignment Oligo) the samples were vortexed, incubated at 95°C for 5 min and then placed at 42°C while preparing the chip for the hybridization. The chip was prepared with the mixer (HX12 mixer for 12x135K array) following the detailed instruction described in the manual and 6 µl of each sample were then loaded on the chip. The hybridization was

carried on at 42°C for 17 hours on the Hybridization System 4 (NimbleGen-Roche, Madison, WI, USA).

4.8.4 Washes and One-color array scanning

The arrays were washed three times with vigorous, constant agitation: Wash I for 2 min, Wash II for 1 min and Wash III for exactly 15 sec (*NimbleGen Wash Buffer Syste*m). The microarray was dried by centrifugation (1 min) and immediately scanned at 2.5 μm resolution by Axon GenePix 4400A scanner (Axon Instruments Inc, Union City, CA, USA). The slide was scanned at 532 nm wavelength and scanned images (TIFF format) were then imported into NimbleScan 2.5 software for grid alignment and expression data analyses.

4.8.5 Gene expression data calculation

Expression data were normalized through quantile normalization and the Robust Multichip Average (RMA) algorithm (Irizarry et al., 2003) included in the NimbleScan software. Statistical analysis on gene expression array-based results was performed with Limma R package (Smyth, 2005) considering a log2 fold change ≥1 or ≤ -1 and a p value adjusted for multiple testing (FDR) ≤0.05 as threshold to define differentially expressed genes (Benjamini and Hochberg, 1995).

4.9 Validation of array-based gene expression data

Gene expression results were validated on seven repressed (*ADH6, BCO2, ESR1, GDF2, HAMP, RDH16* and *SHMT1*) and four induced genes (*DNMT3B, ESM1, NOX4* and *SPINK1*) by RealTime qPCR. Single-strand cDNA was synthesized with TaqMan® Reverse Transcription Reagents (Applied Biosystems, Carlsbad, CA, USA); the reaction conditions are reported below.

Reverse transcriptase reaction (total volume 10 µl):

	µl	final concentration
10x TaqMan RT buffer	1	1x
25 mM $MgCl_2$	2.2	5.5 mM
dNTPs mix	2	500 µM each
random hexamers	0.5	2.5 µM
Rnase inhibitor	0.2	0.4 U/ µl
Reverse Transcriptase (50U/ µl)	0.25	1.25 U/ µl
H2O (nuclease free)	variable	
RNA	variable (0.2 µg)	

Temperature profile: 25°C for 10', 48°C for 30', 95°C for 5'

Real time qPCR was performed in 20 µl reaction volume with TaqMan chemistry on 7500 Real-Time PCR System (Applied Biosystem, Carlsbad, California, USA). The 18S rRNA (Hs99999901_s1) was used as endogenous control (Boujedidi et al., 2012).

Mixture reaction:	TaqMan Universal PCR Master Mix 2x	10 µl
	TaqMan assay (target gene) 40x	0.5 µl
	TaqMan assay (18S) 40x	0.5 µl
	cDNA	5 µl
	H2O (nuclease free)	4 µl

The TaqMan assays utilized are listed below:

ADH6 - Hs00167423_m1
BCO2 - Hs00230564_m1
ESR1 - Hs00174860_m1
GDF2 - Hs00211913_m1
HAMP -Hs00221783_m1
RDH16 - Hs00559712_m1
SHMT1 - Hs00541038_m1
DNMT3B - Hs00171876_m1
ESM1 - Hs00199831_m1
NOX4 - Hs00418356_m1
SPINK1 - Hs00162154_m1

Gene expression data were analyzed by evaluating the difference in mRNA levels from neoplastic and non-neoplastic tissue of each patient. The calculation formula was: $\Delta\Delta Ct=(Ct_{target}-Ct_{18s})_{HCC} - (Ct_{target}-Ct_{18s})_{nonHCC}$.

4.10 Gene expression on RNA extracted from buffy coat

The transcriptional levels of target genes (*ADH6*, *DNMT3B*, *ESM1*, *ESR1*, *GDF2*, *HAMP*, *NOX4*, *RDH16*, *SHMT1*, *SPINK1*) were analyzed on RNA extracted from buffy coat following the protocol described in the previous paragraph. Gene expression data were analyzed calculating $\Delta Ct = Ct_{target} - Ct_{18s}$, instead of $\Delta\Delta Ct$, since there were no reference samples. In the samples where the target gene was undetectable, an arbitrary Ct value of 45 was assigned in order to calculate a ΔCt. The value of 45 was selected because the program settings provides 40 cycles of amplification and we selected a Ct value far from the end point of the reaction. The comparisons were made among HCC patients, alcoholic patients without hepatic neoplasia and healthy subjects.

4.11 Data mining

The large lists of genes obtained with array-based methylation and gene expression analysis were analyzed with PANTHER (Protein Analysis THrough Evolutionary Relationship) classification system, that allowed to cluster genes of interest on the basis of their biological process involvement (Thomas et al., 2003), and with DAVID, the Database for Annotation, Visualization and Integrated Discovery, that was helpful to discover enriched functional-related gene groups and to visualize genes on KEGG pathway maps (Huang da et al., 2009; Huang da et al., 2009). TiGER (Tissue-specific Gene Expression and Regulation) database (Liu et al., 2008) was utilized to check the genes of interest tissue-specific expression.

4.12 Statistical analysis

All the calculations were performed using the SPSS V.17.0 statistical software (SPSS Inc, Chicago, IL, USA). The statistical data analysis was performed applying a T-test for independent samples and the differences were considered statistically significant when $p \geq 0.05$.

5. RESULTS

5.1 Methodological optimization and assessment

5.1.1 Optimization of DNA shearing protocol. DNA shearing was firstly obtained employing a sonication procedure and the best shearing conditions were determined using a DNA sample extracted from buffy coat. The agarose gel reported in Figure 4 contains the same DNA sample but sonicated in 6 different ways.

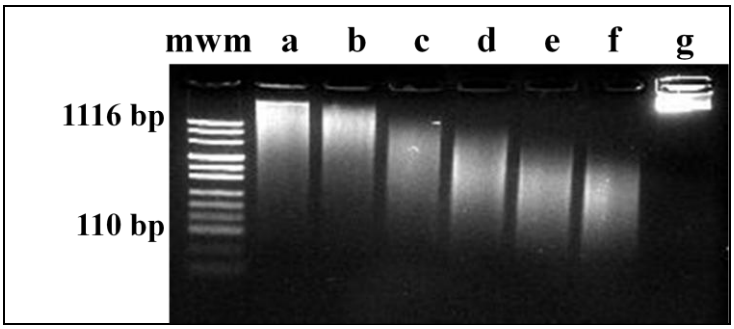

Figure 4. Sonication time-course. 0.2 µg of DNA treated with different protocols of sonication were loaded on a 2% agarose gel : 1 sec "ON" and 1 sec "OFF" repeated 20 fold (**a**); 15 sec "ON" and 15 sec "OFF" for 1 min (**b**), for 3 min (**c**), for 5 min (**d**), for 10 min (**e**) or for 15 min (**f**). Lane **g** contains native DNA; mwm, molecular weight marker.

The optimized sonication method was 15 sec "ON" and 15 sec "OFF" for 5 min (Figure 4 lane **d**) that produced fragments in the size range 150-900 bp. However this technique demonstrated a high variability among different samples, as shown in the Figure 5.

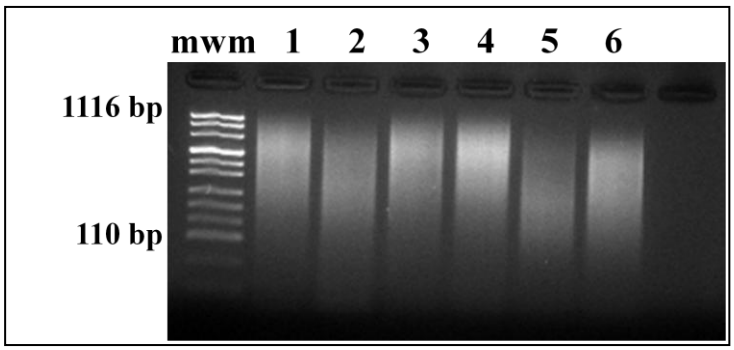

Figure 5. Inter-samples sonication variability. 0.2 µg of different DNA samples extracted from tissues (numbered **1**-**6**) were loaded on 2% agarose gel; the samples were subjected to the same protocol of sonication: 15 sec "ON" and 15 sec "OFF" for 5 min; mwm, molecular weight marker.

To improve the reproducibility of the shearing technique, nebulisation was implemented applying an argon pressure of 3.5 bar for 1 minute. Noteworthy, the variability among the samples was considerably reduced, as demonstrated in Figure 6. In panel A it is shown the electrophoretic pattern of different DNA samples (extracted from liver tissue) fragmented by nebulisation. The fragments length was about 300 to 1,000 bp and it was confirmed also by Bioanalyzer (Figure 6 Panel B).

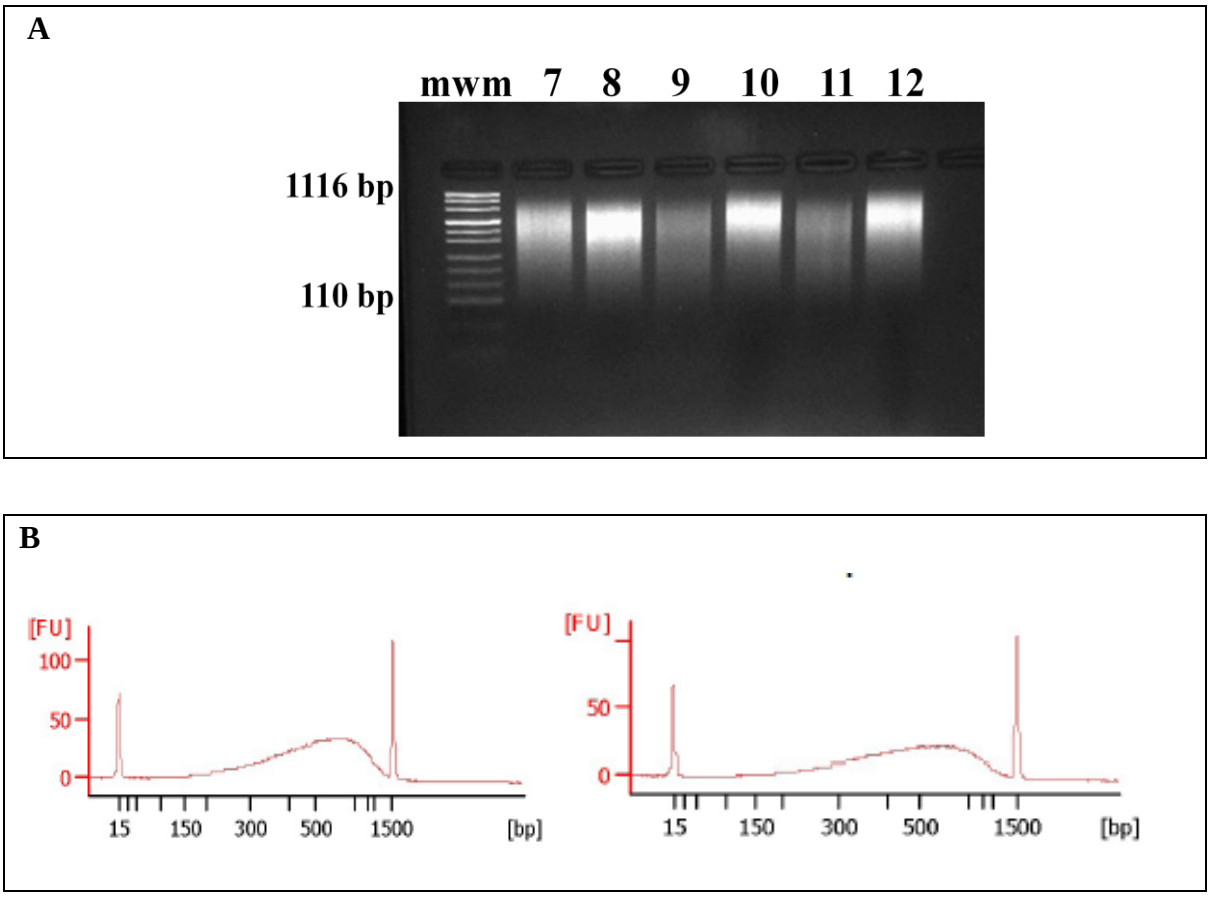

Figure 6. DNA shearing by nebulization. Panel A: 0.2 µg of different nebulised DNA (numbered 7-12) were loaded on 2% agarose gel; mwm, molecular weight marker. Panel B: Bioanalyzer results of two illustrative nebulised samples.

5.1.2 Evaluation of immunoprecipitation efficiency

Immunoprecipitation efficiency was assessed by RealTime qPCR on the controls provided in the MeDIP kit™ (Diagenode); internal controls were represented by human genomic regions either methylated (AlphaX1) or unmethylated (*GAPDH*), while external controls

were DNA specimens totally methylated or totally unmethylated, that were added to the sample before immunoprecipitation. The results are reported in Figure 7.

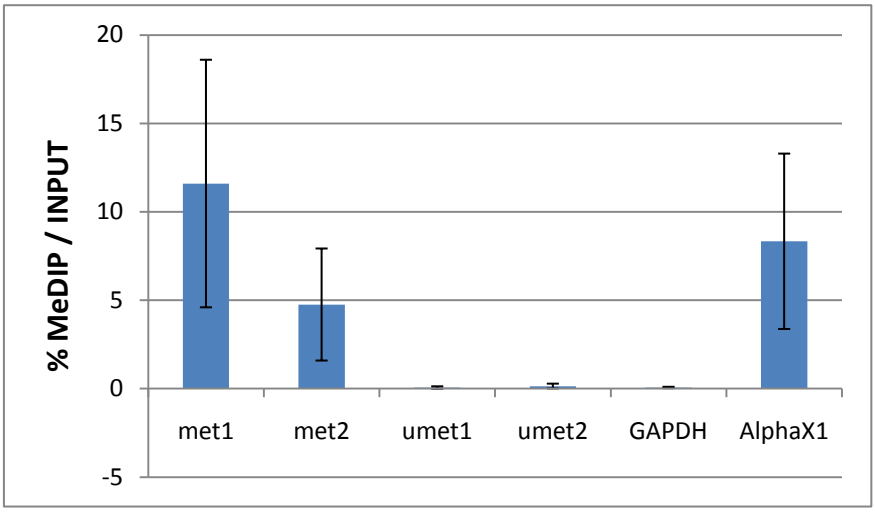

Figure 7. MeDIP efficiency. MeDIP efficiency was calculated according to the formula

$$\% \text{ (MeDNA-IP/Total input)} = 2^{\wedge}[(Ct^{(20\%input)} - 2.322) - Ct^{(MeDNA\text{-}IP)}] \times 100\%$$

Met1 and met2 are positive external controls, umet1 and umet2 are negative external controls, GAPDH is the negative internal control and AlphaX1 is the positive internal control. The graph represents mean values and standard deviations calculated on the 16 (8 neoplastic and 8 non-neoplastic tissue) analyzed samples.

The results were in good accordance with the data reported in the MeDIP kit manual by Diagenode and with the report by Magdalena J. and Goval J.J. (Magdalena and Goval, 2009) despite a rather high variability among different samples, as demonstrated by the standard deviation bars.

5.1.3 Validation of MeDIP-chip data by direct bisulfite sequencing

DNA methylation data were validated analyzing the promoter methylation status of four genes (*ESR1*, *RDH16*, *SHMT1* and *ESM1*) by direct bisulfite sequencing. The differentially methylated ROI of each gene was amplified with *ad hoc* primers in order to obtain a single specific band. The PCR specificity was assessed by agarose gel electrophoresis (Figure 8).

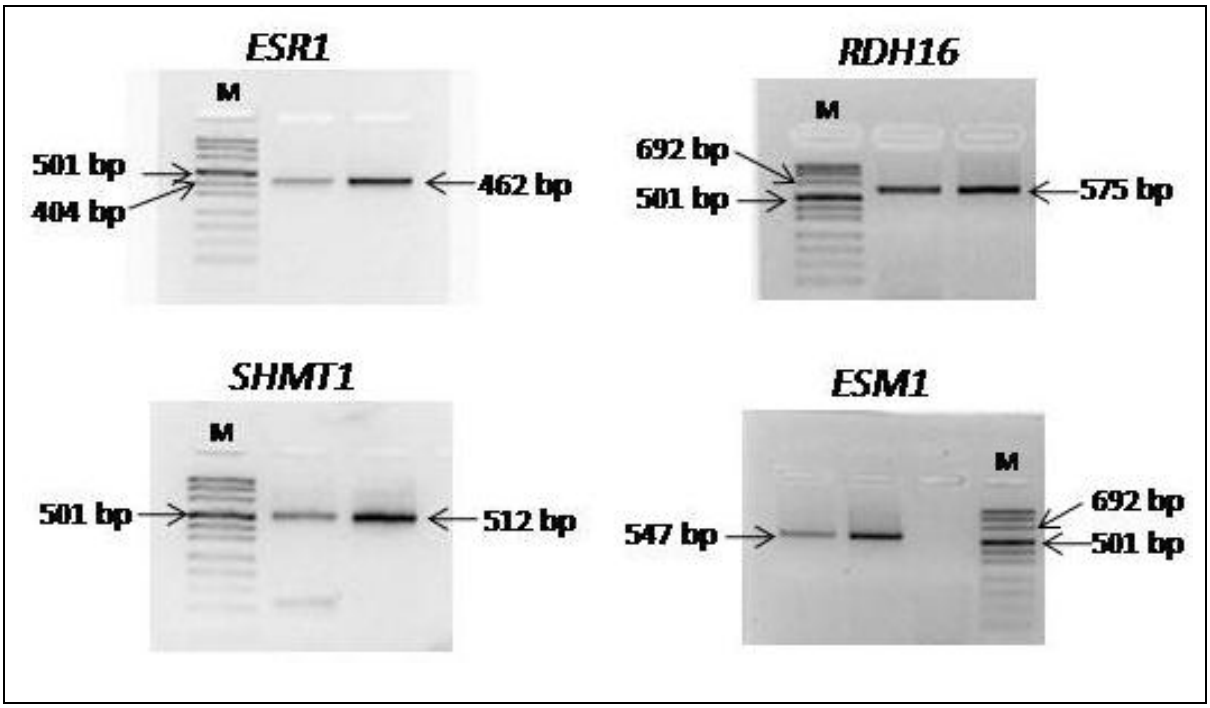

Figure 8. Agarose gel electrophoresis of amplification products used for sequencing. PCR products obtained with specific primers were run on 3% agarose gel along with a molecular weight marker (M). In this figure amplicons of specific length are reported : 462 bp for *ESR1*, 575 bp for *RDH16*, 512 bp for *SHMT1* and 547 bp for *ESM1*. The image colors are inverted, so the bands appear black on a light background.

The amplified DNA was semi-quantified considering that in the molecular weight marker the 501 bp band contains 120 ng, the 692 bp band 45 ng and the 404 bp band 50 ng of nucleic acid.

The bisulfite sequencing of amplified fragments resulted in electropherogram (Figure 9) where unmethylated cytosines were all converted into thymines, except in the CpG sites where cytosines-thymine proportion depended on the methylation level of each specific CG dinucleotide.

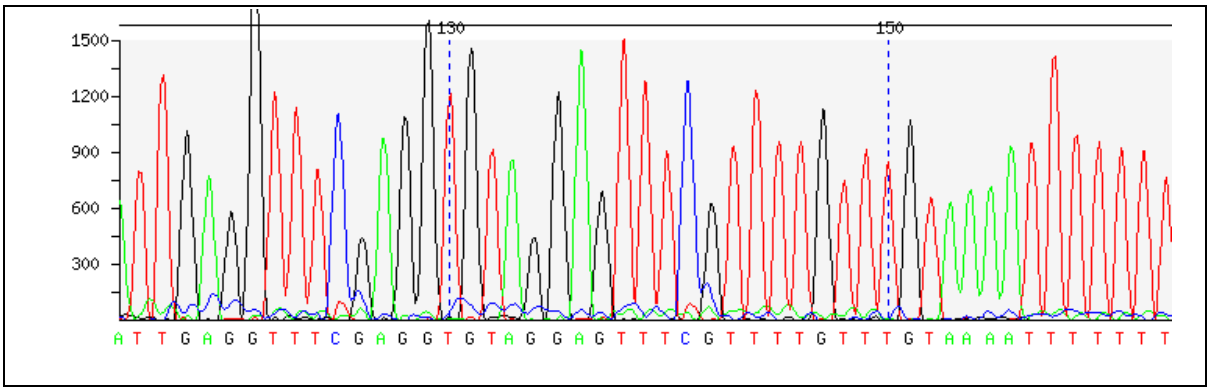

Figure 9. Example of bisulfite sequencing electropherogram. The sequence presents abundance of T residues that derive from the conversion of C residues outside CG dinucleotide.

The methylation percentage at each CpG site was calculated according with the formula:

$$\% \text{ methylation at the CpG site} = 100 - [(T_{CpG} / T_{mean}) \times 100]$$

where T_{CpG} was the height of the T peak at a CpG site and T_{mean} was the average height of two T peaks as illustrated in Figure 10.

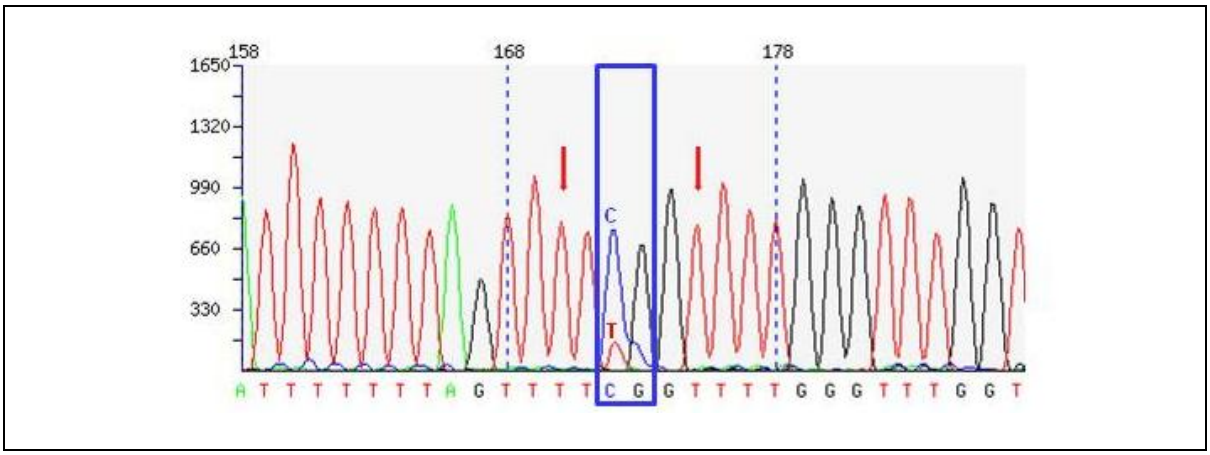

Figure 10. % methylation at a CpG site. Example of Sanger sequence electropherogram where at a CG dinucleotide site C peak (methylated C) and T peak (unmethylated-converted C) are represented by two overlapping peaks. The T peaks indicated by arrows were utilized to calculate T_{mean} while the CG dinucleotide is shown in the blue box.

The validation was performed on a hypomethylated (*ESM1*) and three hypermethylated (*SHMT1*, *RDH16* and *ESR1*) genes. The bisulfite sequencing results are reported in Table 2 and illustrated in Figure 11.

Table 2. Bisulfite sequencing results

GENE	MeDIP-chip results	% mC non-neopl (n=4)	% mC HCC (n=4)	ROI position (bp from TSS)	ROI length (bp)	n°CpGs
ESM1	hypometh	54.5+1.5	41.1+7.3	-139	401	12
SHMT1	hypermeth	76.2+3.9	82.2+2.0	-524	331	10
RDH16	hypermeth	63.6+3.3	67.0+7.8	-623	544	5
ESR1	hypermeth	70.7+7.8	75.0+4.5	-607	484	4

MeDIP-chip results, DNA methylation data obtained by microarray analysis; % mC non-neopl, mean methylation percentage of non-neoplastic tissue; % mC HCC, mean methylation percentage of HCC tissue; ROI position, distance (bp) of the differentially methylated ROI from the TSS; n° CpGs, number of CpGs evaluated by sequencing.

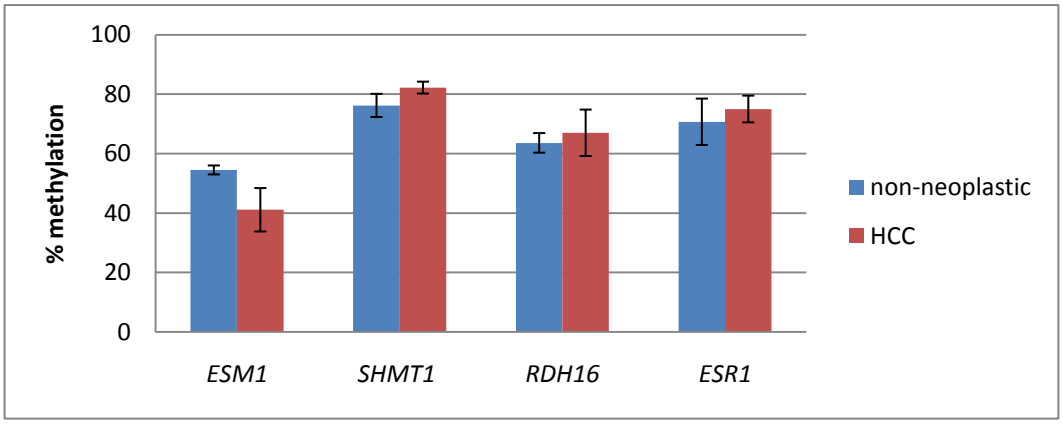

Figure 11. Bisulfite sequencing results. In the graph are reported mean and standard deviation of methylation values in the four analyzed patients.

The results obtained by bisulfite sequencing confirmed that *ESM1* was hypomethylated while *SHMT1*, *RDH16* and *ESR1* were hypermethylated in HCC tissue as compared with homologous non neoplastic tissue.

5.1.4 Validation of array-based gene expression data

The expression level of eleven genes (*ADH6*, *BCO2*, *DNMT3B*, *ESM1*, *ESR1*, *GDF2*, *HAMP*, *NOX4*, *RDH16*, *SHMT1* and *SPINK1*) was assessed by RealTime qPCR comparing HCC and non-neoplastic liver tissue. Results are reported in Figure 12.

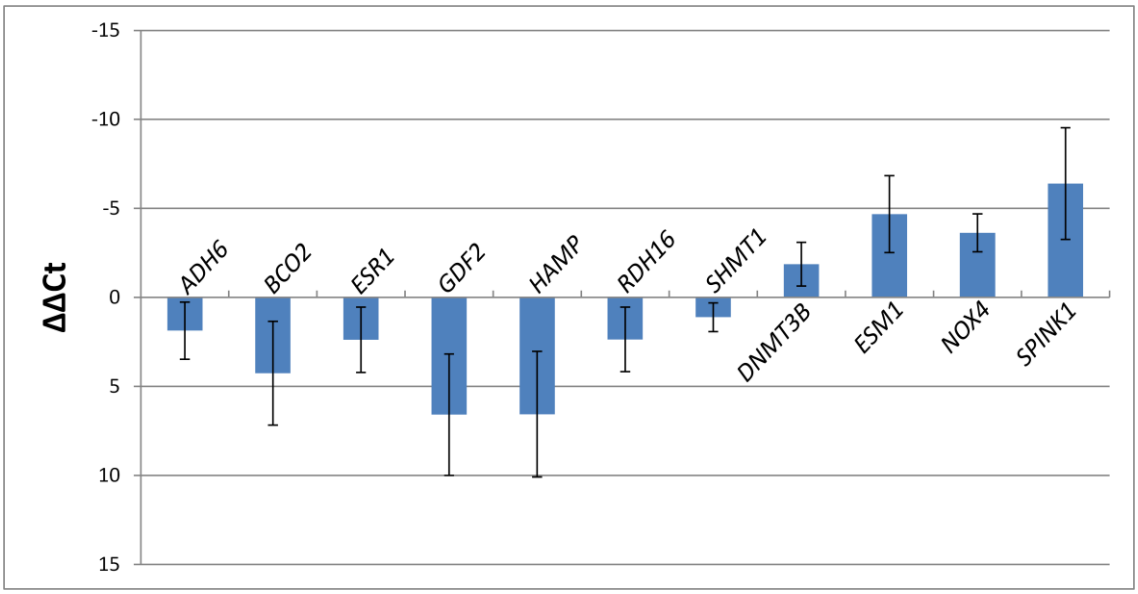

Figure 12. Gene expression results by RealTime qPCR. Gene expression results are expressed according to the formula $\Delta\Delta Ct=(Ct_{target}-Ct_{18s})_{HCC} - (Ct_{target}-Ct_{18s})_{nonHCC.}$ In the graph are reported ΔΔCt mean and standard deviation values in the eight analyzed patients.

Data obtained by RealTime qPCR confirmed the results of the array-based technique, despite a certain degree of variability, demonstrated by the quite high standard deviation. *ADH6*, *BCO2*, *ESR1*, *GDF2*, *HAMP*, *RDH16* and *SHMT1* resulted to be repressed while *DNMT3B*, *ESM1*, *NOX4* and *SPINK1* were induced in the HCC tissue.

5.2 Data analysis

5.2.1 Clinical characteristics of HCC affected patients

The main clinical and biochemical characteristics of the eight HCC patients selected for MeDIP-chip and array-based gene expression analysis are described in Table 3.

Table 3. Clinical and biochemical characteristics of the patients selected for MeDIP-chip analysis

Age (years)	Alcohol intake (Units*)	Smoking	Child Pugh Score	HBsAg	HCV Ab	Hb (g/dL)	MCV (fL)	IgA (g/L)	GGT (U/L)	CHE (U/L)	AST (U/L)	ALT (U/L)	aFP (μg/L)
66	>20	yes	A6	neg	neg	12.1	88.02	2.22	67	4138	25	57	5871
70	6	yes	A5	neg	neg	12.0	69.14	0.91	32	4606	23	37	411
66	16	yes	A5	neg	neg	15.6	89.85	4.08	48	4624	30	33	967
82	11	yes	A5	neg	neg	11.2	88.77	1.66	43	7353	29	28	62
68	5	no	A5	neg	neg	16.0	98.06	3.98	52	7856	25	29	190
60	6	yes	A5	neg	neg	13.8	98.60	1.60	223	3564	38	51	5
75	4	yes	A5	neg	neg	13.7	90.26	3.28	30	7225	28	15	431
71	10	yes	A5	neg	neg	13.3	91.40	1.56	88	8430	24	29	21

*Units, 12 g of ethanol per die (12 g of ethanol are contained in 125 ml wine or 330 ml beer or 40 ml spirit)

Patients were males with an age ranging from 60 to 82 years. All of them were habitual drinkers for a period ≥ 20 years and, according to the guidelines of Italian INRAN (Istituto Nazionale di Ricerca per gli Alimenti e la Nutrizione), classified as heavy drinkers considering a threshold of daily alcohol intake ≥ 3 Units i.e. 36 g ethanol. The stage A Child-Pugh score confirmed the absence of decompensated liver disease. Viral serologic tests for HBV and HCV were confirmed to be negative for all patients. Transaminases values were normal as were GGT and CHE. IgA were also within the normality range in each patient. Hematologic laboratory tests were normal and in particular MCV and haemoglobin levels were within the normality range. Alpha-fetoprotein was considerably higher than normal in all patients but in one (Table 3).

5.2.2 Promoter methylation profiles differentiate HCC versus non-neoplastic tissue.

The MeDIP-chip analysis of HCC *versus* non-neoplastic tissue showed that 2401 gene promoters were hypermethylated and 1244 were hypomethylated in hepatocarcinoma tissue. In Figure 13 the differentially methylated genes are graphically represented by HeatMap (Panel A). The PANTHER classification system identified a large number of differentially methylated genes belonging to several pathways involved in carcinogenesis such as those related to apoptosis, cell communication and adhesion, cell cycle regulation and immune system processes as shown in Figure 13, Panel B.

5.2.3 Gene expression in HCC versus non-neoplastic tissue

The array-based analysis of gene expression of HCC versus non-neoplastic tissue showed 1005 down-regulated and 670 up-regulated genes. Among the repressed genes, notable was the presence of several genes belonging to the retinol metabolism (*ADH1A, ADH1B, ADH6, CYP1A1, CYP1A2, CYP2B6, CYP2C9, CYP26A1, CYP3A4, CYP3A43, CYP4A11, CYP4A22, RDH16, RDH5* and *LRAT*) and a group of genes pertaining one-carbon metabolism (*BHMT1, BHMT2, CBS, GNMT, MTHFD2L* and *SHMT1*).

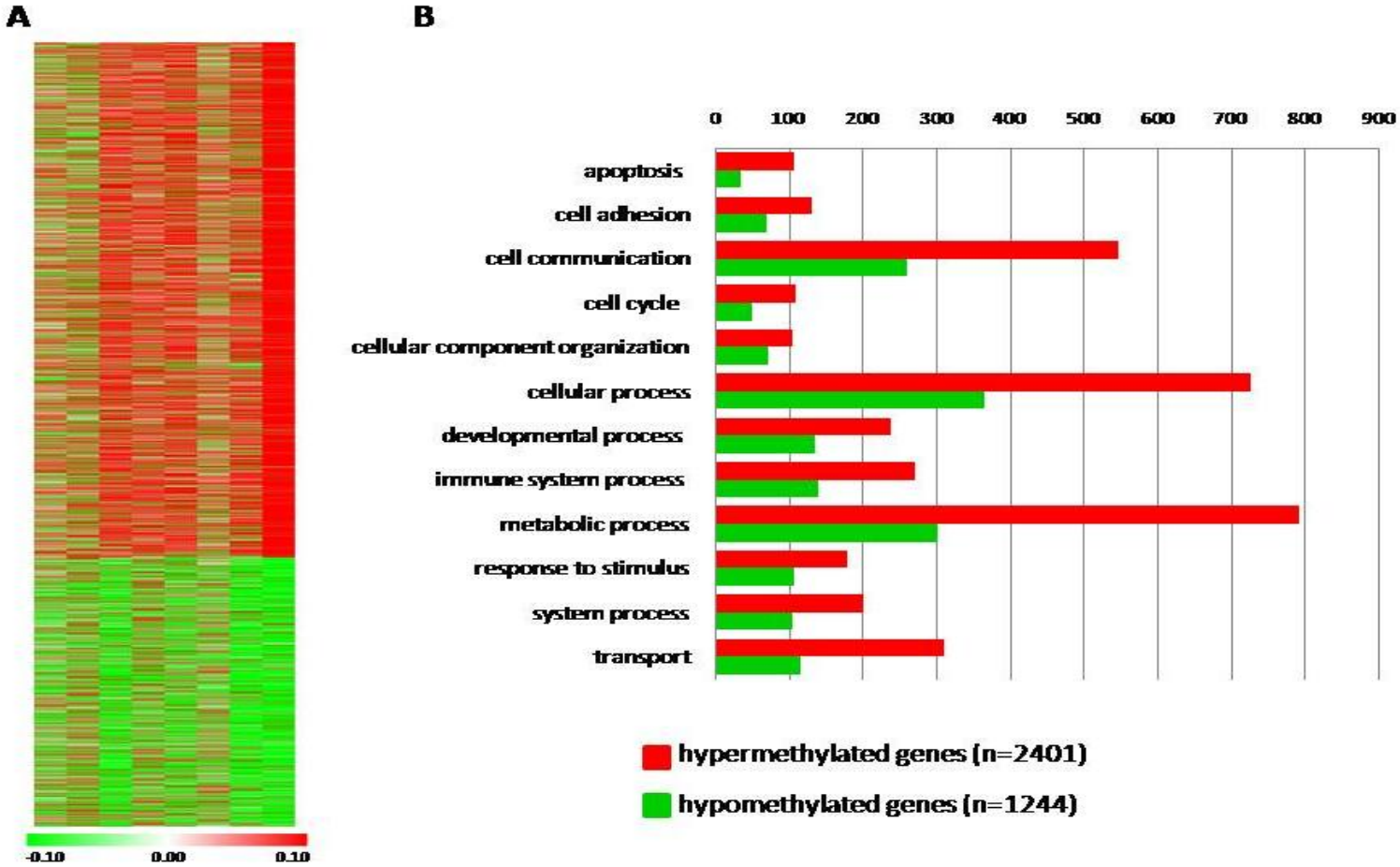

Figure 13. Differentially methylated genes. Panel A: HeatMap of total hypermethylated and hypomethylated genes. Hypermethylated (red) and hypomethylated (green) genes are classified by biological process according to PANTHER classification system.

5.2.4 Promoter DNA methylation profile according to array-based gene expression in HCC versus non-neoplastic tissue.

Promoter DNA methylation data were merged with array-based gene expression results in tumor *versus* tumor-free tissue. The analysis allowed distinguishing four groups of genes, according to both the promoter DNA methylation and gene expression profiles. The analysis highlighted 160 hypermethylated-repressed genes, 31 hypomethylated-induced genes, 50 hypermethylated-induced genes and 56 hypomethylated-repressed genes (Figure 14).

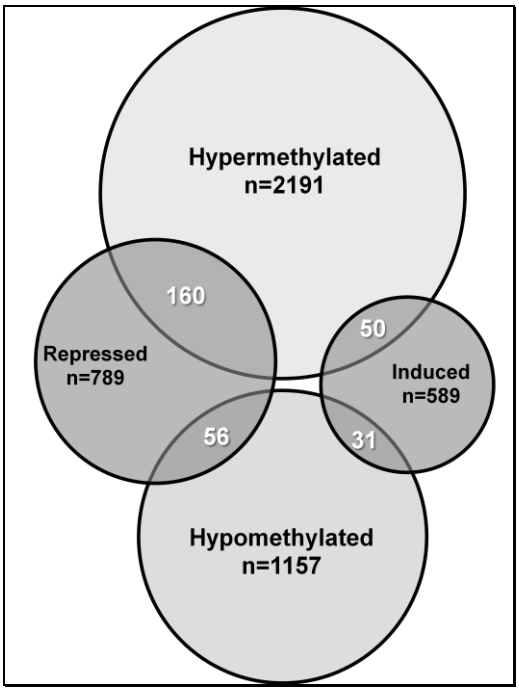

Figure 14. Graphic representation (Venn diagram): merging results of DNA methylation and array-based gene expression data.

Hypermethylated-repressed genes. Table 4 shows the list of hypermethylated-repressed genes subdivided according to their biological function by means of PANTHER classification system. Twenty six genes playing a role in the regulation of cell growth, cycle and proliferation, and in the apoptotic processes were identified by the analysis. Among those, five genes (*FAM107A, IGFALS, MT1G, MT1H* and *RNF180*) likely functioning as candidate tumor-suppressor genes appeared to be highly methylated in the promoter region. A conspicuous number of genes (44 genes) were found to pertain to metabolic and cellular processes regulation, and in particular six genes that are involved in retinol metabolism (*ADH1A, ADH1B, ADH6, CYP3A43, CYP4A22* and *RDH16*) were found to be hypermethylated and repressed. In Figure 15 is reported the retinol metabolism according to Kegg Pathway obtained by David Bioinformatics Database. In the figure are highlighted also the genes that we found down-regulated in HCC tissue, even if the methylation levels were unchanged: *CYP1A1, CYP1A2, CYP2B6, CYP2C9, CYP26A1, CYP3A4, CYP4A11, RDH5* and *LRAT* (Figure 15).

Moreover, PANTHER analysis clustered in this biological process group serine hydroxymethyltransferase 1 (*SHMT1),* a key gene of one-carbon metabolism pathway strictly involved in the methyl groups formation and transfer reactions (Table 4). In the group of hypermethylated and transcriptionally repressed genes in HCC tissue involved in immune response were categorized 23 genes among which was also hepcidin (*HAMP*).

Table 4. Hypermethylated and transcriptionally repressed genes in HCC (160)

Cell communication (17)	Immune response (23)	Metabolic and cellular process (44)	Cell growth, cell cycle and apoptosis (26)	Transport (19)	Others (31)
AKAP2	*ANTXR2*	*ACADS*	*ADORA3*	*ANXA8*	*ADAMTSL2*
AMHR2	*BLNK*	*ACSM5*	*AGTR1*	*APOA5*	*ALPL*
BZRAP1	*C1QTNF1*	*AGMO*	*AR*	*APOL6*	*ANKRD55*
C1orf168	*C1RL*	*AMDHD1*	*AXL*	*AQP7*	*C10orf26*
CLDN1	*C5AR1*	*ANK2*	*CAT*	*CETP*	*C10orf58*
FES	*CCL14*	*ANK3*	*DBH*	*MIP*	*C17orf91*
INHBC	*CCL15*	*ARSD*	*DMD*	*RGN*	*C21orf84*
MORN4	*CD302*	*ATP11C*	*ESR1*	*SLC10A1*	*CCDC68*
OLFML3	*CFI*	*BCO2*	*FGD4*	*SLC22A1*	*CILP*
PDE2A	*CFP*	*CES4A*	*GDF2*	*SLC22A10*	*DNALI1*
PPL	*FCGR2B*	*CHST9*	*JDP2*	*SLC25A25*	*EXPH5*
RIC3	*FCN2*	*CPN1*	*MAP2K3*	*SLC25A47*	*FAM13A*
RND3	*HAMP*	*CYP8B1*	*NAP1L5*	*SLC45A3*	*FAM65C*
SH3D19	*IL13RA2*	*DSE*	*NR4A1*	*SLC47A1*	*FAM83F*
SORBS2	*IL1B*	*EPHX2*	*NUGGC*	*SLC5A1*	*FXYD7*
SUCNR1	*IL1RN*	*FBXO3*	*PTH1R*	*SLC6A12*	*HAPLN4*
VNN1	*KLKB1*	*FMO3*	*PTPN3*	*SLCO1B3*	
	LILRA1	*GLUD2*	*SMAD6*	*SLCO2B1*	
	MBL2	*GPT*	*TBX15*	*TRPV4*	
	MEFV	**Retinol metabolism**	*TNFRSF10D*		
	PGLYRP2	*ADH1A*	*ZBED1*		
	TINAGL1	*ADH1B*	**Candidate tumor-suppressor genes**		
	VSIG4	*ADH6*	*FAM107A*		
		CYP3A43	*MT1G*		
		CYP4A22	*MT1H*		
		RDH16	*RNF180*		
		One-carbon metabolism	*IGFALS*		
		SHMT1			
		GYS2			*INS-IGF2*
		HGFAC			*LINC00574*
		HK3			*LOC339240*
		HOGA1			*LRRC25*
		IDO2			*MYO15A*
		INMT			*PID1*
		INS			*PRSS53*
		IYD			*SMOC1*
		KDM5D			*SPATA18*
		LCAT			*SYNE1*
		LDHD			*TCTEX1D1*
		LPAL2			*TMEM125*
		MOGAT2			*TMEM26*
		OAT			*UNC93A*
		PLIN			*WDR66*
		PSD4			
		TBXA2R			
		UROC1			

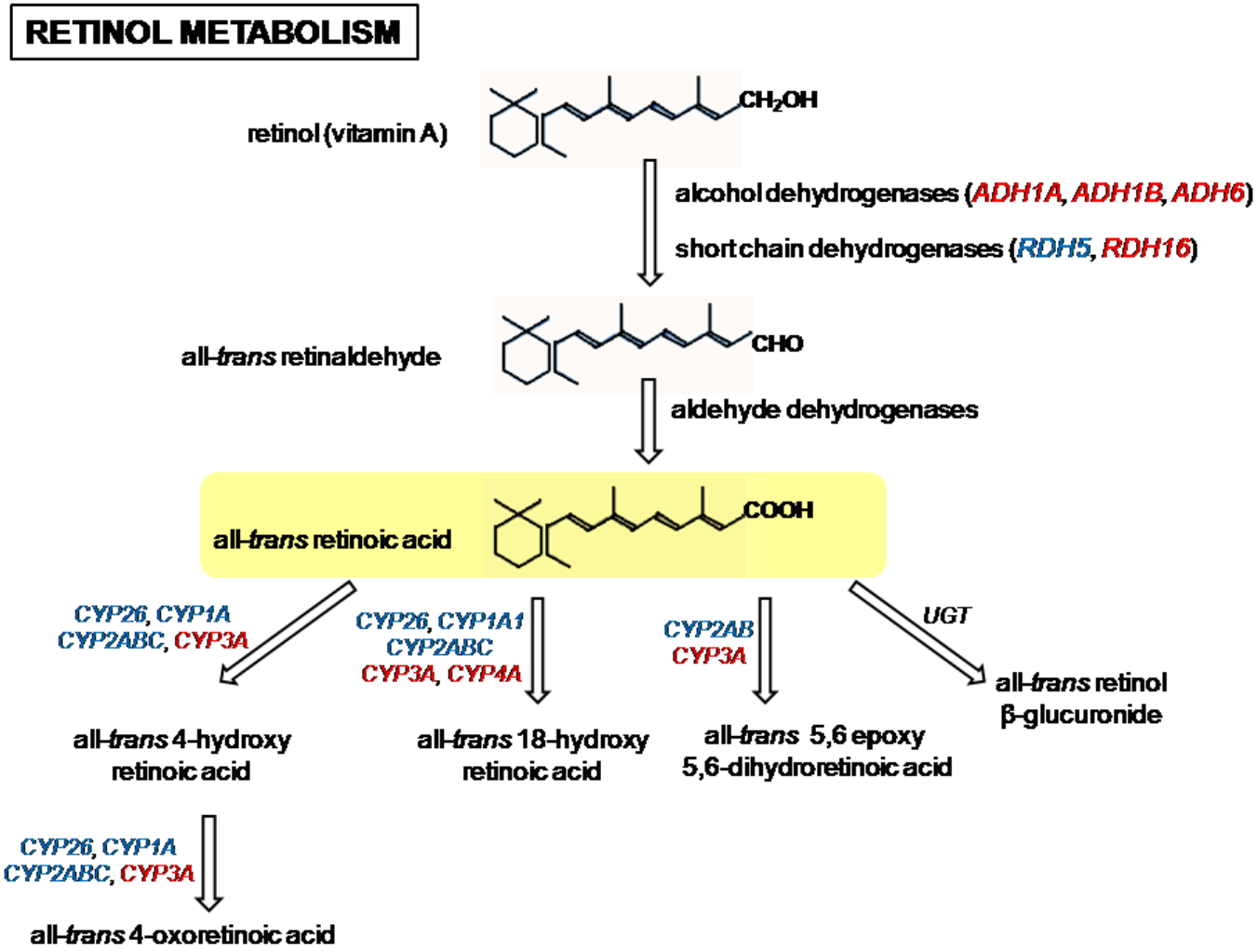

Figure 15. Retinol metabolism. Schematic representation of genes involved in retinol metabolism: genes found hypermethylated and repressed in HCC are in red, transcriptionally repressed genes in blue.

Hypomethylated-induced genes. Thirty-one genes belonged to the group of genes found to be hypomethylated-induced (Table 5). Noticeable was the transcriptional induction associated to low promoter methylation of the following genes: *NOX4* (NADPH oxidase 4) that codifies for a protein implied in the production of various reactive oxygen species, *SPINK1*, also known as Tumor-Associated Trypsin Inhibitor (*TATI*), *ESM1* (endothelial cell-specific molecule 1) that is involved in angiogenesis.

Table 5. Hypomethylated and transcriptionally induced genes in HCC (31)

Cell communication (6)	Immune response (8)	Metabolic and cellular process (4)	Cell growth, cell cycle and apoptosis (4)	Transport (2)	Others (7)
ASAP1	*CD200*	*SPINK1*	*ESM1*	*KIF4B*	*C15orf42*
CD34	*CTLA4*	*DTNA*	*GINS4*	*SLC7A11*	*FBXO32*
GLDN	*CXCL10*	*HIST1H4F*	*LTA*		*KIAA1688*
MYBPC1	*DCSTAMP*	*HIST2H3D*	*MAP2*		*POTEA*
RIMS2	*LRRC69*				*POTEC*
TRIM55	*NOX4*				*VCX2*
	SSX6				*VCX3A*
	SSX8				

Hypermethylated-induced and hypomethylated-repressed genes. The combined DNA methylation-gene expression analysis performed in HCC *versus* non-neoplastic tissue allowed to identify also a number of hypermethylated-induced and hypomethylated-transcriptionally repressed genes (Table 6 and 7). In the group of 50 hypermethylated-induced genes listed in Table 6, emerged the presence of *MMP9* and *MMP12* which are metalloproteinases involved in the breakdown of extracellular matrix. A number of other genes were implicated in the regulation of cell growth, cell cycle and apoptosis. In Table 7 is presented the list of 56 genes found to be hypomethylated-repressed in HCC *versus* non-neoplastic tissue. Notable is the transcriptional repression of two genes possessing a likely function as tumor suppressor, *HEPACAM* (hepatic and glial cell adhesion molecule) and *ABI3BP* (ABI family, member 3 (NESH) binding protein).

Table 6. Hypermethylated and transcriptionally induced genes in HCC (50)

Cell communication (7)	Immune response (3)	Metabolic and cellular process (15)	Cell growth, cell cycle and apoptosis (9)	Transport (5)	Others (11)
BAIAP2L2	*MICB*	*CELF6*	*BAX*	*KIF4A*	*AIM1L*
EPS8L3	*SLAMF8*	*COX7B2*	*BOLA2*	*KPNA2*	*C16orf59*
MCHR1	*VWF*	*DNMT3B*	*BOLA2B*	*SCN4A*	*CSAG1*
PMCH		*HIST1H4l*	*KIAA0101*	*SLC26A6*	*FAM189B*
RASL12		*HKDC1*	*MAGEA5*	*TRIM16L*	*HRCT1*
SEMA3G		*MMP12*	*PLK4*		*MND1*
TNNC1		*MMP9*	*TRAF5*		*PLVAP*
		NEIL3	*TRAIP*		*TRIM31*
		PDE4C	*VRK1*		*VCY*
		PIF1			*VCY1B*
		PLA2G1B			*ZWINT*
		RAB3B			
		S100P			
		UBE2T			
		ZP3			

Table 7. Hypomethylated and transcriptionally repressed genes in HCC (56)

Cell communication (8)	Immune response (9)	Metabolic and cellular process (15)	Cell growth, cell cycle and apoptosis (6)	Transport (7)	Others (11)
CRHBP	*CLEC1B*	*DERA*	*ABI3BP*	*AQP4*	*C14orf105*
DCN	*COLEC10*	*FBXL5*	*CNTN3*	*EHD3*	*DOCK8*
DLG2	*FCRL6*	*FOLH1*	*MACF1*	*LST3TM12*	*ITLN1*
EMR1	*FPR1*	*FRMD4B*	*PDE4DIP*	*LYVE1*	*MBNL2*
GPR128	*IL1RL1*	*GALC*	*PTPN13*	*SLC38A4*	*NEBL*
GRM8	*LILRA5*	*GCNT2*	*TBRG1*	*SLC6A19*	*PAMR1*
IGF1	*MARCO*	*GLYATL1*		*SYTL3*	*PLCXD3*
SPG20	*NLRP12*	*HEPACAM*			*RNF217*
	RAG1	*HSD11B1*			*TMEM100*
		KLHL3			*TMEM133*
		NME5			*ZNF385B*
		PBX1			
		POU6F2			
		RDH14			
		TBXAS1			

5.2.5 Gene expression on RNA extracted from buffy coat

In several genes found differentially methylated and differentially expressed in HCC, gene expression was analyzed in RNA extracted from buffy coat, in order to reveal a possible correlation between blood and hepatic tissue. The subjects analysed were HCC patients, alcoholic patients without hepatic neoplasia and healthy subjects. The gene expression detection of the tested genes is reported in Table 8.

Table 8. Gene expression detection in buffy coat

GENE	Total number of samples	DETECTABLE (n. of samples)	NOT DETECTABLE (n. of samples)
ADH6	24	0	24
DNMT3B	28	24	4
ESM1	24	17	7
ESR1	29	29	0
GDF2	4	0	4
HAMP	4	0	4
NOX4	24	0	24
RDH16	26	20	6
SHMT1	26	25	1
SPINK1	22	0	22

In buffy coat samples *ADH6, GDF2, HAMP, NOX4* and *SPINK1* gene expression levels resulted not detectable while *DNMT3B, ESM1, ESR1, RDH16* and *SHMT1* mRNA levels were detectable. In Figure 16 for each analysed subject the expression levels are represented, calculated according to the formula $\Delta Ct = Ct_{target} - Ct_{18S}$. In all the histograms reported from here on, the values of the ordinate axis (ΔCt) are inverted in order to make data interpretation more immediate.

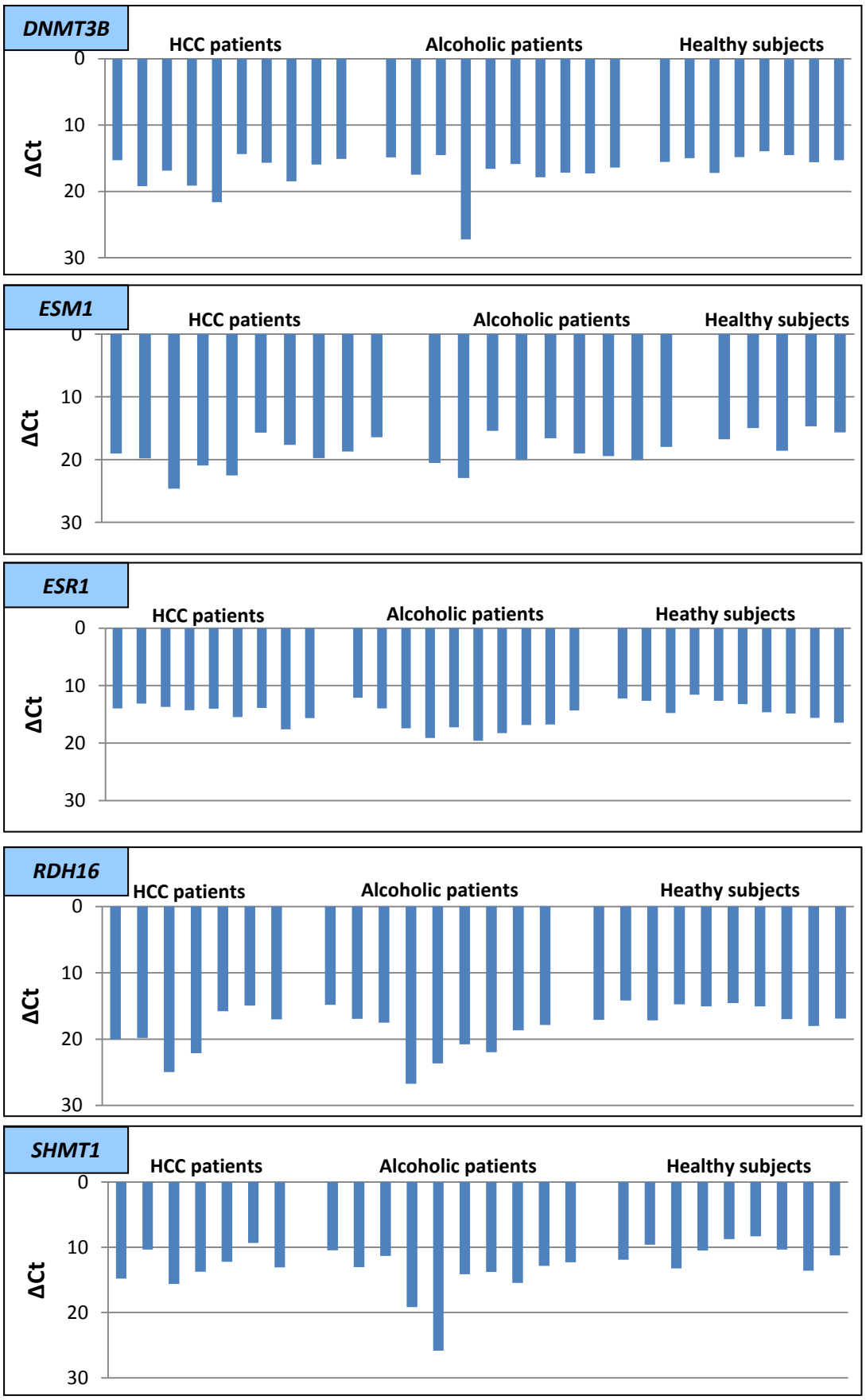

Figure 16. Gene expression levels in buffy coat extracted mRNA. ΔCt values are grouped for HCC patients, alcoholic patients and healthy subjects.

5.2.5.1 Statistical analysis of gene expression data in buffy coat

*HCC patients vs healthy subjects***.** In HCC patients and healthy subjects gene expression was analysed and found to be significantly different for *ESM1* (p=0.024), *RDH16* (p=0.019) and *DNMT3B* (p=0.046); *SHMT1* expression level difference was border line (p=0.086). *RDH16* and *SHMT1* were repressed in HCC patients as expected from the data obtained in the hepatic tissues, while *DNMT3B* and *ESM1* were unexpectedly more expressed in buffy coat of healthy subjects as compared to HCC patients (Figure 17).

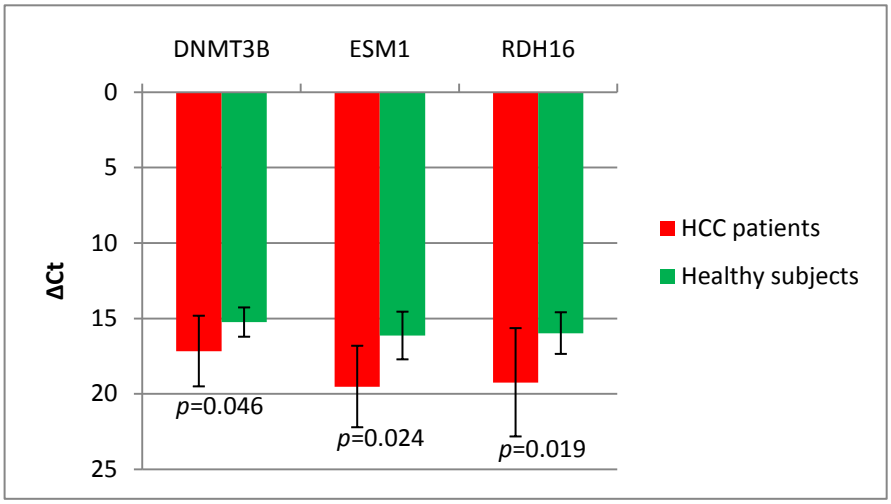

Figure 17. Gene expression in buffy coat of HCC patients compared to healthy subjects. In the histogram are reported genes found differentially expressed in the two groups of subjects.

HCC versus non-HCC. To gain insight into the possible role of cancer in determining the different expression measured for these genes, we compared HCC patients with cancer-free subjects, represented by alcoholic patients and healthy subjects. No statistically significant differences could be demonstrated for all the five genes.

*Alcohol versus non-alcohol***.** The statistical analysis based on alcohol consumption is shown in Figure 18. The comparison was performed between "alcohol", represented by HCC and alcoholic patients, and "non-alcohol", represented by healthy subjects. Statistically significant differences were found for *ESM1* (p=0.011), *ESR1* (p=0.030), *RDH16* (p=0.005) and *SHMT1* (p=0.031). *ESM1*, *ESR1*, *RDH16* and *SHMT1* showed lower mRNA levels in buffy coat of subjects characterized by a high alcohol intake as compared with controls (Figure 18).

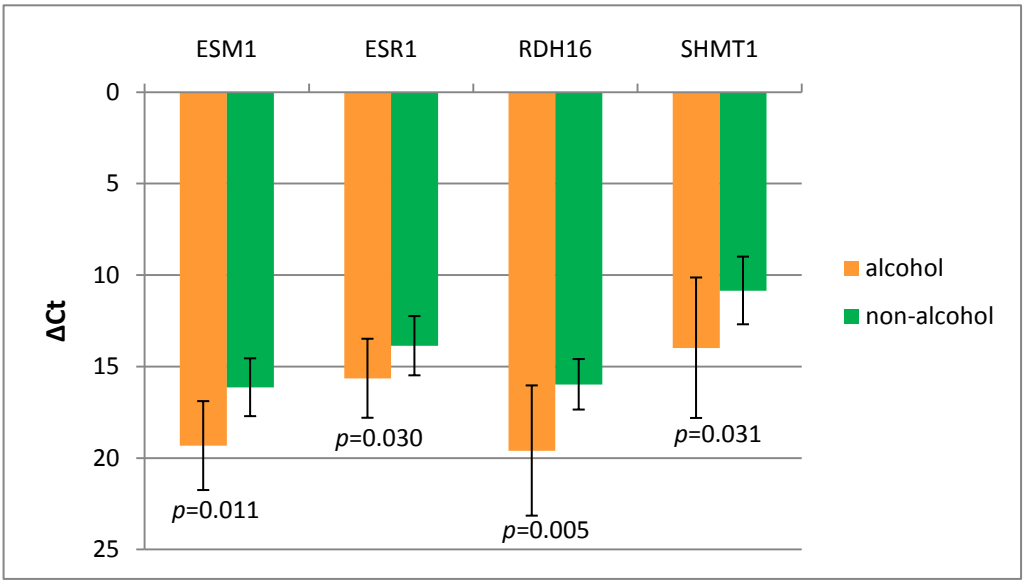

Figure 18. Gene expression in buffy coat of patients with high alcohol intake compared to controls. The histograms represent genes with significant difference in mRNA levels between the two groups. The "alcohol" groups is represented by HCC and alcoholic patients; the "non-alcohol" groups is represented by healthy subjects.

Alcoholic patients vs healthy subjects. When comparing alcoholic patients with healthy subjects, four genes displayed statistically significant differences: *ESM1* (p=0.022), *ESR1* (p=0.009), *RDH16* (p=0.007) and *SHMT1* (p=0.026) (Figure 19). These results once more support the possible role of alcohol in the down-regulation of these genes.

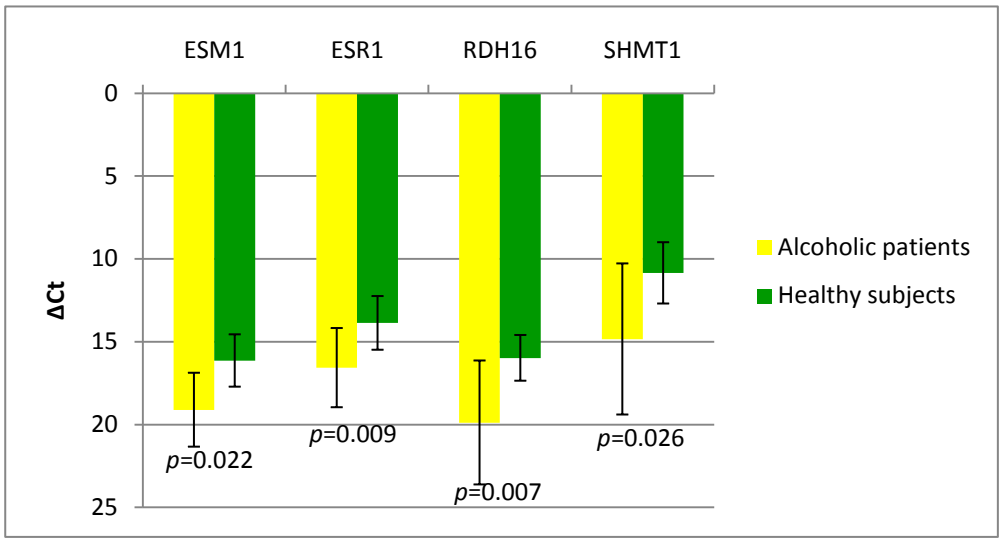

Figure 19. Gene expression in buffy coat of alcoholic patients compared to healthy subjects. In the histogram are reported the genes with statistically significant difference in mRNA levels between the two groups.

6. DISCUSSION

The present study was designed to define promoter DNA methylation and transcription profiles in liver tissue of non-viral, alcohol-related HCC to highlight possible novel epigenetic signatures.

The high interest toward alcohol-linked liver cancer is related to the known inter-relationship between alcohol and epigenetic mechanisms modulation *via* one-carbon metabolism and the potential reversibility of epigenetic mechanisms by influencing nutritional factors including alcohol intake.

To evaluate whether DNA methylation affect gene expression at specific *loci*, methylation and gene expression data obtained from the analysis of cancer and normal liver tissue were combined, showing that the expression of a number of genes is indeed modulated by differential methylation at promoter site.

Furthermore, by the evaluation of specific gene expression profiles in both liver as well as PBMCs we intended to pursue the objective of identifying possible new candidate blood biomarkers for primary liver cancer disease. To our knowledge this is the first study crossing over DNA methylation and gene expression profiling in alcohol-related hepatocellular carcinoma and evaluating gene expression of possible candidate genes in peripheral blood cells.

6.1 Subjects enrolment and methodological optimization

One of the strength of the study is related to the strict criteria of selection of the subjects participating to the study that allowed analyzing the role of alcohol consumption in hepatic carcinogenesis in absence of other confounding factors, *i.e.* mainly HBV and HCV. Autoimmune pathogenesis underlying and potentially leading towards chronic liver disease was also excluded from the present study. Considering the inter-relationships among

alcohol, one-carbon metabolism, carcinogenesis and methylation of DNA, this study was designed precisely with the idea of unraveling markers of disease that, being nutritionally related with an epigenetic regulation may be potentially modifiable.

The MeDIP-chip analysis was performed on neoplastic and homologous non-neoplastic tissues in eight HCC patients that were carefully selected in order to be free of former viral infections to have a positive anamnesis for chronic alcohol intake and to be free of severe liver derangement according to the Child-Pugh score. The comparison of cancer *versus* cancer-free liver tissue from the same subject allowed the exclusion of possible spurious environmental factors that may certainly influence epigenetic features and, therefore, be strong confounders for data interpretation.

The application of high-throughput techniques, both for DNA methylation and gene-expression analysis allowed exploring the methylation profiles of all annotated genes and therefore highlighted possible novel pathways, epigenetically regulated and with a role in alcohol-related HCC. By the identification of genes regulated by methylation in alcohol-linked liver cancer it could be hypothesized to influence the disease risk by nutritional intervention able to influence both alcohol metabolism and epigenetic phenomena such as DNA methylation and consequently influence gene expression.

6.2 DNA methylation and gene expression profile in neoplastic and non-neoplastic tissues

By comparing the methylation profile of neoplastic *versus* homologous non-neoplastic liver tissues, 2,401 hypermethylated and 1,244 hypomethylated genes were identified in primary liver cancer. A large number of genes are related to several carcinogenesis-involved pathways such as apoptosis, cell communication, cell cycle and immune system processes. Since our goal was to investigate the possible role of DNA methylation in

alcohol-mediated carcinogenesis, we narrowed down the analysis of results on differentially methylated genes showing significant variations in transcriptional levels. The merging of promoter methylation values and gene expression results allowed the identification of four groups of genes: 160 hypermethylated-repressed, 31 hypomethylated-induced, 50 hypermethylated-induced and 56 hypomethylated-repressed genes.

The presence of hypermethylated-repressed and hypomethylated-induced genes supports previous reports describing that promoter DNA methylation is able to interfere with the formation of transcriptional complexes leading to gene repression (Bird, 1986; Luczak and Jagodzinski, 2006). On the other hand, promoter demethylation has been described to initiate the transcriptional processes (Luczak and Jagodzinski, 2006). More uncertain is the role of methylation in the enhancement of expression observed in hypermethylated promoter gene regions and, respectively, in the repression of transcription in hypomethylated genes. One could hypothesize that, in certain genes, transcription regulation is independent from promoter methylation, or that DNA methylation might affect transcription by various mechanisms, likely interfering with the bond of transcription enhancer/silencer or involving a different arrangement of chromatin structure.

6.3 FAM107A, RNF180 and *MT1H*: *new candidate tumor-suppressor genes in HCC*

To our knowledge, this is the first report describing the methylation-mediated repression of *FAM107A*, *RNF180* and *MT1H* (Table 4) in HCC tissue. The function of these genes as tumor-suppressors has been already demonstrated in several neoplastic diseases (Awakura et al., 2008; Liu et al., 2009; Wilson et al., 2010; Bell et al., 2011; Cheung et al., 2012). *FAM107A* has been identified in renal cell carcinoma as tumor suppressor gene (Awakura et al., 2008) according to its role in the regulation of apoptotic process. Scarce are the reports about *RNF180*. Very recently, *RNF180* has been characterized and found to be

hypermethylated and silenced in gastric cancer, where it seems to play a role as apoptosis regulator (Cheung et al., 2012). *MT1H* codifies for a metallothionein (MT), a class of proteins involved in the process of ROS and heavy metals cellular detoxification and this function can be involved in tumor suppression mechanisms. The methylation-mediated transcriptional repression of another metallothionein, *MT1G,* that we also found hypermethylated and repressed, has been described previously described in HCC (Kanda et al., 2009), therefore it seems reasonable to hypothesize in HCC a similar role for *MT1H,* that could then represent a new candidate tumor-suppressor gene.

Furthermore our results confirm the finding that the expression of the tumor-suppressor gene *IGFALS* is silenced by methylation in HCC (Neumann et al., 2012).

6.4 DNA methylation: the missing link between retinol metabolism and alcohol

In the cluster of hypermethylated and transcriptionally repressed genes, remarkable was the finding of six genes (*ADH1A*, *ADH1B*, *ADH6*, *CYP3A43*, *CYP4A22* and *RDH16*) associated to retinol metabolism. Moreover, several other genes belonging to the retinol metabolism pathway resulted down-regulated in HCC tissue, even if the methylation levels were unchanged: *CYP1A1, CYP1A2, CYP2B6, CYP2C9, CYP26A1, CYP3A4, CYP4A11, RDH5* and *LRAT*.

Liver is the most important organ involved in retinoid storage and metabolism, and retinoids, namely vitamin A and its derivatives, are known to play important roles in the development of hepatic diseases (steatosis, fibrosis, cirrhosis and HCC) (Shirakami et al., 2012). Retinoids are involved in the regulation of cellular growth, cellular differentiation and apoptosis. Moreover, chronic ethanol intake has been described to impair retinoic acid homeostasis, thus playing a role in the development of alcohol-related cancers (Wang, 2005). In particular, alcohol interferes with retinol metabolism through different

mechanisms: a. as a competitive inhibitor of vitamin A oxidation to retinoic acid involving ADHs (alcohol dehydrogenases) and ALDHs (acetaldehyde dehydrogerases); b. by inducing cytochrome P450 and the catabolism of vitamin A and retinoic acid; c. by increasing vitamin A mobilization from the liver to extrahepatic tissues (Wang, 2005).

Our results suggest that DNA methylation-mediated down-regulation of specific genes may represent a mechanism responsible for the derangement of retinol metabolism associated with chronic alcohol consumption. DNA methylation could, therefore, be regarded as the missing link between alcohol intake, retinol metabolism impairment and hepatic carcinogenesis.

6.5 SHMT1 and one-carbon metabolism

Interestingly, our results showed, for the first time, the occurrence of *SHMT1* gene repression mediated by enhanced promoter methylation. *SHMT1* is a key gene of one-carbon metabolism, that acts to reversibly convert serine and tetrahydrofolate to glycine and 5,10-methylene tetrahydrofolate (Figure 3). SHMT1 enzyme plays a central role in folate metabolism since it operates as a metabolic switch between nucleotide synthesis reactions and biological methylation pathways (Herbig et al., 2002).

Other genes involved in one-carbon metabolism were found transcriptionally repressed in HCC tissue, although the methylation levels were unchanged such as the case of *BHMT1*, *BHMT2*, *CBS*, *GNMT*, MTHFD2L or decreased in *FOLH1* (Table 7). All these genes exert their activity at different crucial nodes of the pathway of methyl units transfer, essential for both biological methylation and nucleotide synthesis reactions. The apparent one-carbon metabolism derangement observed in the present study is in accordance with alcohol-induced alterations of methyl transfer reactions demonstrated in human and animal studies.

6.6 DNA methylation mediates hepcidin down-regulation in HCC

Among hypermethylated-repressed genes very interesting was the finding of hepcidin, a liver peptide hormone involved in iron homeostasis and in the innate immune response (Park et al., 2001; Ganz, 2009). Hepcidin has been already shown to be repressed in unselected HCC cases (Kijima et al., 2008) and transcriptionally repressed in the liver of alcoholic patients (Costa-Matos et al., 2012). According to our results, we report a putative role for DNA methylation in the transcriptional repression of hepcidin in non-viral alcohol-related HCC.

6.7 Hypomethylation mediates up-regulation of NOX4, SPINK1 and ESM1 in HCC

In hepatocellular carcinoma the overexpression of *NOX4* (Carmona-Cuenca et al., 2008; Caja et al., 2011), *SPINK1* (Lee et al., 2007) and *ESM1* (Chen et al., 2010; Kang et al., 2011) has been related to their role in oxidative stress defense, regulation of tumor growth and angiogenesis. For the first time we can hypothesize the possible role of promoter hypomethylation in the transcriptional up-regulation of these genes in HCC.

6.8 Gene expression on RNA extracted from buffy coat

The transcriptional levels of several genes that were found to be differentially methylated and diversely expressed in HCC tissue, were also studied in PBMCs samples from mRNA obtained from buffy coat, in order to determine a possible correlation between liver and buffy coat gene expression pattern. Among the ten genes analyzed, five were not detectable in the buffy coat analysis, *i.e. ADH6, GDF2, HAMP*, *NOX4* and *SPINK1*, and this is in accordance with the data reported in TiGER (Tissue-specific Gene Expression and Regulation) database (Liu et al., 2008).

On the contrary, *DNMT3B, ESM1, ESR1, RDH16* and *SHMT1* mRNA levels were detectable in buffy coat samples. *DNMT3B* and *ESR1* are reported by TiGER database as expressed in blood cells, while *ESM1, RDH16* and *SHMT1* expression levels in blood are described here for the first time.

The gene expression evaluation was performed in three groups of subjects: HCC patients, alcoholic patients without hepatic neoplasia and healthy subjects.

In the comparison between HCC patients and healthy subjects, very interesting was the finding of *RDH16*, a gene belonging to the retinol metabolism, significantly less expressed in PBMCs of HCC patients compared to healthy subjects. Moreover, it should be also highlighted that *SHMT1* expression was also lower in PBMCs of HCC patients although with a borderline p-value. One may assume that the statistical power was impaired by the relatively small number of subjects analyzed and further studies in a larger sample set may be helpful to better clarify the role of SHMT1 in HCC.

RDH16, also named *RoDH4*, encodes for a short chain dehydrogenase/reductase (SDR) that converts retinol to retinaldehyde (Perlmann, 2002); this enzyme is not inhibited by ethanol, as it usually occurs with medium chain alcohol dehydrogenases (ADHs), since it does not utilize ethanol as a substrate (Shirakami et al., 2012). Our results on alcohol-related HCC suggest a role for DNA methylation in the down-regulation of the gene instead of an inhibition of the enzyme mediated by alcohol. Furthermore, the finding of *RDH16* gene repression also in PBMCs of patients affected by HCC as compared to healthy subjects suggests the possibility of using differential *RDH16* gene expression data as a possibly useful biomarker for the predisposition to liver cancer development.

To better clarify the role of alcohol intake in gene expression regulation in PBMCs , we compared gene expression levels in alcoholic patients with or without liver cancer *versus*

not alcohol drinkers. Interestingly enough, high alcohol intake was associated with a statistically significant down-regulation of *RDH16, ESR1* and *SHMT1* expression.

While the *RDH16* gene down-regulation was associated to the presence of HCC, *ESR1* and *SHMT1* gene repression in PBMCs seemed to be linked only to alcohol intake and not to the presence of cancer disease.

ESR1 encodes for an estrogen receptor, a nuclear transcription factor that is activated by the binding of estrogens and regulates the expression of specific target genes. Estrogens and their receptors appear to play a significant role in carcinogenetic processes of all hormone-sensitive organs. In particular, in HCC, a cancer displaying higher incidence in males than in females, the role of estrogens and of their receptors has been widely investigated but it remains still poorly understood (Kalra et al., 2008). The role of estrogens in HCC has been controversial with evidence suggesting both carcinogenic and protective effects in the liver. Nevertheless, recent studies highlighted the protective role (Naugler et al., 2007) and the suppressive effects of estrogens in HCC development (Xu et al., 2012). Our findings, showing a hypermethylated and transcriptionally repressed *ESR1* gene in non-viral, alcohol-related HCC may suggest that the down-regulation of estrogen receptors could contribute to make HCC tissue less responsive to the suppressive effects of estrogens. Moreover, our data on PBMCs, displaying a significant less expression in alcoholic patients as compared to healthy controls, suggest a possible influence of alcohol on *ESR1* gene expression even before cancer development.

SHMT1, a key gene of one-carbon metabolism, has been studied in relation to different type of cancers, such as cancer of the gastroenteric tract (Macfarlane et al., 2011) and, more specifically, in rectal cancer (Komlosi et al., 2010) and also in lymphoma (Weiner et al., 2011). SHMT1 regulates the partitioning of folate-activated one-carbons between thymidylate and S-adenosylmethionine biosynthesis. Therefore, changes in SHMT1

expression enable the determination of the specific contributions made by thymidylate and S-adenosylmethionine biosynthesis to colorectal cancer risk. However, it should be taken into account that all those studies mostly evaluated indirectly the function of this gene by studying the association of the *SHMT1 1420C>T* functional polymorphism with tumour development. The presence of this polymorphism appears mainly to be repressive for gene function (Heil et al., 2001) although data are not always consistent in this regard. Our results correlating alcohol intake and *SHMT1* gene repression in PBMCs of alcoholic patients are new findings suggesting that a deeper study of the role of this gene may open up new interesting perspectives as a possible biomarkers of non-viral alcohol-related HCC.

7. CONCLUSIONS

Results from the present study allowed the observation of epigenetic regulation by methylation at promoter site of retinol-associated genes, namely *ADH1A, ADH1B, ADH6, CYP3A43, CYP4A22* and *RDH16* in alcohol-related, non-viral HCC.

One-carbon metabolism, well-known to be linked to both alcohol metabolism and DNA methylation, is also involved through the epigenetic transcriptional repression of a major enzyme, *SHMT1* that has been found to be repressed by methylation at his promoter site in HCC. Noteworthy, *ESR1*, a transcription factor with a hormone-binding domain involved in cell cycle regulation, and hepcidin, a liver peptide hormone involved in iron homeostasis were also identified as epigenetically regulated through DNA methylation inducing transcriptional repression. Furthermore, the gene expression analysis on mRNA extracted from PBMCs rich-buffy coat of HCC patients, alcoholic patients without liver cancer and healthy subjects revealed that transcriptional repression of *RDH16* was significantly associated with hepatic cancer. Thus suggesting that specific genes from PBMCs DNA may be useful biomarkers for HCC.

Moreover, the expression of *RDH16, SHMT1*and *ESR1* was associated to chronic alcohol intake compared to controls. Considering that epigenetic phenomena are potentially reversible and influence by nutritional factors such as alcohol intake, these results may have important implications for preventive strategies.

8. REFERENCES

Awakura Y., Nakamura E., Ito N., Kamoto T. and Ogawa O. (2008). Methylation-associated silencing of TU3A in human cancers. Int J Oncol 33(4): 893-9.

Bell A., Bell D., Weber R. S. and El-Naggar A. K. (2011). CpG island methylation profiling in human salivary gland adenoid cystic carcinoma. Cancer 117(13): 2898-909.

Benjamini Y. and Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Statist. Soc. Ser.B 57: 289-300.

Bird A. (2007). Perceptions of epigenetics. Nature 447(7143): 396-8.

Bird A. P. (1986). CpG-rich islands and the function of DNA methylation. Nature 321(6067): 209-13.

Boland M. J. and Christman J. K. (2009). Mammalian DNA methyltransferases. Nutrients and Epigenetics. S. W. Choi and S. Friso. Boca Raton, CRC Press

Boujedidi H., Bouchet-Delbos L., Cassard-Doulcier A. M., Njike-Nakseu M., Maitre S., Prevot S., Dagher I., Agostini H., Voican C. S., Emilie D., Perlemuter G. and Naveau S. (2012). Housekeeping gene variability in the liver of alcoholic patients. Alcohol Clin Exp Res 36(2): 258-66.

Caja L., Sancho P., Bertran E. and Fabregat I. (2011). Dissecting the effect of targeting the epidermal growth factor receptor on TGF-beta-induced-apoptosis in human hepatocellular carcinoma cells. J Hepatol 55(2): 351-8.

Carmona-Cuenca I., Roncero C., Sancho P., Caja L., Fausto N., Fernandez M. and Fabregat I. (2008). Upregulation of the NADPH oxidase NOX4 by TGF-beta in hepatocytes is required for its pro-apoptotic activity. J Hepatol 49(6): 965-76.

Chen L. Y., Liu X., Wang S. L. and Qin C. Y. (2010). Over-expression of the Endocan gene in endothelial cells from hepatocellular carcinoma is associated with angiogenesis and tumour invasion. J Int Med Res 38(2): 498-510.

Cheung K. F., Lam C. N., Wu K., Ng E. K., Chong W. W., Cheng A. S., To K. F., Fan D., Sung J. J. and Yu J. (2012). Characterization of the gene structure, functional significance, and clinical application of RNF180, a novel gene in gastric cancer. Cancer 118(4): 947-59.

Choi S. W. and Mason J. B. (2002). Folate status: effects on pathways of colorectal carcinogenesis. J Nutr 132(8 Suppl): 2413S-2418S.

Costa-Matos L., Batista P., Monteiro N., Simoes M., Egas C., Pereira J., Pinho H., Santos N., Ribeiro J., Cipriano M. A., Henriques P., Girao F., Rodrigues A. and Carvalho A. (2012). Liver hepcidin mRNA expression is inappropriately low in alcoholic patients compared with healthy controls. Eur J Gastroenterol Hepatol.

Cravo M. L., Gloria L. M., Selhub J., Nadeau M. R., Camilo M. E., Resende M. P., Cardoso J. N., Leitao C. N. and Mira F. C. (1996). Hyperhomocysteinemia in chronic

alcoholism: correlation with folate, vitamin B-12, and vitamin B-6 status. Am J Clin Nutr 63(2): 220-4.

Down T. A., Rakyan V. K., Turner D. J., Flicek P., Li H., Kulesha E., Graf S., Johnson N., Herrero J., Tomazou E. M., Thorne N. P., Backdahl L., Herberth M., Howe K. L., Jackson D. K., Miretti M. M., Marioni J. C., Birney E., Hubbard T. J., Durbin R., Tavare S. and Beck S. (2008). A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 26(7): 779-85.

Ehrlich M. (2002). DNA methylation in cancer: too much, but also too little. Oncogene 21(35): 5400-13.

Ehrlich M. (2006). Cancer-linked DNA hypomethylation and its relationship to hypermethylation. Curr Top Microbiol Immunol 310: 251-74.

Feber A., Wilson G. A., Zhang L., Presneau N., Idowu B., Down T. A., Rakyan V. K., Noon L. A., Lloyd A. C., Stupka E., Schiza V., Teschendorff A. E., Schroth G. P., Flanagan A. and Beck S. (2011). Comparative methylome analysis of benign and malignant peripheral nerve sheath tumors. Genome Res 21(4): 515-24.

Feinberg A. P. (2007). Phenotypic plasticity and the epigenetics of human disease. Nature 447(7143): 433-40.

Feinberg A. P. and Tycko B. (2004). The history of cancer epigenetics. Nat Rev Cancer 4(2): 143-53.

Friso S. and Choi S. W. (2002). Gene-nutrient interactions and DNA methylation. J Nutr 132(8 Suppl): 2382S-2387S.

Friso S., Lotto V., Choi S. W., Girelli D., Pinotti M., Guarini P., Udali S., Pattini P., Pizzolo F., Martinelli N., Corrocher R., Bernardi F. and Olivieri O. (2012). Promoter methylation in coagulation F7 gene influences plasma FVII concentrations and relates to coronary artery disease. J Med Genet 49(3): 192-9.

Friso S., Udali S., Guarini P., Pellegrini C., Pattini P., Moruzzi S., Girelli D., Pizzolo F., Martinelli N., Corrocher R., Olivieri O. and Choi S. W. (2013). Global DNA hypomethylation in peripheral blood mononuclear cells as a biomarker of cancer risk. Cancer Epidemiol Biomarkers Prev.

Ganz T. (2009). Iron in innate immunity: starve the invaders. Curr Opin Immunol 21(1): 63-7.

Goldberg A. D., Allis C. D. and Bernstein E. (2007). Epigenetics: a landscape takes shape. Cell 128(4): 635-8.

Hamid A. and Kaur J. (2006). Chronic alcoholism alters the transport characteristics of folate in rat renal brush border membrane. Alcohol 38(1): 59-66.

Heil S. G., Van der Put N. M., Waas E. T., den Heijer M., Trijbels F. J. and Blom H. J. (2001). Is mutated serine hydroxymethyltransferase (SHMT) involved in the etiology of neural tube defects? Mol Genet Metab 73(2): 164-72.

Herbig K., Chiang E. P., Lee L. R., Hills J., Shane B. and Stover P. J. (2002). Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. J Biol Chem 277(41): 38381-9.

Herceg Z. and Paliwal A. (2011). Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome. Mutat Res 727(3): 55-61.

Heyn H. and Esteller M. (2012). DNA methylation profiling in the clinic: applications and challenges. Nat Rev Genet 13(10): 679-92.

Hou L., Wang H., Sartori S., Gawron A., Lissowska J., Bollati V., Tarantini L., Zhang F. F., Zatonski W., Chow W. H. and Baccarelli A. (2010). Blood leukocyte DNA hypomethylation and gastric cancer risk in a high-risk Polish population. Int J Cancer 127(8): 1866-74.

Huang da W., Sherman B. T. and Lempicki R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37(1): 1-13.

Huang da W., Sherman B. T. and Lempicki R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4(1): 44-57.

Illingworth R. S. and Bird A. P. (2009). CpG islands--'a rough guide'. FEBS Lett 583(11): 1713-20.

Irizarry R. A., Hobbs B., Collin F., Beazer-Barclay Y. D., Antonellis K. J., Scherf U. and Speed T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4(2): 249-64.

Jemal A., Bray F., Center M. M., Ferlay J., Ward E. and Forman D. (2011). Global cancer statistics. CA Cancer J Clin 61(2): 69-90.

Jia D. and Cheng X. (2009). Methylation on the nucleosome. Nutrients and Epigenetics. F. S. Choi S.W. Boca Raton.

Jones P. A. (1986). DNA methylation and cancer. Cancer Res 46(2): 461-6.

Jones P. A. and Baylin S. B. (2007). The epigenomics of cancer. Cell 128(4): 683-92.

Jones P. A. and Laird P. W. (1999). Cancer epigenetics comes of age. Nat Genet 21(2): 163-7.

Kalra M., Mayes J., Assefa S., Kaul A. K. and Kaul R. (2008). Role of sex steroid receptors in pathobiology of hepatocellular carcinoma. World J Gastroenterol 14(39): 5945-61.

Kanda M., Nomoto S., Okamura Y., Nishikawa Y., Sugimoto H., Kanazumi N., Takeda S. and Nakao A. (2009). Detection of metallothionein 1G as a methylated tumor suppressor gene in human hepatocellular carcinoma using a novel method of double combination array analysis. Int J Oncol 35(3): 477-83.

Kang Y. H., Ji N. Y., Lee C. I., Lee H. G., Kim J. W., Yeom Y. I., Kim D. G., Yoon S. K., Park P. J. and Song E. Y. (2011). ESM-1 silencing decreased cell survival, migration, and invasion and modulated cell cycle progression in hepatocellular carcinoma. Amino Acids 40(3): 1003-13.

Kijima H., Sawada T., Tomosugi N. and Kubota K. (2008). Expression of hepcidin mRNA is uniformly suppressed in hepatocellular carcinoma. BMC Cancer 8: 167.

Komlosi V., Hitre E., Pap E., Adleff V., Reti A., Szekely E., Biro A., Rudnai P., Schoket B., Muller J., Toth B., Otto S., Kasler M., Kralovanszky J. and Budai B. (2010). SHMT1 1420 and MTHFR 677 variants are associated with rectal but not colon cancer. BMC Cancer 10: 525.

Lee Y. C., Pan H. W., Peng S. Y., Lai P. L., Kuo W. S., Ou Y. H. and Hsu H. C. (2007). Overexpression of tumour-associated trypsin inhibitor (TATI) enhances tumour growth and is associated with portal vein invasion, early recurrence and a stage-independent prognostic factor of hepatocellular carcinoma. Eur J Cancer 43(4): 736-44.

Lim U., Flood A., Choi S. W., Albanes D., Cross A. J., Schatzkin A., Sinha R., Katki H. A., Cash B., Schoenfeld P. and Stolzenberg-Solomon R. (2008). Genomic methylation of leukocyte DNA in relation to colorectal adenoma among asymptomatic women. Gastroenterology 134(1): 47-55.

Liu Q., Zhao X. Y., Bai R. Z., Liang S. F., Nie C. L., Yuan Z., Wang C. T., Wu Y., Chen L. J. and Wei Y. Q. (2009). Induction of tumor inhibition and apoptosis by a candidate tumor suppressor gene DRR1 on 3p21.1. Oncol Rep 22(5): 1069-75.

Liu X., Yu X., Zack D. J., Zhu H. and Qian J. (2008). TiGER: a database for tissue-specific gene expression and regulation. BMC Bioinformatics 9: 271.

Lu S. C., Martinez-Chantar M. L. and Mato J. M. (2006). Methionine adenosyltransferase and S-adenosylmethionine in alcoholic liver disease. J Gastroenterol Hepatol 21 Suppl 3: S61-4.

Lu S. C. and Mato J. M. (2005). Role of methionine adenosyltransferase and S-adenosylmethionine in alcohol-associated liver cancer. Alcohol 35(3): 227-34.

Luczak M. W. and Jagodzinski P. P. (2006). The role of DNA methylation in cancer development. Folia Histochem Cytobiol 44(3): 143-54.

Macfarlane A. J., Perry C. A., McEntee M. F., Lin D. M. and Stover P. J. (2011). Shmt1 heterozygosity impairs folate-dependent thymidylate synthesis capacity and modifies risk of Apc(min)-mediated intestinal cancer risk. Cancer Res 71(6): 2098-107.

Magdalena J. and Goval J. J. (2009). Methyl DNA immunoprecipitation. Methods Mol Biol 567: 237-47.

Naugler W. E., Sakurai T., Kim S., Maeda S., Kim K., Elsharkawy A. M. and Karin M. (2007). Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. Science 317(5834): 121-4.

Neumann O., Kesselmeier M., Geffers R., Pellegrino R., Radlwimmer B., Hoffmann K., Ehemann V., Schemmer P., Schirmacher P., Bermejo J. L. and Longerich T. (2012). Methylome analysis and integrative profiling of human HCCs identify novel protumorigenic factors. Hepatology.

Park C. H., Valore E. V., Waring A. J. and Ganz T. (2001). Hepcidin, a urinary antimicrobial peptide synthesized in the liver. J Biol Chem 276(11): 7806-10.

Perlmann T. (2002). Retinoid metabolism: a balancing act. Nat Genet 31(1): 7-8.

Pogribny I. P. and Rusyn I. (2012). Role of epigenetic aberrations in the development and progression of human hepatocellular carcinoma. Cancer Lett.

Poschl G. and Seitz H. K. (2004). Alcohol and cancer. Alcohol Alcohol 39(3): 155-65.

Pufulete M., Al-Ghnaniem R., Leather A. J., Appleby P., Gout S., Terry C., Emery P. W. and Sanders T. A. (2003). Folate status, genomic DNA hypomethylation, and risk of colorectal adenoma and cancer: a case control study. Gastroenterology 124(5): 1240-8.

Rakyan V. K., Down T. A., Thorne N. P., Flicek P., Kulesha E., Graf S., Tomazou E. M., Backdahl L., Johnson N., Herberth M., Howe K. L., Jackson D. K., Miretti M. M., Fiegler H., Marioni J. C., Birney E., Hubbard T. J., Carter N. P., Tavare S. and Beck S. (2008). An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Res 18(9): 1518-29.

Seitz H. K. and Stickel F. (2007). Molecular mechanisms of alcohol-mediated carcinogenesis. Nat Rev Cancer 7(8): 599-612.

Shirakami Y., Lee S. A., Clugston R. D. and Blaner W. S. (2012). Hepatic metabolism of retinoids and disease associations. Biochim Biophys Acta 1821(1): 124-36.

Smyth G. K. (2005). Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions using R and Bioconductor. R. Gentleman, V. Carey, S. Dudoit, R. Irizarry and W. Huber. New York, Springer**:** 397–420.

Terry M. B., Delgado-Cruzata L., Vin-Raviv N., Wu H. C. and Santella R. M. (2011). DNA methylation in white blood cells: association with risk factors in epidemiologic studies. Epigenetics 6(7): 828-37.

Thomas P. D., Kejariwal A., Campbell M. J., Mi H., Diemer K., Guo N., Ladunga I., Ulitsky-Lazareva B., Muruganujan A., Rabkin S., Vandergriff J. A. and Doremieux O. (2003). PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Res 31(1): 334-41.

Udali S., Guarini P., Moruzzi S., Choi S. W. and Friso S. Cardiovascular epigenetics: From DNA methylation to microRNAs. Mol Aspects Med.

Wang X. D. (2005). Alcohol, vitamin A, and cancer. Alcohol 35(3): 251-8.

Wani N. A., Hamid A., Khanduja K. L. and Kaur J. (2012). Folate malabsorption is associated with down-regulation of folate transporter expression and function at colon basolateral membrane in rats. Br J Nutr 107(6): 800-8.

Weber M., Davies J. J., Wittig D., Oakeley E. J., Haase M., Lam W. L. and Schubeler D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37(8): 853-62.

Weiner A. S., Beresina O. V., Voronina E. N., Voropaeva E. N., Boyarskih U. A., Pospelova T. I. and Filipenko M. L. (2011). Polymorphisms in folate-metabolizing genes and risk of non-Hodgkin's lymphoma. Leuk Res 35(4): 508-15.

Wilson A. J., Chueh A. C., Togel L., Corner G. A., Ahmed N., Goel S., Byun D. S., Nasser S., Houston M. A., Jhawer M., Smartt H. J., Murray L. B., Nicholas C., Heerdt B. G., Arango D., Augenlicht L. H. and Mariadason J. M. (2010). Apoptotic sensitivity of colon cancer cells to histone deacetylase inhibitors is mediated by an Sp1/Sp3-activated transcriptional program involving immediate-early gene induction. Cancer Res 70(2): 609-20.

Xu H., Wei Y., Zhang Y., Xu Y., Li F., Liu J., Zhang W., Han X., Tan R. and Shen P. (2012). Oestrogen attenuates tumour progression in hepatocellular carcinoma. J Pathol 228(2): 216-29.

Zhang Y. J., Wu H. C., Shen J., Ahsan H., Tsai W. Y., Yang H. I., Wang L. Y., Chen S. Y., Chen C. J. and Santella R. M. (2007). Predicting hepatocellular carcinoma by detection of aberrant promoter methylation in serum DNA. Clin Cancer Res 13(8): 2378-84.