**UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II**

DOTTORATO DI RICERCA IN:

Scienze Chimiche XXII Ciclo

# *Transcriptional patterns and pathways characterizing HCV-associated neoplastic lesions*

**Dottoranda**: Valeria De Giorgi

**Tutore**: Prof. Piero Pucci

**Co-Tutore**: Dott. Franco Buonaguro

**Relatore**: Prof. Paola Giardina

**Coordinatore**: Prof. Aldo Vitagliano

“Non è la più forte delle specie che sopravvive, né la più intelligente, ma quella più reattiva ai cambiamenti”.

*Charles Darwin*

INTRODUCTION....6
Omics and Cancer....6
The Hepatitis C virus....9
Pathogenesis of Human hepatocellular carcinoma....10
The Microarray....12
Microarray applied to Cancer Biology....14
Data Analysis....15
Aim of the PhD-thesis Project....18
MATERIALS AND METHODS....19
Patients....19
RNA Isolation....19
Quantification and Analysis of Total RNA....19
Probe Preparation and microarray hybridization....20
Statistical Analysis of Microarray Data....24
Time Course Analysis....24
Ingenuity Pathways Analysis....25
RESULTS....26
Quality Control of RNA....26
Unsupervised Analysis....27
Differential Gene Expression Patterns between HCV positive Liver Tissue with and without HCC and Normal control Liver Tissue in the first analysis....29
Differential Gene Expression Patterns between HCV positive Liver Tissue with and without HCC Metastasis and Normal control Liver Tissue in the second analysis....30
Gene signature Involved in HCC progression....37
Ingenuity Pathway Analysis....39
DISCUSSION....41
REFERENCES....46

## SUMMARY

The use of gene expression microarrays is particularly important in cancer. This is because the accumulation and combinatorial effects of abnormalities that drive the initiation and malignant progression of cancer result from the altered sequence or expression level of cancer-causing genes. Biological research on HCC mainly concentrates on early detection and diagnosis, elucidation of hepatocarcinogenesis by varieties of etiological factors, and prognosis prediction. Investigations have been conducted at different molecular levels including DNA level, RNA level and protein level, with regard to chromosomal imbalance and genetic instability, epigenetic alteration, gene expression, and gene regulation and translation. Numbers of omics-based methods have been developed and applied. Hepatitis C virus (HCV) infection is a major cause of hepatocellular carcinoma (HCC) worldwide. The precise molecular mechanism underlying the progression of chronic hepatitis viral infections to HCC is currently unknown. The direct or indirect HCV role in HCC pathogenesis is still a controversial issue and additional efforts need to be made aimed to specifically dissect the relationship between stages of HCV chronic infection and progression to HCC. The present study has been focused on investigating the genes/protein and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients, to elucidate the molecular mechanisms underlying cancer progression and to identify possible marker for diagnostic purposes trough DNA microarray.

In a first approach a pair of liver biopsies from fourteen HCV-positives HCC patients and seven HCV-negative non-liver cancer control patients (during laparoscopic cholecystectomy) were obtained, to investigate genes and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients. In a second approach to verify the consistency of the previous data obtained in a very limited sample and to identify a set of genes sufficient for the molecular signature of liver diseases, a pair of liver biopsies from twenty HCV-positive HCC patients, fifteen metastatic patients and six HCV-negative non-liver cancer control patients were collected. Gene expression profiling of liver tissues has been performed using a high-density microarray containing 36'000 oligos, representing 90% of the human genes.

Transcriptional profiles identified in liver biopsies from HCC nodules and paired non-adjacent non HCC liver tissue of the same HCV-positive patients and from metastatic patients were compared to those from HCV-negative controls by the Cluster program. The pathway analysis was performed using the BRB-Array-Tools based on the "Ingenuity System Database". Significance threshold of *t*-test was set at 0.001.

The top canonical pathways in HCV-related HCC samples include protein ubiquitination (*p*=1.67E-05), antigen presentation (*p*=9.52E-04) and Aryl Hydrocarbon receptor signaling pathway (*p*=1.37E-03). The top canonical pathways in HCV-related non HCC samples include Interferon Signaling Genes(*p*=1.12E-05), SAPK/JNK Signaling (*p*=1.07E-03) and NF-kB Activation by viruses pathway (*p*=1.19E-03). The top canonical pathways in metastatic samples include Integrin Signaling (*p=7.75E-04*) and Actin Cytoskeleton Signaling Pathway (*p*=4.43E-04)

In addition a time course analysis was performed to identify markers of tumoral progression between normal liver samples, HCV-related non HCC and HCV-related HCC liver samples. Several molecular markers for early HCC diagnosis have been recognized. In this study, informative data on the global gene expression pattern in HCV-related HCC as well as HCV-related non-HCC counterpart liver tissues have been obtained compared to normal controls. A traditional HCC diagnosis has relied on the use of a single biomarker approach (e.g., AFP). The use of multiple markers in combination to improve the accuracy of identifying HCC cases has been proposed.

All these data altogether suggested developing a specific gene-chip along with genes showing the highest fold up-regulation in common in two group of analysed samples representing the different stage of disease. The identification of the lesions and the evaluation of their neoplastic progression will be based on the gene pattern expression on the gene-chip.

# INTRODUCTION

## Omics and Cancer

Systems biology is an emergent field that aims at system-level understanding of biological systems. With the progress of genome sequence project and range of other molecular biology project that accumulate in-depth knowledge of molecular nature of biological system, we are now at the stage to seriously look into possibility of system-level understanding solidly grounded on molecular-level understanding. Systems biology focuses on systems that are composed of molecular components.

There are numbers of exciting and profound issues that are actively investigated, such as robustness of biological systems, network structures and dynamics, and applications to drug discovery.

Cancer arises when cells escape normal cell cycle and differentiations, being able to proliferate rapidly, to invade surrounding tissues, and to metastasize to distant sites. The development of cancer is proposed to be a multi-step process in which several genes and other environmental and hormonal factors play a role (Rui X.*et al* 2004).

The use of gene expression microarrays is particularly important in cancer. This is because the accumulation and combinatorial effects of abnormalities that drive the initiation and malignant progression of cancer result from the altered sequence or expression level of cancer-causing genes. These genetic abnormalities, which may be inherited or acquired, lead to the 'big six' hallmark traits of cancer, namely: (a) self-sufficiency in proliferative growth signals; (b) insensitivity to growth inhibitory signals; (c) evasion of apoptosis; (d) acquisition of limitless replicative potential; (e) induction of angiogenesis; and (f) induction of invasion and metastasis (Figure 1) (Brent R 2000, Hanahan D *et al* 2000).

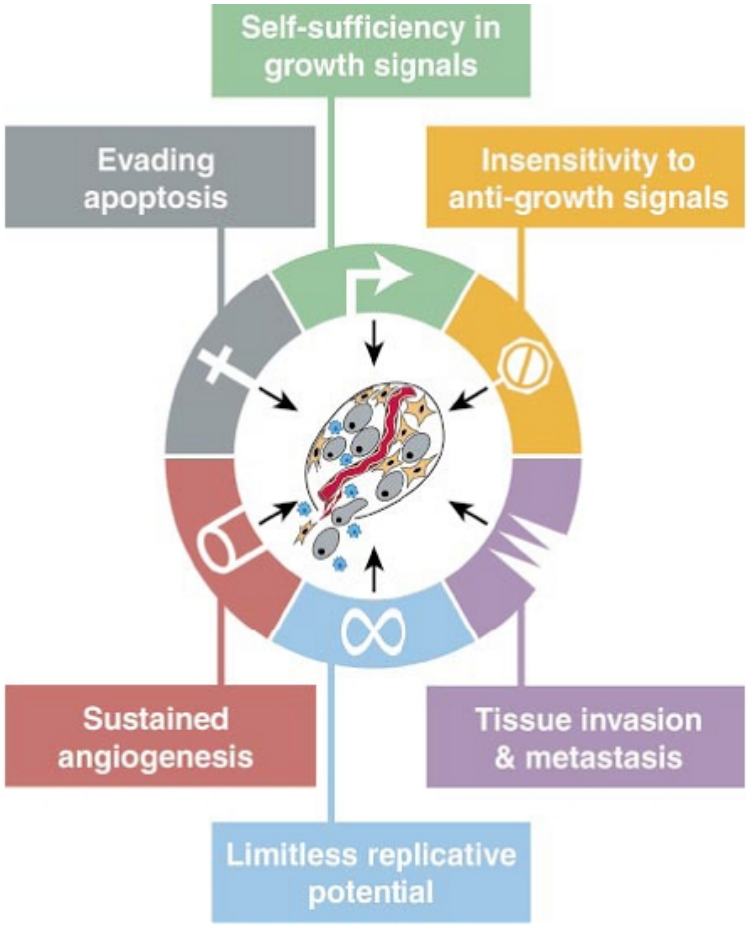


**Figure 1** *Acquired Capabilities of Cancer*

The use of microarrays can be extremely valuable both in understanding the basic biology and in the treatment of cancer. Important applications include: development of a more global understanding of the gene expression changes that contribute to malignant progression; discovery of diagnostic and prognostic indicators and biomarkers of response; identification and validation of new molecular targets; provision of an improved understanding of the molecular mode of action during lead identification and optimization; prediction of potential side-effects during preclinical development and toxicology studies; confirmation of the molecular mode of action during hypothesis-testing early clinical trials; identification of genes involved in conferring drug sensitivity and resistance; and prediction of patients most likely to benefit from the drug and use in general pharmacogenomic studies.

Gene expression can be assessed by measuring the quantity of the final product, i.e. the protein, or its intermediate, the mRNA template. Changes in the molecular

phenotype of the cell should be accurately reflected by its transcriptional profile, and the evaluation of gene expression by measuring mRNA should provide a molecular signature of the state of the activity of the cell and by extension the activity of proteins that regulate that state. However, although analysis of global mRNA expression (the transcriptome) undoubtedly generates much valuable information, the ideal scenario would be one where this is measured alongside global protein expression (the proteome) within the same experiment (Lockhart DJ *et a*l 2000).

Biological research on HCC mainly concentrates on early detection and diagnosis, elucidation of hepatocarcinogenesis by varieties of etiological factors, and prognosis prediction. Investigations have been conducted at different molecular levels including DNA level, RNA level and protein level, with regard to chromosomal imbalance and genetic instability, epigenetic alteration, gene expression, and gene regulation and translation (Lau SH *et al* 2005, Midorikawa Y et l 2007, Jiang J *et al* 2006, Herath NI *et al* 2008). Numbers of omics-based methods have been developed and applied. Large scale profiling technologies, including comparative genomic hybridization (CGH), array-based CGH, microarray and 2D electrophoresis (2DE), mass spectrometry (MS) and other proteomic analysis methods, have been used to detect change of different molecular levels (Feitelson MA *et al* 2002, Thorgeirsson SS *et al* 2006, Sun S *et al* 2007), and computational methods began to play important roles (He X *et al* 2006, Poon TC *et al* 2006, Oh JH *et al* 2008, Xu XQ *et al* 2004, Zhang X *et al* 2006). A variety of HCC-associated molecular alterations have been detected. However, because of the lack of good diagnostic markers and treatment strategies and because of clinical heterogeneity, a coherent understanding of the mechanism of HCC development is still limited (Blum HE *et al* 2005). The assessment of complex multigenic molecular pathways in HCC remains a difficult challenge. Integrating observations from multiple aspects is an essential step toward the systematic understanding of the disease.

**The Hepatitis C virus**

Hepatitis C virus is a member of the Flaviviridae family of enveloped, positive-strand RNA viruses and is the only member of the genus Hepacivirus. The HCV genome consists of an RNA molecule, of approximately 9.6 kb, that contains a large open-reading frame flanked by structured 50 and 30 no-translated regions (NTRs). Viral proteins are translated as a polyprotein precursor from an internal ribosome entry site (IRES) located in the 50 NTR. The polyprotein undergoes a complex series of co- and posttranslational cleavage events catalysed by both host and viral proteinases to yield the individual HCV proteins. The structural proteins include the core protein and the envelope glycoproteins E1 and E2. The non-structural proteins include the P7 polypeptide, the NS2-3 autoprotease and the NS3 serine protease, an RNA helicase located in the C-terminal region of NS3, the NS4A polypeptide, the NS4B and NS5A proteins, and the NS5B RNA-dependent RNA polymerase (Tellinghuisen and Rice, 2002) (Figure 2)

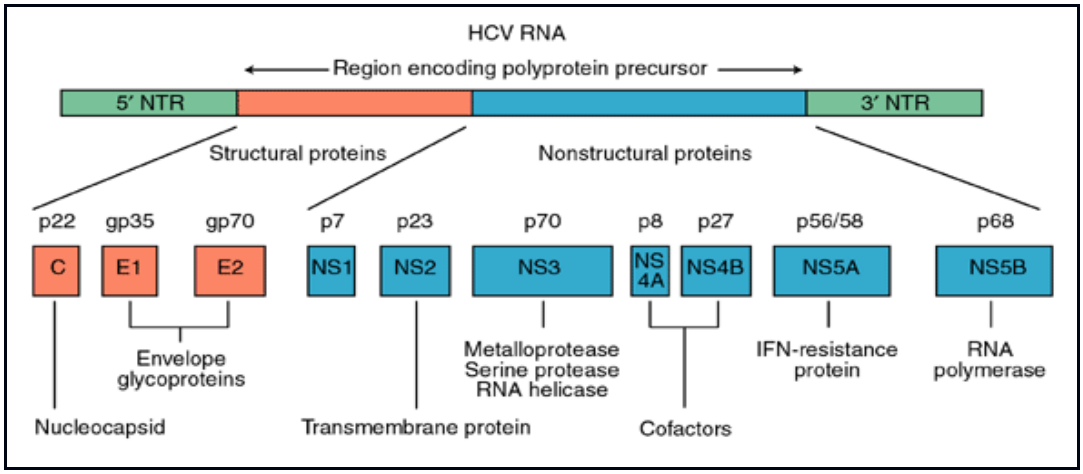


**Figure 2** *Hepatitis C virus (HCV) proteins*

The course of disease varies widely among infected persons. HCV infection is hardly ever diagnosed during the acute phase. Progression to chronic disease occurs in about 70-80% of infected persons, whereas 20-30% show spontaneous recovery. The early stage of chronic infection is typically characterized by a prolonged asymptomatic period. Spontaneous clearance of viremia, once chronic infection has been established, is rare. Most chronic infections will lead to

hepatitis and to some degree of fibrosis. 10-20% of those infected chronically develop liver cirrhosis. At this stage of the disease the risk of developing hepatocellular carcinoma (HCC) is 1-4% per year (Tsukuma *et al*, 1999)

**Pathogenesis of Human hepatocellular carcinoma**

Hepatocellular carcinoma (HCC) ranks among the most common cancers in many countries (Bosch *et al*., 1999). A recent estimate indicates that HCC represents the fifth most common cancer of males, and the eighth most common cancer in female candidates, with a total of 560 000 new cases each year, 83% of which occurring in developing countries, and more than one-half in China alone. Moreover, because of its very poor prognosis, HCC represents the third leading cause of cancer death worldwide. Chronic hepatitis B and C and associated liver cirrhosis represent major risk factors for HCC development, being implicated in more than 70% of HCC cases worldwide. A large analysis of HCC in Europe, based on both serology and molecular tests, has demonstrated the major impact of hepatitis B virus (HBV) and hepatitis C virus (HCV). Only 29% of HCC cases were found negative for these viruses. The hepatitis B surface antigen (HBsAg) and anti-HCV antibodies were detected in 19 and 40.1% of the patients, respectively, with HCV 1b being the most prevalent genotype (Brechot *et al*., 1998) Additional etiological factors that often represent co-factors of an underlying HBV- or HCVrelated chronic liver disease include toxins and drugs (e.g., alcohol, aflatoxins, microcystin, anabolic steroids), metabolic liver diseases (e.g., hereditary haemochromatosis, α1-antitrypsin deficiency), steatosis (Ohata *et al*., 2003) and non-alcoholic fatty liver diseases (Brunt *et al*., 2004), diabetes (Davila *et al*., 2005). In general, HCCs are more frequent in men than in women and the incidence increases with age. The prevalence of HCC in Italy, and in Southern Italy in particular, is significantly higher compared to other Western countries. Hepatitis virus infection, long-term alcohol and tobacco consumption account for 87% of HCC cases in Italian population and, among these, 61% of HCC are attributable to HCV. In particular, a recent seroprevalence surveillance study conducted in the general population of Southern Italy Campania Region reported 7.5% positivity for HCV infection which peaked at 23.2% positivity in the 65 years or older age group (Fusco M *et al* 2008).

As for most types of cancer, hepatocarcinogenesis is a multistep process involving different genetic alterations that ultimately lead to malignant transformation of the hepatocyte. Malignant transformation of hepatocytes is believed to occur, regardless of the etiological agent, through a pathway of increased liver cell turnover, induced by chronic liver injury and regeneration, in a context of inflammation and oxidative DNA damage. This microenvironment facilitates the occurrence of genetic and epigenetic alterations. Chronic viral hepatitis, alcohol, metabolic liver diseases such as hemochromatosis and α1-antitrypsin deficiency, as well as non-alcoholic fatty liver disease may act predominantly through this pathway of chronic liver injury, regeneration and cirrhosis. Accordingly, the major clinical risk factor for HCC development is liver cirrhosis and 70–90% of all HCCs develop in a cirrhotic liver. The risk of HCC in patients with liver cirrhosis depends on the activity, duration and the etiology of the underlying liver disease. The co-existence of multiple etiologies, for example, HCV infection with overt or occult HBV, aflatoxin B1(AFB1) and HBV infection, HCV infection and alcohol or HCV infection and liver steatosis, increases the relative risk of HCC development.

**The Microarray**

DNA microarrays are created by robotic machines that arrange minuscule amounts of hundreds or thousands of gene sequences on a single microscope slide. There is a database of over 40,000 gene sequences that its possible to use for this purpose. When a gene is activated, cellular machinery begins to copy certain segments of that gene. The resulting product is known as messenger RNA (mRNA), which is the body's template for creating proteins. The mRNA produced by the cell is complementary, and therefore will bind to the original portion of the DNA strand from which it was copied. To determine which genes are turned on and which are turned off in a given cell, a researcher must first collect the messenger RNA molecules present in that cell. Then each mRNA molecule is labelled by using a reverse transcriptase enzyme (RT) that generates a complementary cDNA to the mRNA. During that process fluorescent nucleotides are attached to the cDNA. The tumour and the normal samples are labelled with different fluorescent dyes. Next, the labelled cDNAs is placed onto a DNA microarray slide. The labelled cDNAs that represent mRNAs in the cell will then hybridize or bind to their synthetic complementary DNAs attached on the microarray slide, leaving its fluorescent tag. A special scanner is used to measure the fluorescent intensity for each spot/areas on the microarray slide. If a particular gene is very active, it produces many molecules of messenger RNA, thus, more labelled cDNAs, which hybridize to the DNA on the microarray slide and generate a very bright fluorescent area. Genes that are somewhat less active produce fewer mRNAs, thus, less labelled cDNAs, which results in dimmer fluorescent spots. If there is no fluorescence, none of the messenger molecules have hybridized to the DNA, indicating that the gene is inactive. Frequently use this technique is used to examine the activity of various genes at different times. When co-hybridizing Tumour samples (Red Dye) and Normal sample (Green dye) together, they will compete for the synthetic complementary DNAs on the microarray slide. As a result, if the spot is red, this means that specific gene is more expressed in tumour than in normal (up-regulated in cancer). If a spot is green, that means that gene is more expressed in the Normal tissue (down-regulated in cancer). If a spot is yellow that means that specific gene is equally

expressed in normal and tumour (Figure3). The log ratio between the two intensities of each dye is used as the gene expression data *gene expression = log 2 Int(Cy5)/ Int(Cy3)* (Lashkari *et al*. 1997, Derisi *et al*. 1997, Eisen *et al*. 1998).

Microarray image processing uses differential excitation and emission wavelengths of the two fluors to obtain a scan of the array for each emission wavelength, typically as two 16-bit grey scale TIFF images. These images are then analysed to identify the spots, calculate their associated signal intensities, and assess local background noise. Most image acquisition software packages also contain basic filtering tools to flag spots such as extremely low intensity spots, ghosts spots (where background is higher than spot intensity), or damaged spots (e.g., dust artifacts). These results allow an initial ratio of the evaluated channel/reference channel intensity to be calculated for every spot on the chip. The products of the image acquisition are the TIFF image pairing and a quantified data file that has not yet been normalized.

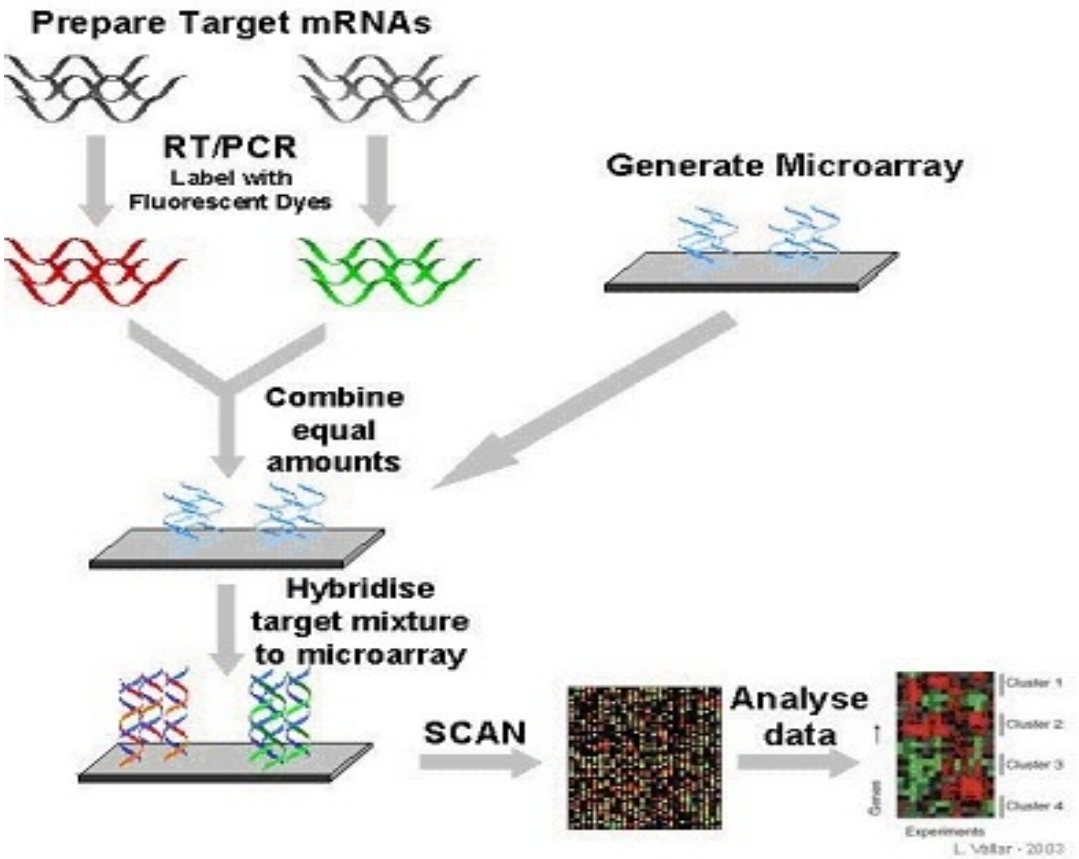


**Figure 3** *Microarray scheme*

**Microarray applied to Cancer Biology**

Cancer is caused by the accumulation of genetic and epigenetic changes resulting from the altered sequence or expression of cancer-related genes, such as oncogenes or tumour suppressor genes, as well as genes involved in cell cycle control, apoptosis, adhesion, DNA repair, and angiogenesis. Because gene expression profiles provide a snapshot of cell functions and processes at the time of sample preparation, comprehensive combinatorial analysis of the gene expression patterns of thousands of genes in tumour cells and comparison to the expression profile obtained with healthy cells should provide insights concerning consistent changes in gene expression that are associated with tumour cellular dysfunction and any concomitant regulatory pathways.

The new bioinformatics tools and development of genome-wide microarray analyses both in human, mice and other model organisms have opened new windows in cancer research. The field of gene expression studies has greatly advanced the identification of novel tumour susceptibility genes, classification of tumours, prediction of outcome, treatment response, discovery of potential markers and targets for diagnosis of this malignant disease. Microarray technology and the statistical tools developed for it are an excellent option to study mRNA expression differences of normal and tumour tissues or various tumours and model organisms on a global scale. Class discovery methods such as hierarchical or K-means clustering or self-organizing maps (MacQueen J. 1967, Kohonen T. 2001) provide a global overview of the cell transcript levels and can be very useful in identifying novel markers for cancer or to identify important genes or pathways for tumorigenesis. Class prediction methods, on the other hand, provide detailed information of specific genetic signatures in various tumour subtypes, which may previously have been very difficult to characterize by conventional methods. Additionally, expression array technology can provide a tool to diagnose clinical cases which may have been difficult to identify otherwise. DNA microarray technology has expanded rapidly and has been applied to study several different types of human cancer, such as breast, (Perou CM *et al* 1999,Sorlie *et al* 2001, van't Veer LJ *et al* 2002, Hedenfalk *et al* 2001, Cunliffe HE *et al* 2003) prostate, (Bubendorf L *et al* 1999, Dhanasekaran SM *et*

*al* , 2001, 412, Singh D et 2002) colorectal (Alon U *et al* 1999) and ovarian cancer (Ono K *et al* 2000) as well as hematological malignancies (Golub TR *et al* 1999, Alizadeh A, *et al* 1999, Alizadeh AA *et al* 2000, Golub TR. *et al* 2001).

Yet another study of importance is the one by van't Veer *et al.* (2002), where they were able to classify lymph node negative breast tumours into those with poor or good prognosis using a signature of 70 genes, with power that outstripped the available clinical prognostic markers. The impact of classification can be clinically very significant, as in prostate cancer (PC) where the surgical removal and risk that the surgery poses to these often older men has to be assessed. With molecular signatures being able to determine the possible outcome of the cancer these decisions can be greatly facilitated.

**Data Analysis**

Microarrays can be used to investigate problems in cell biology in various ways, with a range of differential approaches. At the other end, the main interest lies in finding a single change in gene expression that might be a key to a given alteration in phenotype. At the other extreme, the aim is to look at overall patterns of gene expression in order to understand the architecture of genetic regulatory networks. The basic idea behind the statistical analysis is to characterize the structure of the experimental data and extract statistically significant patterns from it. Because of the complexity of the problems and datasets generated by microarray experiments, the use of data analysis software is essential. A large number of commercial and non-commercial software tools for statistical analysis and visualization of gene expression data have been developed, which all offer their own solutions to the problem at hand. GeneSpring (Agilent Technologies), Cluster and Treeview (Eisen MB *et al* 1998, Tamayo *et al* 1999, Golub TR *et al* 1999), SAM (Tusher VG et al, 2001), and dCHIP (Wong WH *et al* 2001), are examples of these software tools, to name a few. Methods utilized in the data analysis vary considerably. The analysis of microarray data is explorative by nature, and the components of the analysis depend upon the purpose of the experiment. Tools that are generally used include filters to remove redundant genes from the experiment, statistical tests to find differentially expressed genes and classification methods to discover pathway level expression patterns and find

distinguished expression profile signatures. The expression of a large number of genes that are irrelevant or unchanged add a high level of noise and uncertainty to the data, which makes the use of classical statistical tests problematic. Reducing the number of genes in the analyses benefits the power of statistical testing. Gene filtering and differentiation approaches can efficiently reduce the dimensionality of the data and help remove redundant genes. This helps to highlight the genes that are truly differential for the investigated trait. Typical tests used in microarray analysis are parametric tests such as Student's t-test or analysis of variance (ANOVA), which assume normal distribution of the data and try to estimate whether the variance in the data comes from the normal distribution or not. In addition, microarray experiments typically have large number of observations but only few samples that leads to testing of multiple hypotheses. As some of the observed differences are expected to happen by chance alone, correction for multiple testing is desired. These adjustments to the statistical tests include corrections such as the Bonferroni method and the false discovery rate (FDR) suggested by Benjamini and Hochberg (Benjamini *et al* 1995). Permutation-based models are another approach to validate the results. Other methods for analysis include data transformations such as principal component analysis (PCA) and singular value decomposition (SVD), which reduce the dimensionality of the data and aim to find the major components explaining the variance in the data. Fold change was among the first methods used to evaluate whether genes were differentially expressed, but is nowadays considered an inadequate test statistic when used alone, as it does not account for the variance and offers no associated level of confidence. The choice of an appropriate correction can be challenging, as many of the popular correction methods, such as the Bonferroni method, have not been designed for microarray data, where there are few cases but many observations per sample. This may lead to very stringent correction and loss of data, with no false positive findings, but also very few true positive findings. Therefore, permutations and FDR based methods with adjustable threshold levels have gained popularity in validation of microarray analyses.

Classification is a widely used analysis method for gene expression data, used either to discover new categories within a dataset or to assign cases to a given

category, and is often referred to as clustering. Two principal categories of clustering exist: the unsupervised and supervised methods. In the unsupervised clustering (or class discovery), objects such as genes or samples are grouped into classes based on some sort of similarity metric that is computed for one or more variables. Typically, genes are grouped into classes on the basis of the similarity in their expression profiles across cases, tissues or conditions. Unsupervised clustering can further be split to hierarchical clustering methods, which produce a tree diagram (dendrogram) and non-hierarchical clustering methods such as self-organizing maps (SOM) or K-means clustering, (MacQueen J *et al* 1967, Kohonen T 2001) which typically divide the cases into a predetermined number of groups in a manner that maximizes a specific function. In the supervised clustering (or class prediction) methods, algorithms are developed to assign objects to predetermined categories. The supervised methods generally involve the use of a training data set and an independent validation data set, and aim to obtain a function or rule that uses expression data to predict whether a case is of one type or another.

Biological interpretation of transcriptomes from human liver tissues containing multiple cell types, each type with its own expression program, is notoriously difficult. With gene expression data from human liver biopsies as input, was used IPA to confidently reconstruct and integrate cellular regulatory networks and canonical pathways. Ingenuity Pathways Analysis (IPA) is a software/database search tool for finding function and pathway for specific biological states. It is web-delivered application makes use of the Ingenuity Pathways Knowledge Base, the curated database consisting of millions of individually modelled relationships between proteins, genes, complexes, cells, tissues, drugs, and diseases. It currently has data for human, rat, and mouse.

**Aim of the PhD-thesis Project**

The precise molecular mechanism underlying the progression of chronic hepatitis viral infections to HCC is currently unknown. The direct or indirect HCV role in HCC pathogenesis is still a controversial issue and additional efforts need to be made aimed to specifically dissect the relationship between stages of HCV chronic infection and progression to HCC

The present study has been focused on investigating the genes/protein and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients, to elucidate the molecular mechanisms underlying cancer progression and to identify possible marker for diagnostic purposes trough DNA microarray.

# MATERIALS AND METHODS

## Patients

In a first approach liver biopsies from fourteen HCV-positive HCC patients and seven HCV-negative non-liver cancer control patients (during laparoscopic cholecystectomy) were obtained with informed consent at the liver unit of the INT "Pascale" in Naples, to investigate genes and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients.

In a second approach to verify the consistency of the previous data obtained in a very limited sample and to identify a set of genes sufficient for the molecular signature of liver diseases, liver biopsies from twenty HCV-positive HCC patients, fifteen metastatic patients and six HCV-negative non-liver cancer control patients were collected. In particular, from each of the HCV-positive HCC patients, and metastatic samples a pair of liver biopsies from HCC nodule and non-adjacent non-HCC counterpart were surgically excised.

All liver biopsies were stored in RNA Later at -80°C (Ambion, Austin, TX). Confirmation of the histopathological nature of the biopsies was performed by the Pathology lab at INT before the processing for RNA extraction. The non-HCC tissues from HCV-positive patient were an heterogeneous sample representing the prevalent liver condition of each subject (ranging from persistent HCV-infection to cirrhotic lesions). Furthermore, laboratory analysis confirmed that the 11 controls were seronegative for hepatitis C virus antibodies (HCV Ab).

## RNA Isolation

Samples were homogenized in disposable tissue grinders (Kendall, Precision). Total RNA was extracted by TRIzol solution (Life Technologies, Rockville, MD) according to manufacturer's instructions. RNA was aliquoted and stored at -80°C.

## Quantification and Analysis of Total RNA

The purity of the RNA preparation was verified by the 260:280 nm ratio (range, 1.8-2.0) at spectrophotometric reading with NanoDrop (Thermo Fisher Scientific, Waltham, MA). Integrity of extracted RNA was evaluated by Agilent 2100

Bioanalyzer (Agilent Technologies, Palo Alto, CA), analysing the presence of 28S and 18S ribosomal RNA bands as well as the 28S/18S rRNA intensity ratio equal or close to 1.5 (Figure 4). In addition, phenol contamination was checked and a 260:230 nm ratio (range, 2.0-2.2) was considered acceptable.

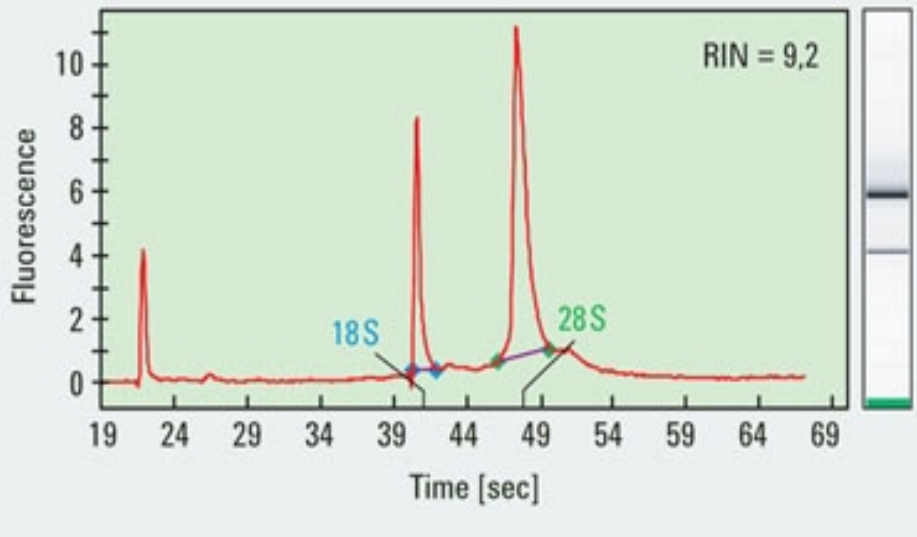


**Figure 4** *Representative Electropherogram of total RNA*

**Probe Preparation and microarray hybridization**

Double-stranded cDNA was prepared from 3 µg of total RNA (T-RNA) in 9 µl DEPC -treated H2O using the Super script II Kit (Invitrogen) with a (dT) (5' AAA CGA CGG CCA GTG AAT TGT AAT ACG ACT CAC TAT AGG CGC $T_{(15)}$ oligonucleotide primer. cDNA synthesis was completed at 42°C for 1 h (Table 1)

| *Component* | *Volume(µl)* |
|---|---|
| 5 X First strand buffer | 4ul |
| 0.1- 0.5ug/ul TS | 1ul |
| 0.1M DTT | 2ul |
| 10mM dNTP | 2ul |
| Superscript II | 1ul |
| | |
| *Time (min)* | *Temperature (°C)* |
| 90 | 42 |

**Table 1** *First strand cDNA synthesis*

Full-length dsDNA was synthesized incubating the produced cDNA with 2 U of RNase-H (Promega) and 3 µl of Advantage cDNA Polymerase Mix (Clontech), in Advantage PCR buffer (Clontech), in presence of 10 mM dNTP and DNase-free water (Table 2).

| *Component* | *Volume (µl)* |
|---|---|
| DEPC treated H2O | 106 ul |
| Advantage PCR buffer | 15ul |
| 10 mM dNTP mix | 3ul |
| RNase H | 1ul |
| Advantage cDNA Polymerase mix | 3ul |
| | |
| *Time (min)* | *Temperature (°C)* |
| 5 | 37 |
| 2 | 94 |
| 1 | 65 |
| 30 | 75 |

**Table 2** *Second strand cDNA synthesis*

dsDNA was extracted with phenolchloroform- isoamyl, precipitated with ethanol in the presence of 1 µl linear acrylamide (0.1 µg/µl, Ambion, Austin, TX) and aRNA (amplified-RNA) was synthesized using Ambion's T7 MegaScript in Vitro Transcription Kit (Ambion, Austin, TX) (Table 3)

| *Component* | *Volume (µl)* |
|---|---|
| 75mM NTP | 2ul |
| reaction buffer | 2ul |
| enzyme mix | 2ul |
| double stranded cDNA | 8ul |
| | |
| *Time (hours)* | *Temperatue (°C)* |
| 16 | 37 |

**Table 3** *In Vitro Tanscription*

aRNA recovery and removal of template dsDNA was achieved by TRIzol purification. For the second round of amplification, aliquots of 1 µg of the aRNA were reverse transcribed into cDNA using 1 µl of random hexamer under the conditions used in the first round. (Table 4)

| *Component* | *Volume (µl)* |
|---|---|
| 5 X First strand buffer | 4ul |
| (0.5ug/ul) oligo dT-T7 primer | 1ul |
| 0.1M DTT | 2ul |
| RNAsin | 1ul |
| 10mM dNTP | 2ul |
| Superscript II | 2ul |
| | |
| *Time (min)* | *Temperature (°C)* |
| 90 | 42 |

**Table 4** *Second round of amplification*

Second-strand cDNA synthesis was initiated by 1 µg oligo-dT-T7 primer and the resulting dsDNA was used as template for in vitro transcription of aRNA in the same experimental conditions as for the first round.

6 µg of this aRNA was used for probe preparation, in particular test samples were labelled with USL-Cy5 (Kreatech) and pooled with the same amount of reference sample (control donor peripheral blood mononuclear cells), PBMC, seronegative for hepatitis C virus antibodies (HCV Ab) labelled with USL-Cy3 (Kreatech) (Figure 5).

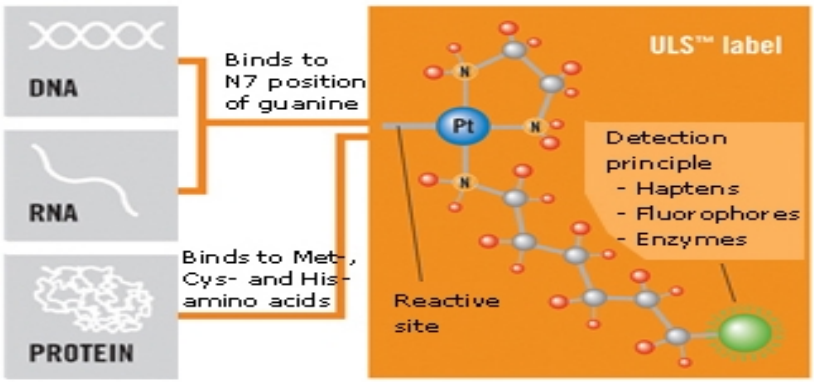


**Figure 5** *Structure of ULS_reagent*

The two labelled aRNA probes were separated from unincorporated nucleotides by filtration, fragmented, mixed and co-hybridized to a custom-made 36 K oligoarrays at 42°C for 24 h. The oligo-chips were printed at the Immunogenetics Section Department of Transfusion Medicine, Clinical Center, National Institutes of Health (Bethesda, MD) (Table 5)

| *Component* | *Volume (µl)* |
|---|---|
| 2nd amp. RNA (6 µg ) | 3 µl |
| Cy-5 | 2 µl |
| Cy-3 | 2 µl |
| 10x labeling solution | 1 µl |
| H20 | 2 µl |
| | |
| *Time (min)* | *Temperature (°C)* |
| 85 | 15 |

**Table 5** *Target labelling*

After hybridization the slides were washed with 2 × SSC/ 0.1%SDS for 1 min, 1 × SSC for 1 min, 0.2 × SSC for 1 min, 0.05 × SSC for 10 sec., and dried by centrifugation at 800 g for 3 minutes at RT. (Table 6).

| | Buffer | Volume | Time |
|---|---|---|---|
| 1st Wash | 2x SSC + 0.1% SDS | 450ml dd H2O, 50ml 20XSSC 5ml 0.1%SDS | 10 secondi |
| 2nd Wash | 1x SSC | 475ml dd H2O, 25ml 20XSSC | 1min |
| 3rd Wash | 0.2 x SSC | 495ml dd H2O, 5ml 20XSSC | 30 min |
| 4th Wash | 0.05x SSC | 500ml dd H2O 1.25ml 20XSSC | 10 secondi |

**Table 6** *Array washing*

Hybridized arrays were scanned at 10-µm resolution with a GenePix 4000 scanner (Axon Instruments) at variable photomultiplier tube (PMT) voltage to obtain maximal signal intensities with less than 1% probe saturation. Image and data files were deposited at microarray data base (mAdb) at http://madb.nci.nih.gov/

and retrieved after median centred, filtering of intensity (>200) and spot elimination (bad and no signal). Data were further analysed using Cluster and TreeView software (Stanford University, Stanford, CA).

**Statistical Analysis of Microarray Data**

For the Unsupervised analysis, a low-stringency filtering was applied, selecting the genes differentially expressed in 80% of all experiments with a >3 fold change ratio in at least one experiment. Hierarchical cluster analysis was conducted on these genes according to Eisen *et al*. (Eisen MB *et al* 1998); differential expressed genes were visualized by Treeview and displayed according to the central method (Ross DT *et al* 2000). Supervised class comparison was performed using the BRB ArrayTool developed at NCI, Biometric Research Branch, Division of Cancer Treatment and Diagnosis.

Paired samples were analysed using a two-tailed paired Student's *t*-test. Unpaired samples were tested with a two-tailed unpaired Student's *t*-test assuming unequal variance or with an *F*-test as appropriate. Gene clusters identified by the univariate *t*-test were challenged with two alternative additional tests, a univariate permutation test (PT) and a global multivariate PT. The multivariate PT was calibrated to restrict the false discovery rate to 10%. Genes identified by univariate *t*-test as differentially expressed ($p$-value < 0.001 and $p$-value < 0.01) and a PT significance <0.05 were considered truly differentially expressed. Gene function was assigned based on Database for Annotation, Visualization and Integrated Discovery (DAVID) and Gene Ontology http://www.geneontology.org/

**Time Course Analysis**

A time course analysis was performed to identify markers of tumoral progression between normal liver samples, HCV-related non HCC and HCV-related HCC liver samples. For this analysis, normal liver samples (CTR) were taken as the early time point, HCV-related non HCC the intermediate point and the HCV-related HCC the last time point. Parameters for gene selection were: *F* test $p$-value < 0.005, 80% presence call, ratio of > 2 and false discovery rate < 0.1.

**Ingenuity Pathways Analysis**

The pathway analysis was performed using the gene set expression comparison kit implemented in BRB-Array- Tools. The human pathway lists determined by "Ingenuity System Database" was selected. Significance threshold of *t*-test was set at 0.001. The Ingenuity Pathways Analysis (IPA) is a system that transforms large data sets into a group of relevant networks containing direct and indirect relationships between genes based on known interactions in the literature.

# RESULTS

## Quality Control of RNA

The quality of extracted total RNA was verified by Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA), showing discrete 28S and 18S rRNA bands (Figure 6A) as well as a 28S/18S rRNA intensity ratio equal or close to 1.5 which is considered appropriate for total RNA extracted from liver tissue biopsies ("Assessing RNA Quality", http://www.ambion.com/techlib/tn/111/8.html). Based on this parameter, all extracted total RNA samples met the quality control criteria (Figure 6B).

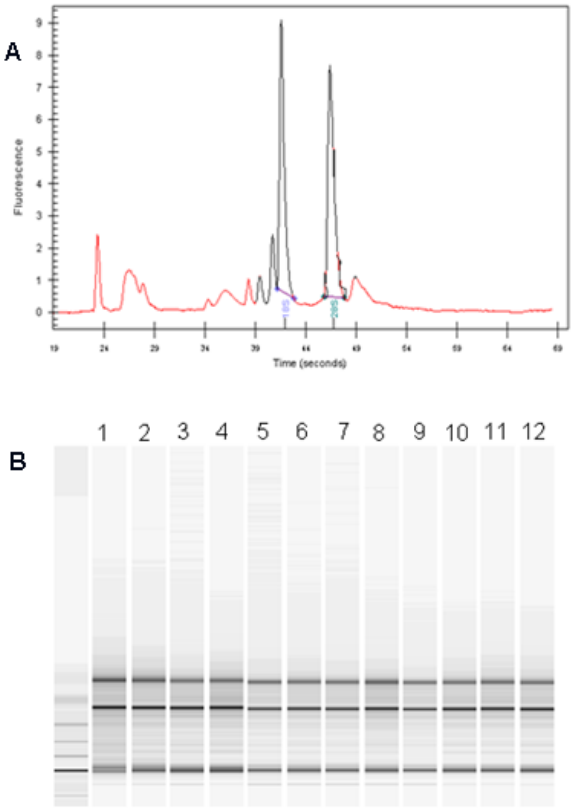


**Figure 6** *(A) Electropherogram of total RNA extracted from samples included in the analysis. (B) Gel image evaluation of RNA integrity and 28S/18S rRNA ratio.*

**Unsupervised Analysis**

The gene expression profiles of tissue samples from the three groups of analysed samples (the HCV-related HCC, their non-HCC counterpart, as well as samples from control patients) were compared by an unsupervised analysis. No clear separation of the 3 different groups was observed, although control samples clustered mainly with samples from HCV-related non-HCC paired tissue, which includes dysplastic lesion in cirrhotic liver, representing a pre-neoplastic step (Figure 7A).

In order to identify genes differentially modulated in HCV-related lesions compared to normal liver tissue samples, an unsupervised analysis was then performed including only paired samples from HCV-related HCC and normal control samples and from the HCV-related non- HCC counterpart and control samples (Figures 7B and 7C). According to filtering described in Material and Methods, HCV-related HCC and normal control samples showed 5'473 genes differentially expressed, with a perfect clustering according to histological characteristics (Figure 6B). Similarly, HCV-related non-HCC tissue and normal control samples showed 6'069 genes differentially expressed with a perfect clustering according to histological characteristics also in this case (Figure 6C). The only exception to this pattern is represented by the normal control sample (CTR#80) which did not fall in the control cluster (CTR).

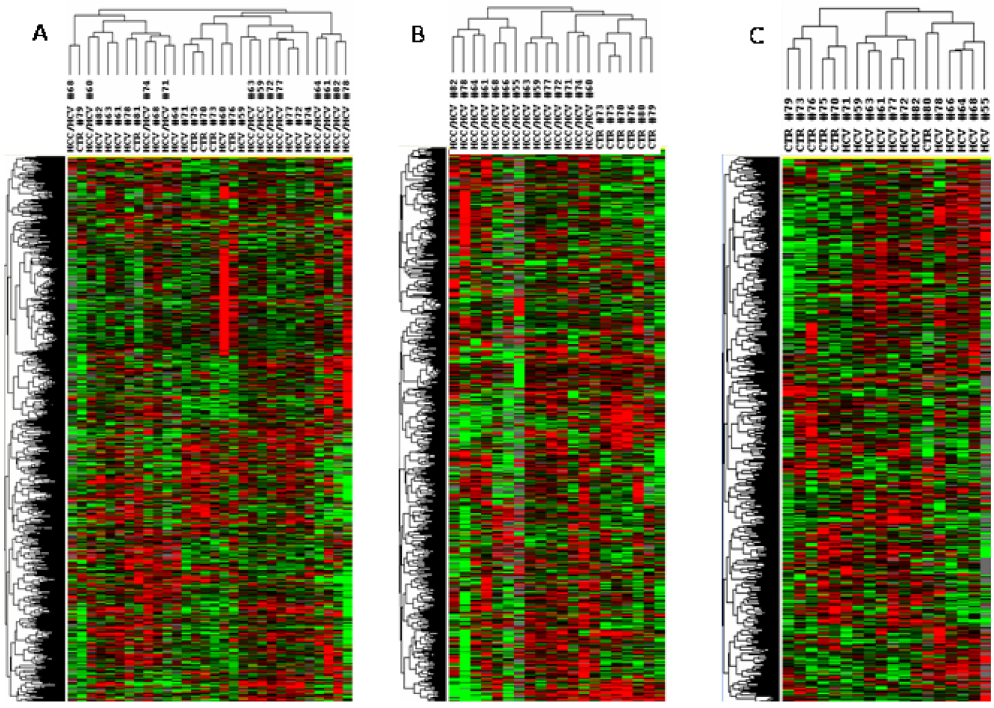


**Figure 7** *Unsupervised hierarchical clustering. Red indicates over-expression; green indicates under-expression; black indicates unchanged expression; grey indicates no detection of expression (intensity of both Cy3 and Cy5 below the cut-off value). Each row represents a single gene; each column represents a single sample. The dendrogram at the left of matrix indicates the degree of similarity among the genes examined by expression patterns. The dendrogram at the top of the matrix indicates the degree of similarity between samples. (A), unsupervised analysis including all three set of samples; (B), unsupervised analysis including HCV-related HCC and normal control liver samples; (C), unsupervised analysis including HCV-related non-HCC counterpart and normal control liver samples.*

The gene expression profiles of tissue samples from the second groups of analysed samples (the HCV-related HCC, their non-HCC counterpart, metastasis, as well as samples from control patients) were compared by an unsupervised analysis. 5210 genes were found to be significantly differentially expressed among the four diagnostic groups. The HCC samples and the metastatic cancers clustered into two distinct groups, based on differences in their patterns of gene expression (Figure 8).

The expression patterns of gastrointestinal liver metastasis are clearly distinct from those of HCV-related primary HCC, allowing a definite molecular characterization of the two liver diseases.

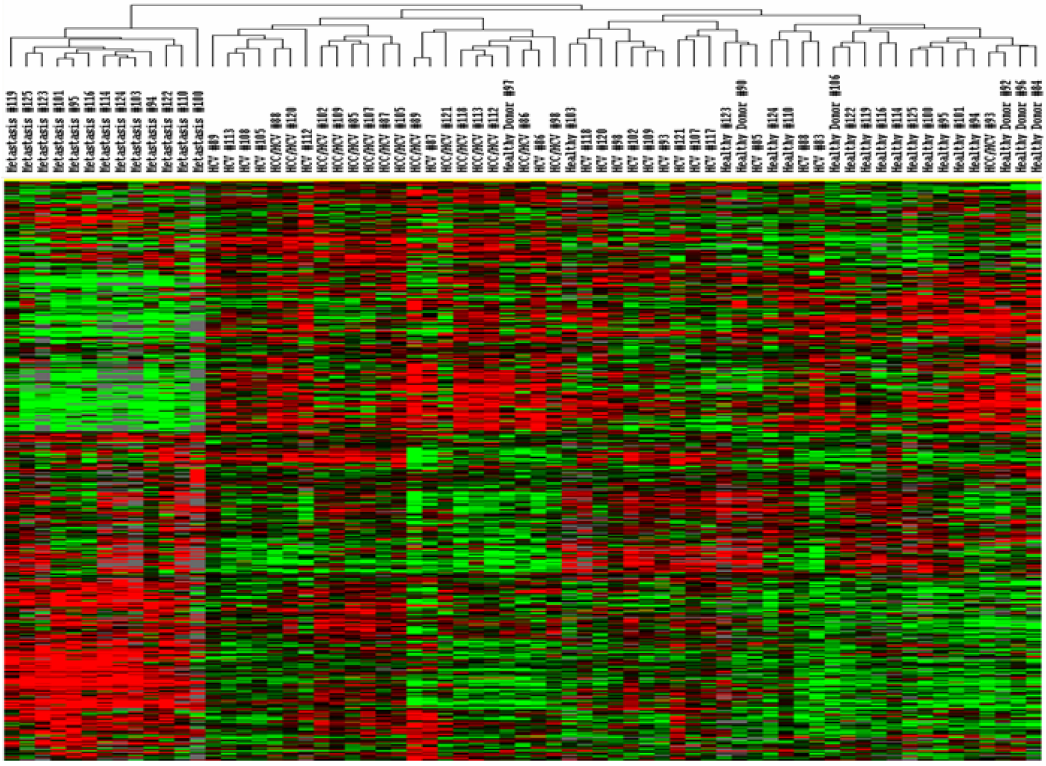

**Figure 8** *Unsupervised hierarchical clustering. Overall patterns of gene expression across the 20 HCV-related HCC and non-HCC counterparts, 15 liver metastasis and their non metastatic counterparts as well as 6 HCV-negative control patients.*

### Differential Gene Expression Patterns between HCV positive Liver Tissue with and without HCC and Normal control Liver Tissue in the first analysis

The supervised analysis was performed comparing pairs of gene sets using an unpaired Student's *t* test with a cut-off set at $p < 0.01$. The analysis comparing gene sets in liver tissues from HCV related HCC and normal controls identified 825 genes differentially expressed. Among them, 465 were shown to be up-regulated and 360 down-regulated in HCV-related HCC liver tissues (Figure 9A) The analysis comparing gene sets in liver tissues from HCV-related non-HCC tissue and controls identified 151 genes differentially expressed. Among them, 127 were shown to be up-regulated and 24 down-regulated in HCV related non-HCC liver tissues (Figure 9B)

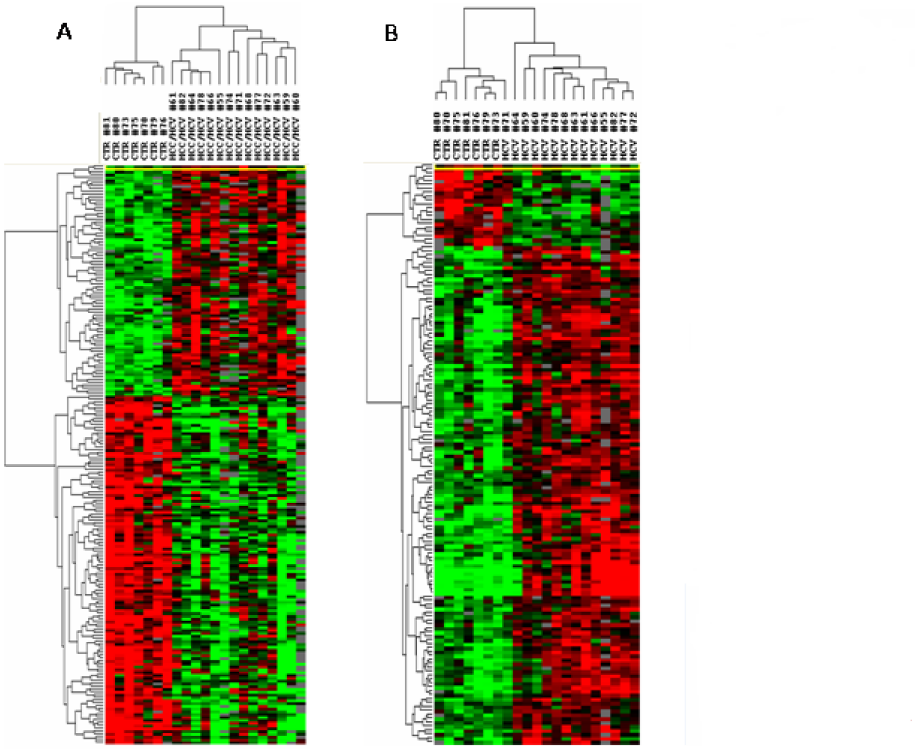


**Figure 9** *Heat map of the genes differentially expressed, identified by Class Comparison Analysis (A), analysis including HCV-related HCC and normal control liver samples; (B), analysis including HCV-related non-HCC liver tissues and control liver samples. The expression pattern of the genes is shown, each row represents a single gene.*

### Differential Gene Expression Patterns between HCV positive Liver Tissue with and without HCC Metastasis and Normal control Liver Tissue in the second analysis

The supervised analysis was performed comparing pairs of gene sets using an unpaired Student's *t*-test with a cut-off set $p<0.05$. The analysis comparing gene sets in liver tissues from HCV-related HCC and normal controls identifies 1182 genes differentially expressed. Among them, 704 were shown to be up-regulated and 478 down regulated in HCV-related HCC liver tissues (Figure 10A). The analysis comparing gene sets in liver tissues from HCV-related non HCC and normal controls identify 875 genes differentially expressed. Among them, 540 were shown to be up-regulated and 335 down regulated in HCV-related non HCC liver tissues (Figure 10B).

The analysis comparing gene sets in liver tissues from metastasis and their paired non metastatic normal counterpart identify 2794 genes differentially expressed.

Among them, 1919 were shown to be up-regulated and 647 down-regulated in metastatic liver tissues (Figure 10C)

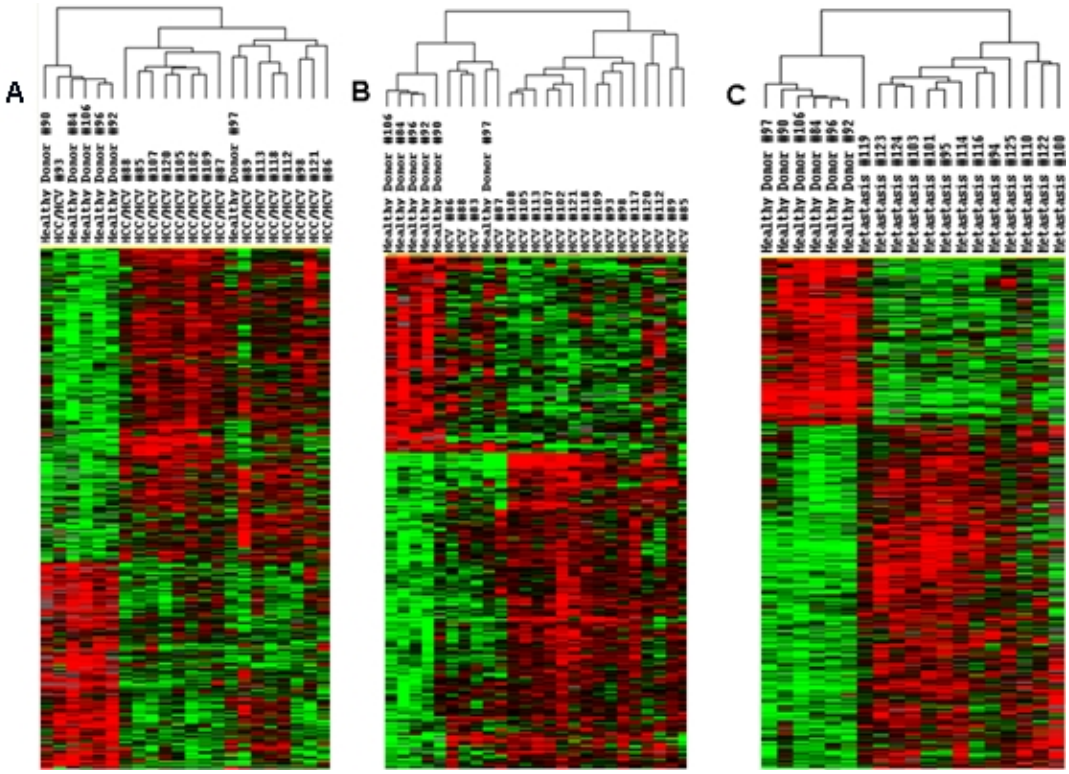


**Figure 10** *Heat map of the genes differentially expressed, identified by Class Comparison Analysis. (A) analysis including HCV-related HCC and normal control liver samples; (B), analysis including HCV-related non-HCC liver tissues and normal control liver samples; (C), analysis including liver metastasis and their paired non metastatic normal counterpart. The expression pattern of the genes is shown each row representing a single gene.*

The genes showing the highest fold up-regulation in HCV related HCC, in HCV related non HCC in common in two works and in metastatic liver samples are listed respectively in Tables 7, 8, 9.

| NAME | DESCRIPTION | "ENTREZ GENE" |
|---|---|---|
| ADPRHL2 | ADP-ribosylhydrolase like 2 | 54936 |
| ARID4B | AT rich interactive domain 4B transcript variant 1 | 51742 |
| ARL8A | ribosylation factor-like 8A | 127829 |
| ATG3 | autophagy related 3 homolog | 64422 |
| ATP5G1 | synthase, H+ transporting, mitochondrial F0 complex, subunit C1 (subunit 9) | 516 |
| AXIN1 | axin 1, transcript variant 2 | 8312 |
| BOLA2 | bolA homolog 2 (E. coli) | 552900 |
| BUD31 | homolog (S. cerevisiae) | 8896 |
| CAD | carbamoyl-phosphate synthetase 2 | 790 |
| CD52 | CD52 molecule | 1043 |
| CDC2L6 | cell division cycle 2-like 6 | 23097 |
| CDC7 | cell division cycle 7 homolog | 8317 |
| CNIH4 | cornichon homolog 4 (Drosophila) | 29097 |
| COX6B1 | cytochrome c oxidase subunit Vib polypeptide 1 | 1340 |
| CRIP1 | Cysteine-rich protein 1 | 1396 |
| CROCC | ciliary rootlet coiled-coil, rootletin | 9696 |
| DDX39 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 39 | 10212 |
| DR1 | down-regulator of transcription 1, TBP-binding (negative cofactor 2) | 1810 |
| EIF1AY | eukaryotic translation initiation factor 1A, Y-linked | 9086 |
| G6PD | glucose-6-phosphate dehydrogenase | 2539 |
| GBP2 | guanylate binding protein 2, interferon-inducible | 2634 |
| GNG5 | guanine nucleotide binding protein (G protein), gamma 5 (GNG5) | 2787 |
| GPR172A | protein-coupled receptor 172A | 79581 |
| GRN | granulin (GRN) | 2896 |
| HLA-F | Major histocompatibility complex, class I, F | 3134 |
| HLA-H | Major histocompatibility complex, class I, H (pseudogene) | 3136 |
| HN1 | hematological and neurological expressed 1, transcript variant 2, | 51155 |
| IER5 | immediate early response 5 | 51278 |
| IFI27 | interferon, alpha-inducible protein 27 | 3429 |
| ING1 | inhibitor of growth family, member 1, transcript variant 3 | 3621 |
| IRF5 | interferon regulatory factor 5, transcript variant 7, | 3663 |
| ISG15 | ubiquitin-like modifier | 9636 |
| KRTCAP2 | keratinocyte associated protein 2 | 200185 |
| LRP10 | low density lipoprotein receptor-related protein 10 | 26020 |
| M6PRBP1 | mannose-6-phosphate receptor binding protein 1 | 10226 |
| MAFB | maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian) | 9935 |
| MAP4K2 | mitogen-activated protein kinase kinase kinase kinase 2 | 5871 |
| MZF1 | myeloid zinc finger 1, transcript variant 1 | 7593 |
| NUP62 | nucleoporin 62kDa, transcript variant 4, | 23636 |
| OGFR | opioid growth factor receptor | 11054 |
| PIK3IP1 | phosphoinositide-3-kinase interacting protein 1 | 113791 |
| PLCB1 | phospholipase C, beta 1 (phosphoinositide-specific), transcript variant 1 | 23236 |
| PLXDC2 | plexin domain containing 1 | 84898 |
| POP7 | processing of precursor 7, ribonuclease P/MRP subunit (S. cerevisiae) | 10248 |
| PPP1CC | protein phosphatase 1, catalytic subunit, gamma isoform | 5501 |
| PRDM1 | PR domain containing 1, with ZNF domain, transcript variant 2 | 639 |
| PRKD2 | protein kinase D2, transcript variant 4 | 25865 |
| PSIP1 | PC4 and SFRS1 interacting protein 1, transcript variant 2 | 11168 |
| PURA | purine-rich element binding protein A | 5813 |

| | | |
|---|---|---|
| RBP7 | retinol binding protein 7, cellular | 116362 |
| RELB | v-rel reticuloendotheliosis viral oncogene homolog B | 5971 |
| RNF31 | ring finger protein 31 | 55072 |
| RRAGD | Ras-related GTP binding D | 58528 |
| SERPINB1 | serpin peptidase inhibitor, clade B (ovalbumin), member 1 | 1992 |
| SGSH | N-sulfoglucosamine sulfohydrolase (sulfamidase) | 6448 |
| SMARCC2 | SWI/SNF related, matrix associated, actin dependent regulator of chromatin | 6601 |
| SQSTM1 | sequestosome 1 | 8878 |
| ST6GALNAC4 | T6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1, 3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4 | 27090 |
| STAT1 | signal transducer and activator of transcription 1, 91kDa, transcript variant beta | 6772 |
| TERF1 | telomeric repeat binding factor (NIMA-interacting) 1, transcript variant 2 | 7013 |
| TK1 | thymidine kinase 1, soluble | 7083 |
| TMEM106C | transmembrane protein 106C | 79022 |
| TXN | thioredoxin | 7295 |
| TYK2 | tyrosine kinase 2 | 7297 |
| USP14 | ubiquitin specific peptidase 14 (tRNA-guanine transglycosylase), transcript variant 2 | 9097 |
| USP3 | ubiquitin specific peptidase 3 | 9960 |
| VWF | von Willebrand factor | 7450 |
| WBP5 | WW domain binding protein 5 | 51186 |
| ZBTB17 | zinc finger and BTB domain containing 17 | 7709 |
| ZNF580 | zinc finger protein 580, transcript variant 2 | 51157 |
| ZNF652 | zinc finger protein 652 | 22834 |
| ZNF706 | zinc finger protein 706, transcript variant 3 | 51123 |

**Table 7** *Table of genes associated with HCV-related HCC*

| NAME | DESCRIPTION | "ENTREZ GENE" |
|---|---|---|
| ANXA4 | annexin A4 | 307 |
| APOL3 | apolipoprotein L, transcript variant alpha/d | 80833 |
| AQP1 | aquaporin 1 (Colton blood group) | 358 |
| ARL8A | ADP-ribosylation factor-like 8A | 127829 |
| B2M | beta-2-microglobulin | 567 |
| BET1L | blocked early in transport 1 homolog (S. cerevisiae)-like, transcript variant 1 | 51272 |
| BTN3A3 | butyrophilin, subfamily 3, member A3, transcript variant 2 | 10384 |
| CKB | creatine kinase, brain | 1152 |
| CNN2 | calponin 2, transcript variant 2 | 1265 |
| CRIP1 | Cysteine-rich protein 1 (intestinal) | 1396 |
| CROCC | CROCC--ciliary rootlet coiled-coil, rootletin | 9696 |
| HLA-C | major histocompatibility complex, class I, C | 3107 |
| HLA-DMA | DMA--major histocompatibility complex, class II, DM alpha | 3108 |
| HLA-DPB1 | DPB1--major histocompatibility complex, class II, DP beta 1 | 3115 |
| HLA-F | Major histocompatibility complex, class I, F | 3134 |
| HLA-H | Major histocompatibility complex, class I, | 3136 |
| IFI6 | interferon, gamma-inducible protein 16 | 2537 |
| IGHG1 | Immunoglobulin heavy constant mu | 3500 |
| IGL@ | Immunoglobulin lambda joining 3 | 3535 |
| ISG15 | ubiquitin-like modifier | 9636 |
| ISG20 | interferon stimulated exonuclease gene 20kDa | 3669 |
| KIAA0746 | KIAA0746 protein | 23231 |
| LIPT1 | lipoyltransferase 1, nuclear gene encoding mitochondrial protein, transcript variant 5 | 51601 |
| OAS1 | 2',5'-oligoadenylate synthetase 1, 40/46kDa , transcript variant 2 | 4938 |
| OASL | 2'-5'-oligoadenylate synthetase-like, transcript variant 2 | 8638 |
| PKM2 | pyruvate kinase, muscle, transcript variant 1 | 5315 |
| PSMB9 | Proteasome (prosome, macropain) subunit, beta type, 9 (large multifunctional peptidase 2) | 5698 |
| RARRES3 | retinoic acid receptor responder (tazarotene induced) 3 | 5920 |
| SAFB2 | scaffold attachment factor B2 | 9667 |
| SEMA4D | sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) | 10507 |
| STAT1 | signal transducer and activator of transcription 1, 91kDa, transcript variant beta | 6772 |
| TAP1 | transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) | 6890 |
| TESK1 | testis-specific kinase 1 | 7016 |
| TMEM55A | transmembrane protein 55A | 55529 |
| UBD | ubiquitin D | 10537 |

**Table 8** *Table of genes associated with HCV-related non HCC*

| NAME | DESCRIPTION | "ENTREZ GENE" |
|---|---|---|
| ACSS1 | acyl-CoA synthetase short-chain family member 1 | 84532 |
| ACTG1 | actin, gamma 1 | 71 |
| AFAP1 | actin filament associated protein 1, transcript variant 1 | 60312 |
| ALDOA- | aldolase A, fructose-bisphosphate , transcript variant 2 | 226 |
| ANXA2 | annexin A2, transcript variant 3 | 302 |
| AREG | amphiregulin (schwannoma-derived growth factor) | 374 |
| ARFGAP 1 | ADP-ribosylation factor GTPase activating protein 1, transcript variant 1 | 55738 |
| ASNS | asparagine synthetase, transcript variant 2 | 440 |
| ATP9A | ATPase, Class II, type 9A | 10079 |
| AURKA | aurora kinase A, transcript variant 6 | 6790 |
| BOLA2 | bolA homolog 2 (E. coli), transcript variant 1 | 552900 |
| BSG | basigin (Ok blood group), transcript variant | 682 |
| CCT6A | chaperonin containing TCP1, subunit 6A (zeta 1), transcript variant 2 | 908 |
| CD44 | CD44 molecule (Indian blood group), transcript variant 5 | 960 |
| CD9 | CD9 molecule | 928 |
| CDH3 | cadherin 3, type 1, P-cadherin (placental) | 1001 |
| CEACA M5 | Carcinoembryonic antigen-related cell adhesion molecule 5 | 1048 |
| CENPO | centromere protein O | 79172 |
| CKB | creatine kinase, brain | 1152 |
| CLIC1 | chloride intracellular channel 1 | 1192 |
| COL3A1 | collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant) | 1281 |
| CSDA | cold shock domain protein A | 8531 |
| CTBP2 | terminal binding protein 2, transcript variant | 1488 |
| DPEP1 | dipeptidase 1 (renal) | 1800 |
| ELF3 | E74-like factor 3 (ets domain transcription factor, epithelial-specific ). | 1999 |
| FAM60A | FAM60A--family with sequence similarity 60, member A | 58516 |
| FHL2 | four and a half LIM domains 2, transcript variant 4 | 2274 |
| GPR160 | G protein-coupled receptor 160 | 26996 |
| GPR56 | G protein-coupled receptor 56, transcript variant 3 | 9289 |
| GSTP1 | GSTP1--glutathione S-transferase pi | 2950 |
| HEPH | hephaestin, transcript variant 2 | 341208 |
| HKDC1 | hexokinase domain containing 1 | 80201 |
| HN1 | hematological and neurological expressed 1, transcript variant 2 | 51155 |
| IER3 | immediate early response 3 | 8870 |
| IL8 | interleukin 8 | 3576 |
| KIAA074 6 | KIAA0746 protein | 23231 |
| KRT18 | keratin 18, transcript variant 2 | 3875 |
| LAMB3 | laminin, beta 3, transcript variant 1 | 3914 |
| LGALS3 | lectin, galactoside-binding, soluble, transcript variant 1 | 3958 |
| LGALS4 | lectin, galactoside-binding, soluble, 4 (galectin 4) | 3960 |
| MCM2 | minichromosome maintenance complex component 2 | 4171 |
| MYC | v-myc myelocytomatosis viral oncogene homolog (avian) | 4609 |
| NBL1 | neuroblastoma, suppression of tumorigenicity 1, transcript variant 2 | 4681 |
| NQO1 | NAD(P)H dehydrogenase, quinone 1, transcript variant 3 | 1728 |
| NUP93 | nucleoporin 93kDa | 9688 |
| OGT | O-linked N-acetylglucosamine | 8473 |

| | | |
|---|---|---|
| OR2T10 | olfactory receptor, family 2, subfamily T, member 10 | 127069 |
| PAQR8 | progestin and adipoQ receptor family member VIII | 85315 |
| PHLDA2 | pleckstrin homology-like domain, family A, member 2 | 7262 |
| PKM2 | pyruvate kinase, muscle, transcript variant 1 | 5315 |
| PTMA | prothymosin, alpha (gene sequence 28), transcript variant 2 | 5757 |
| PTP4A3 | protein tyrosine phosphatase type IVA, member 3 , transcript variant 2 | 11156 |
| PYGB | phosphorylase, glycogen; brain | 5834 |
| RALGDS | ral guanine nucleotide dissociation stimulator, transcript variant 1 | 5900 |
| S100A10 | S100 calcium binding protein A10 | 6281 |
| S100A11 | S100 calcium binding protein A11 | 6282 |
| S100A6 | S100A6--S100 calcium binding protein A6 | 6277 |
| S100P | S100P--S100 calcium binding protein P | 6286 |
| SERPINB 1 | serpin peptidase inhibitor, clade B (ovalbumin), member 1 | 1992 |
| SMARC A4 | SWI/SNF related, matrix associated, actin dependent regulator of chromatin | 6597 |
| SNRPB | small nuclear ribonucleoprotein polypeptides B and B1, transcript variant 2 | 6628 |
| SOX9 | SRY (sex determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal) | 6662 |
| SPINT2 | serine peptidase inhibitor, Kunitz type, 2 | 10653 |
| STK24 | serine/threonine kinase 24 (STE20 homolog, yeast), transcript variant 1 | 8428 |
| STMN3 | stathmin-like 3 | 50861 |
| SYK | spleen tyrosine kinase | 6850 |
| TAX1BP 3 | Tax1 (human T-cell leukemia virus type I) binding protein 3 | 30851 |
| TFDP1 | transcription factor Dp-1 | 7027 |
| THY1 | Thy-1 cell surface antigen | 7070 |
| TIMP1 | TIMP metallopeptidase inhibitor 1 | 7076 |
| TK1 | thymidine kinase 1, soluble | 7083 |
| TMED3 | transmembrane emp24 protein transport domain containing 3 | 24423 |
| TMSB10 | thymosin, beta 10 | 9168 |
| TRIM5 | tripartite motif-containing 5, transcript variant alpha | 85363 |
| UTX | ubiquitously transcribed tetratricopeptide repeat, X chromosome | 7403 |
| VASP | VASP--vasodilator-stimulated phosphoprotein | 7408 |
| VDR | vitamin D (1,25- dihydroxyvitamin D3) receptor, transcript variant 1 | 7421 |
| VIL2 | villin 2 (ezrin) | 7430 |
| YWHAZ | tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein | 7534 |

**Table 9** *Table of genes associated with Liver metastasis.*

**Gene signature Involved in HCC progression**

Time course analysis was conducted considering the CTR the early time point, the cirrhosis the intermedian point and the HCC the last time point. All samples were average belonging at the same group as it was a single array.

Testing the significant gene expression increase or decrease across the tissue types from normal (n = 6) to HCV related non HCC (n =20) to HCV-related HCC (n = 20) 450 genes with a significant decreasing trend and 136 genes with a significant increasing trend in expression values were identified. Genes with a significant increasing trend in expression values were considered as possible diagnostic and prognostic markers (Figure 11).

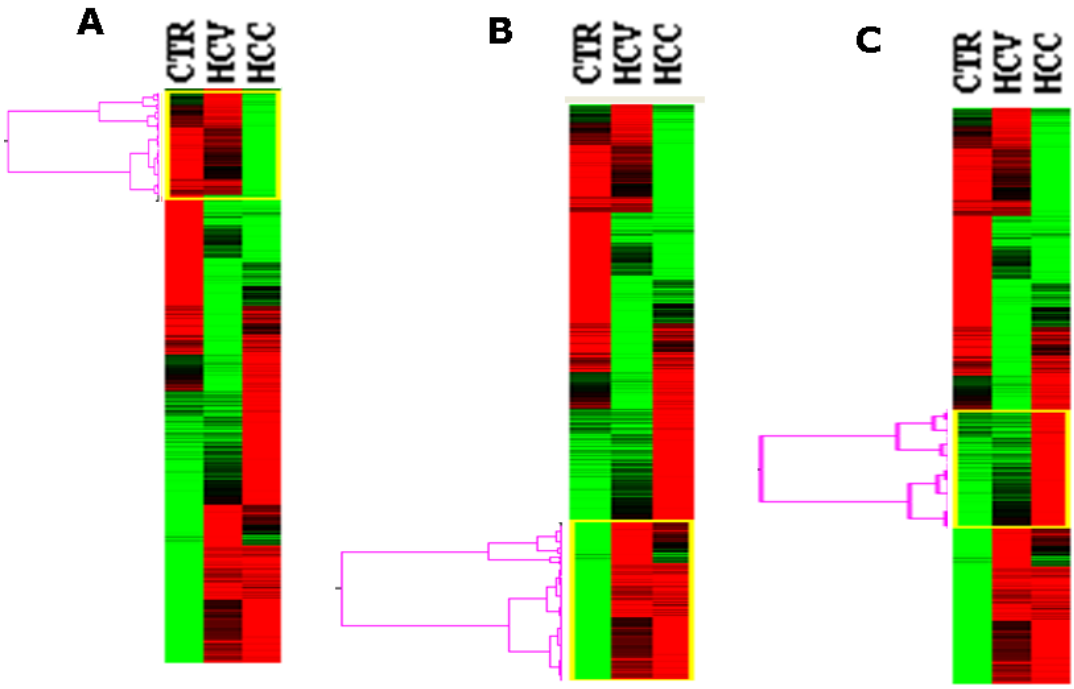


**Figure 11** *Time course analysis. (A) Genes up regulated in the CTR and then lost completely in the HCC(decreasing trend);(B) Genes switched off in the control and up regulated in HCV positive and in the HCC; (C) Genes were down-regulated in CTR and HCV positive and up-regulated in HCC (increasing trend).*

Relevant genes are reported in Table 10.

| NAME | DESCRIPTION | "ENTREZ GENE" |
|---|---|---|
| ADH1A. | alcohol dehydrogenase 1A (class I), alpha polypeptide | 124 |
| AKR1B10 | aldo-keto reductase family 1, member B10 (aldose reductase) | 57016 |
| ANKRD29 | ankyrin repeat domain 29 | 147463 |
| APOA1 | apolipoprotein A-I | 335 |
| ASCC3 | activating signal cointegrator 1 complex subunit 3, transcript variant 1 | 10973 |
| AURKC | Aurora kinase C | 6795 |
| BHMT | betaine-homocysteine methyltransferase | 635 |
| CLEC4G | C-type lectin superfamily 4, member G | 339390 |
| COL1A1 | collagen, type I, alpha 1 | 1277 |
| COL4A2 | collagen, type IV, alpha 2 | 1284 |
| CXCL12 | chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1), transcript variant 1 | 6387 |
| CXCL2 | chemokine (C-X-C motif) ligand 2 | 2920 |
| DNASE1L3 | deoxyribonuclease I-like 3 | 6387 |
| DPT | dermatopontin | 1805 |
| ENO3 | enolase 3 (beta, muscle), transcript variant 2 | 2027 |
| FOS | v-fos FBJ murine osteosarcoma viral oncogene homolog | 2353 |
| GLUL | glutamate-ammonia ligase (glutamine synthetase), transcript variant 2 | 2752 |
| GPC3 | glypican 3 | 2719 |
| HAL | histidine ammonia-lyase | 3034 |
| HAO2 | hydroxyacid oxidase 2 (long chain), transcript variant 1 | 51179 |
| HMGCS2 | 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 | 3158 |
| IFI27 | interferon, alpha-inducible protein 27 | 3429 |
| IGF2 | insulin-like growth factor 2 (somatomedin A) (IGF2), transcript variant 2, mRNA. | 3481 |
| IGKC | Netrin 2-like | 3514 |
| IQCH | IQ motif containing H, transcript variant 1 | 64799 |
| ISG15 | ISG15 ubiquitin-like modifier | 9636 |
| ITIH3 | inter-alpha (globulin) inhibitor H3 (ITIH3), mRNA. | 3699 |
| LIPT1 | lipoyltransferase 1, nuclear gene encoding mitochondrial protein, transcript variant 5 | 51601 |
| LRRC46 | leucine rich repeat containing 46 | 90506 |
| LRRC8D | leucine rich repeat containing 8 family, member D | 55144 |
| RGL3 | Ral guanine nucleotide dissociation stimulator-like 3 | 57139 |
| RNF125 | ring finger protein 125 | 54941 |
| SPINK1 | serine peptidase inhibitor, Kazal type 1 | 6690 |
| STAT1 | signal transducer and activator of transcription 1, 91kDa , transcript variant beta | 6772 |
| TDRD1 | tudor domain containing 1 | 56165 |
| THY1 | Thy-1 cell surface antigen | 7070 |
| TRIM55 | tripartite motif-containing 55, transcript variant 4 | 84675 |
| UBD | ubiquitin D | 10537 |

**Table 10** *Genes associated with liver disease progression to HCC*

**Ingenuity Pathway Analysis**

To define the biological significance of the genes with relevant expression modifications their biological interactions was investigated using the IPA tool and all genes were mapped to their molecular/cellular functions and to relevant canonical pathways.

The more important molecular and cellular functions (High *p* value) of the genes up-regulated in HCV-related HCC samples were related to cell death (*p*=1.11E-04-3.28E-02) [81 molecules]), cell to cell signaling and interaction (*p*=8.09E-06-3.28E-02) [116 molecules]) and Antigen Presentation (*p*=1.25E-04 - 3.28E-02 56 [56 molecules]).The top canonical pathways included protein ubiquitination (*p*=1.67E-05), antigen presentation (*p*=9.52E-04) and Aryl Hydrocarbon receptor signaling pathway (*p*=1.37E-03) (Figure 12).

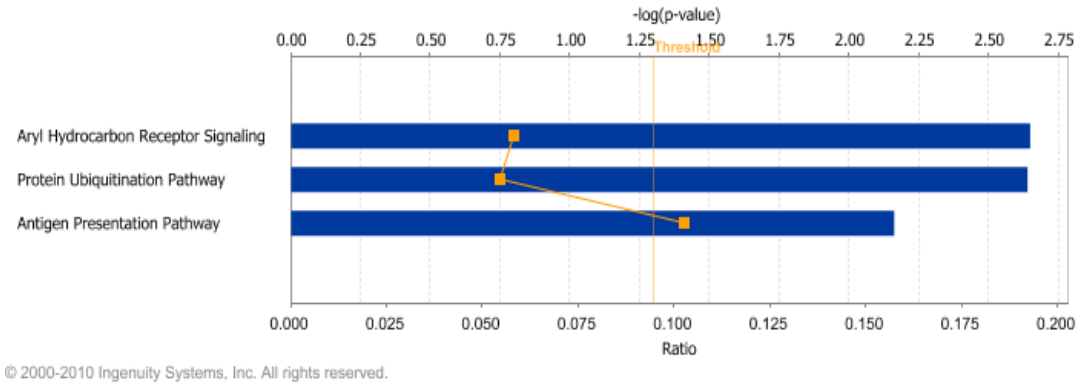


**Figure 12** *The human pathway lists determined by "Ingenuity System Database" in HCV-related HCC samples.*

The more important molecular and cellular functions (High *p* value) of genes up-regulated in HCV-related non HCC samples were related to cellular growth and proliferation (*p* = 2.60E-09 - 1.59E-02) [112 molecules]), Antigen Presentation (*p*=3.63E-08 - 1.19E-02) [77 molecules]) and cell to cell signaling and interaction (*p*=1.30E-05 - 1.59E-02 [76 molecules]).The top canonical pathways included Interferon Signaling Genes(*p*=1.12E-05), SAPK/JNK Signaling (*p*=1.07E-03) and NF-kB Activation by viruses pathway (*p*=1.19E-03) (Figure 13).

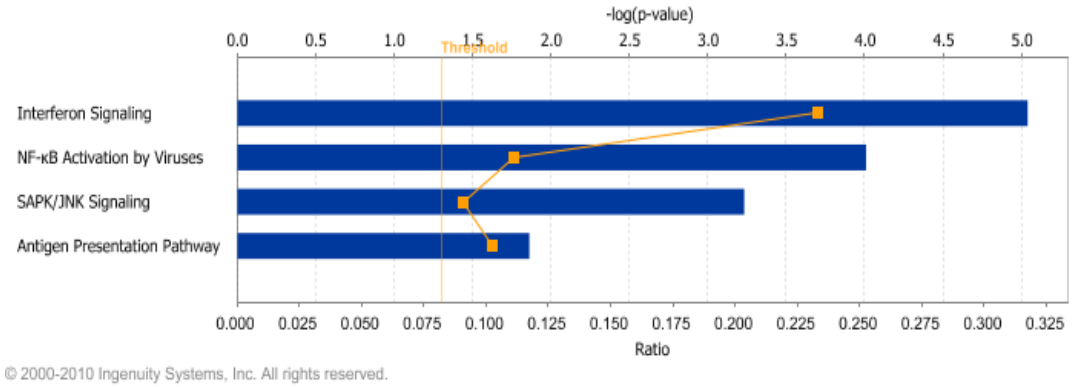



**Figure 13** *The human pathway lists determined by "Ingenuity System Database" in HCV-related non HCC samples*

The more important molecular and cellular functions (High *p* value) of genes up-regulated in metastasis were related to cellular growth and proliferation (*p*=1.83E-09 - 1.28E-02 [264 molecules])cell death (*p*=5.92E-10 - 1.24E-02 [259 molecules]), and cellular assembly and organization (*p*=1.99E-08 - 1.12E-02 [111 molecules]).The top canonical pathways included Integrin Signaling (*p=7.75E-04*) and Actin Cytoskeleton Signaling Pathway (*p*=4.43E-04).

## DISCUSSION

HCC is a common and aggressive malignant tumour worldwide with a dismal outcome. Early detection and resection may offer an opportunity to improve the long-term survival for HCC patients. Unfortunately, with current diagnostic approaches, only about 10% to 20% of HCC patients are eligible for resection (Lai EC *et al* 1995).

The present study has been focused on investigating the genes/protein and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients, to elucidate the molecular mechanisms underlying cancer progression and to identify possible marker for diagnostic purposes trough DNA microarray. The study was conducted performing two set of experiments.

First the study was conducted on liver biopsies with a small number of HCV-positive HCC patients and HCV-negative non-liver cancer control patients to investigate genes and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients. Then, in order to verify the consistency of the previous data obtained in a very limited sample and to identify a set of genes sufficient for the molecular signature of liver diseases, the study was conducted on a larger number of liver biopsies from HCV-positive HCC patients, metastatic patients and HCV-negative non-liver cancer control patient.

In the first batch of samples, microarray analyses of liver biopsies from HCC nodules and paired non-adjacent non-HCC liver tissue of the same HCV-positive patients were compared to biopsies from HCV-negative control subjects. The class comparison analysis used in that study successfully identified a set of genes significant differentially expressed. Moreover the up-regulated genes identified within the individual class comparison analysis were evaluated and classified by a pathway analysis, according to the "Ingenuity System Database".

The genes up-regulated in samples from HCV-related HCC were classified in metabolic pathways, and the most represented are the Aryl Hydrocarbon receptor signaling (AHR) and, protein Ubiquitination pathways, which have been previously reported to be involved in cancer, and in particular in HCC, progression.

The Aryl Hydrocarbon receptor signal transduction Pathway (AHR) is involved in the activation of the cytosolic aryl hydrocarbon receptor by structurally diverse xenobiotic ligands (including dioxin, and polycyclic or halogenated aromatic hydrocarbons) and mediating their toxic and carcinogenic effects *(*Safe S *et al 2001,* Okey AB 2007*)*. More recently AHR pathway has been shown to be involved in apoptosis, cell cycle regulation, mitogen-activated protein kinase cascades *(*Puga A *et al* 2009*)*. In particular, studies on liver tumour promotion have shown that dioxin-induced AHR activation mediates clonal expansion of initiated cells by inhibiting apoptosis and bypassing AHR-dependent cell cycle arrest *(*Bock KW *et al* 2005*)*. Furthermore, it has been shown that changes in mRNA expression of specific genes in the AHR pathway are linked to progression of HCV-associated hepatocellular carcinoma *(*Tsunedomi R *et al* 2005*)*. Moreover, the HCV-induced AHR signal transduction pathway, could be directly involved in the increased severity of hepatic lesions in patients with chronic hepatitis C induced by smoking *(*Pessione F *et al* 2001*,* Hezode C *et al* 2003*)*. The ubiquitin and ubiquitin-related proteins of the ubiquitination pathway play instrumental roles in cell-cycle regulation *(*Jentsch S, et al 2000*)* as well as cell death/apoptosis *(*Jesenberger V *et al 2002*) through modification of target proteins. In particular, ubiquitin-like proteins, i.e. FAT10, has been reported to bind non-covalently to the human spindle assembly checkpoint protein, MAD2 *(*Liu YC, *et al* 1999*),* which is responsible for maintaining spindle integrity during mitosis *(*Shah JV *et al 2000)* and whose inhibited function has been associated with chromosomal instability *(*Wang X *et al* 2000*,* Gemma A *et al* 2001*)*. Moreover, FAT10 overexpression has been previously shown in hepatocellular carcinoma *(*Lee CG *et al* 2003*)*.

The genes up-regulated in samples from HCV-related non-HCC tissue were classified in several pathways prevalently associated to inflammation and native/adaptive immunity and most of the over expressed genes belong to the Antigen Presentation pathway.

The analysis of second batch of samples was performed with the same statistical analysis under the same condition to confirm the first data. To elucidate the genes and molecular pathways involved in the HCV-related HCC a class comparisons

analysis were performed on new samples set. This analysis allowed us to identify the unique probe sets characterizing the pathological status, in fact as expected, the gene expression patterns were found to vary significantly among the HCC and normal control liver samples. Genes associated with cell death, cell to cell signaling and interaction, were found to have increased expression in HCC samples. The molecular events linked to the development and progression of HCC is not well known. Malignant hepatocytes are the result of sequential changes accumulated in mature hepatocytes or can derive from stem cells. The most accepted hypotheses (Buendia MA *et al* 2000, Thorgeirsson SS *et a*l 2002) describes a step-by-step process in which external stimuli induce genetic alterations in mature hepatocytes leading to cell death, cellular proliferation, and the production of monoclonal populations. These populations harbor dysplastic hepatocytes that evolve to dysplastic nodules. (Theise ND *et a*l 2002)

Canonical pathways prevalently associated with HCV-related HCC included protein ubiquitination, antigen presentation and Aryl Hydrocarbon receptor signaling pathway, confirming the first data.

Cellular growth and proliferation and antigen presentation were the more important cellular and molecular functions when HCV-related non HCC samples were compared with normal control liver tissue. These data agree with the numerous regulatory roles reported for the HCV core, that affect signal transduction, expression of viral and cellular genes, cell growth and proliferation (Lohmann V *et al* 2006, Blight KJ *et al* 1998).

Several viruses target specific components of the MHC class I pathway, leading to diminished cell surface expression of MHC class I molecules. Other viruses block the transport of MHC class I molecules through the endoplasmic reticulum (ER), inhibit TAP-mediated translocation of cytoplasmic peptides into the ER, or interfere with proteasomal degradation of their own proteins (Rosenberg W 1999). Other viruses, like human cytomegalovirus, escape CD8_T-cell recognition by down regulating cellular MHC class I molecules (Falk CS, *et al* 2002) and simultaneously inducing the expression of virus-encoded MHC class I homologues capable of engaging inhibitory receptors that give a negative signal blocking NK cell function. Flaviviruses can upregulate MHC class I cell surface

expression by increased peptide supply to the ER (Momburg F *et al* 2001, Mullbacher A *et al* 1995). Viruses may use these strategies to evade and counteract a potential NK cell attack. Some studies demonstrated that HCV core protein induced the up regulation of antigen presentation and immune response mechanisms (Herzer K *et al* 2003).

Canonical pathways mainly associated with HCV-related non-HCC tissue included Interferon Signaling, SAPK/JNK Signaling and NF-kB Activation by viruses pathway. These pathways are prevalently associated with inflammation and native/adaptive immunity.

A traditional HCC diagnosis has relied on the use of a single biomarker approach (e.g., AFP). The use of multiple markers in combination to improve the accuracy of identifying HCC cases has been proposed. Among the three different class comparison analysis (HCV-related HCC, HCV-related non HCC and Metastatic liver tissue vs normal control) a gene-set that distinguish the different cases of liver disease has been found, in particular with time course analysis the genes that should be candidate as a possible progression markers (e.g., OCM, PTK6, FAM38B, FBLN1, STEAP1, MOXD1, MALAT1, GPC3, CCL20, SPINK1, GLUL,THSD7A, RARRES1, ALPK2, DEFB1, SERPINA7, UBD, GSTA1, TM4SF1, DPT, SAA2, SCD, MAL2, OS9, DPYS, TMED9, COL4A2, COL4A1) has been identified. All these data altogether suggested developing a specific gene-chip along with genes showing the highest fold up-regulation in common with the first batch of samples representing the different stage of disease.

The identification of the lesions and the evaluation of their neoplastic progression will be based on the gene pattern expression on the gene-chip.

In conclusion in this study:

1. Informative data on the global gene expression pattern in HCV-related HCC as well as HCV-related non-HCC counterpart and metastatic liver tissues have been obtained compared to normal controls.

**Research Article 1**
"Gene profiling, biomarkers and pathways characterizing HCV-related hepatocellular carcinoma" **Valeria De Giorgi**, Alessandro Monaco, Andrea Worchech, Maria Lina Tornesello, Francesco Izzo, Luigi Buonaguro, Francesco M Marincola, Ena Wang, Franco M Buonaguro *Journal of Translational Medicine 2009, 7:85*

2. Gene signatures that distinguish the different pathological stage of liver disease and potential molecular progression markers for early HCC diagnosis in HCV positive patients were identified.

**Research Article 2**
"Molecular Signature Associated To Human Liver Cancer" **V. De Giorgi**, L. Buonaguro, A. Worschech, M.L. Tornesello, F. Izzo, F.M. Marincola, E. Wang and F.M. Buonaguro (Submitted)

**Patent**
"Method for biomolecular detection of human liver diseases composition and kits used in said methods". **De Giorgi V**.*et al* PCT/EP2009/062716

Analysis of the same samples in relation to new contexts (miRNA, aCGH, proteomic) to develop increasingly sophisticated gene expression indicators of specific types or stages of liver disease has been planned.

## REFERENCES

Alizadeh A, Eisen M, Davis RE, Ma C, Sabet H, Tran T, Powell JI, Yang L, Marti GE, Moore DT, Hudson JR,Jr, Chan WC, Greiner T, Weisenburger D, Armitage JO, Lossos I, Levy R, Botstein D, Brown PO, Staudt LM. The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harb. Symp. Quant. Biol. 1999, 64, 71-78. 125.*

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J,Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature 2000, 403, 503-511. 126*.

Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, Levine AJ. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A. 1999, 96, 6745-6750.*

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J.Roy.Stat.Soc.B Met. 1995, 57, 289-300.*

Blight KJ, Kolykhalov AA, Reed KE, Agapov EV, Rice CM: Molecular virology of hepatitis C virus: an update with respect to potential antiviral targets. *Antivir Ther 1998, 3: 71-81.*

Blum HE, Hepatocellular carcinoma: therapy and prevention. *World J Gastroenterol 2005, 11: 7391–7400.*

Bock KW, Kohle C. Ah receptor- and TCDD-mediated liver tumor promotion: clonal selection and expansion of cells evading growth arrest and apoptosis. *Biochem Pharmacol.* 2005;69:1403–1408.

Bosch FX, Ribes J, Borràs J. Epidemiology of primary liver cancer Semin Liver Dis. *1999;19(3):271-85.*

Brechot C. Molecular mechanisms of hepatitis B and C viruses related to liver carcinogenesis Hepatogastroenterology. *1998 Aug;45 Suppl 3:1189-96.*

Brent R. Genomic biology. *Cell 2000;100:169–83*.

Brunt EM: Nonalcoholic steatohepatitis. *Semin Liver Dis 2004, 24:3-20*

Bubendorf L, Kolmer M, Kononen J, Koivisto P, Mousses S, Chen Y, Mahlamaki E, Schraml P, Moch H, Willi N, Elkahloun AG, Pretlow TG, Gasser TC, Mihatsch MJ, Sauter G, Kallioniemi OP. Hormone therapy failure in human

prostate cancer: analysis by complementary DNA and tissue microarrays. *J. Natl. Cancer Inst. 1999, 91, 1758- 1764. 120.*

Buendia MA: Genetics of hepatocellular carcinoma. *Semin Cancer Biol 2000, 10: 185-200*.

Chung C.H., Bernard P.S., Perou C.M.: Molecular portraits and the family tree of cancer. *Nat Genet 2002 Dec, VOL:32 (Suppl), P:533*.

Cunliffe HE, Ringner M, Bilke S, Walker RL, Cheung JM, Chen Y, Meltzer PS. The gene expression response of breast cancer to growth regulators: patterns and correlation with tumor expression profiles. *Cancer Res. 2003, 63, 7158-7166*.

Davila JA, Morgan RO, Shaib Y, McGlynn KA, El-Serag HB: Diabetes increases the risk of hepatocellular carcinoma in the United States: a population based case control study. *Gut 2005, 54:533-539.*

Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM. Delineation of prognostic biomarkers in prostate cancer. *Nature 2001, 412, 822-826. 121.*

Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA 1998, 95:14863-14868.*

Falk CS, Mach M, Schendel DJ, Weiss EH, Hilgert I, Hahn G: NK cell activity during human cytomegalovirus infection is dominated by US2-11-mediated HLA class I down-regulation. *J Immunol 2002, 169: 3257-3266*

Feitelson MA, Sun B, Satiroglu Tufan NL, Liu J, Pan J and Lian Z. Genetic mechanisms of hepatocarcinogenesis. *Oncogene 2002, 21: 2593–2604*.

Fusco M, Girardi E, Piselli P, Palombino R, Polesel J, Maione C. Epidemiology of viral hepatitis infections in an area of southern Italy with high incidence rates of liver cancer. *Eur J Cancer 2008, 44:847-853.*

Gemma A, Hosoya Y, Seike M, Uematsu K, Kurimoto F, Hibino S, et al. Genomic structure of the human MAD2 gene and mutation analysis in human lung and breast cancers. *Lung Cancer. 2001;32:289–295*

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science 1999, 286, 531-537. ,124-126 124.*

Golub TR. Genomic approaches to the pathogenesis of hematologic malignancy. *Curr. Opin. Hematol. 2001, 8, 252-261*

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science 1999, 286, 531-537*

Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell 2000;100: 57–70.*

He X, Wei Q, Sun M, Fu X, Fan S and Li Y. LS-CAP: an algorithm for identifying cytogenetic aberrations in hepatocellular carcinoma using microarray data. *Front Biosci 2006, 11: 1311–1322.*

Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A, Trent J. Gene expression profiles in hereditary breast cancer. *N. Engl. J. Med. 2001, 344, 539-548.*

Herath NI, Leggett BA and MacDonald GA. Review of genetic and epigenetic alterations in hepatocarcinogenesis. *J Gastroenterol Hepatol 2006, 21: 15–21.*

Herzer K, Falk CS, Encke J, Eichhorst ST, Ulsenheimer A, Seliger B et al.: Upregulation of major histocompatibility complex class I on liver cells by hepatitis C virus core protein via p53 and TAP1 impairs natural killer cell cytotoxicity. *J Virol 2003, 77: 8299-8309.*

Hezode C, Lonjon I, Roudot-Thoraval F, Mavier JP, Pawlotsky JM, Zafrani ES, et al. Impact of smoking on histological liver lesions in chronic hepatitis C. *Gut. 2003;52:126–129.*

Jentsch S, Pyrowolakis G. Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol. 2000;10:335–342.*

Jesenberger V, Jentsch S. Deadly encounter: ubiquitin meets apoptosis. *Nat Rev Mol Cell Biol. 2002;3:112–121*.

Jiang J, Gusev Y, Aderca I, Mettler TA, Nagorney DM, Brackett DJ and Roberts LR. Association of MicroRNA expression in hepatocellular carcinomas with hepatitis infection, cirrhosis, and patient survival. *Clin Cancer Res 2008, 14: 419–427.*

Kohonen T. Self-Organizing Maps. 3rd extended. *New York, Springer, 2001*

Lai EC, Fan ST, Lo CM, Chu KM, Liu CL, Wong J: Hepatic resection for hepatocellular carcinoma. An audit of 343 patients. *Ann Surg 1995, 221: 291-298*
Lau SH and Guan XY. Cytogenetic and molecular genetic alterations in hepatocellular carcinoma. *Acta Pharmacol Sin 2005, 26: 659–665*.

Lee CG, Ren J, Cheong IS, Ban KH, Ooi LL, Yong TS, et al. Expression of the FAT10 gene is highly upregulated in hepatocellular carcinoma and other gastrointestinal and gynecological cancers. *Oncogene. 2003;22:2592–2603*

Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. U. S. A.* 2001, 98, 31-36

Liu YC, Pan J, Zhang C, Fan W, Collinge M, Bender JR, et al. A MHC-encoded ubiquitin-like protein (FAT10) binds noncovalently to the spindle assembly checkpoint protein MAD2. *Proc Natl Acad Sci USA. 1999;96:4313–4318*

Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. *Nature 2000;405:827–36.*

Lohmann V, Korner F, Koch J, Herian U, Theilmann L, Bartenschlager R: Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science 1999, 285: 110-113*

MacQueen J. *Some* Methods for classification and Analysis of Multivariate Observations. *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability 1967, 1, 281-297. 56.*

Mazumder A., Wang Y.: Gene-expression signatures in oncology diagnostics. *Pharmacogenomics 2006 Dec, VOL: 7(8), P:1167.*

Midorikawa Y, Makuuchi M, Tang W and Aburatani H. Microarray-based analysis for hepatocellular carcinoma: from gene expression profiling to new challenges. *World J Gastroenterol 2007, 13: 1487–1492*.

Momburg F, Mullbacher A, Lobigs M: Modulation of transporter associated with antigen processing (TAP)-mediated peptide import into the endoplasmic reticulum by flavivirus infection. *J Virol 2001, 75: 5663-5671*

Mullbacher A, Lobigs M: Up-regulation of MHC class I by flavivirus-induced peptide translocation into the endoplasmic reticulum. *Immunity 1995, 3: 207-214*.

Oh JH, Kim YB, Gurnani P, Rosenblatt K and Gao JX. Biomarker selection and sample prediction for multi-category disease on MALDI-TOF data. *Bioinformatics 2008, 24: 1812–1818.*

Ohata K, Hamasaki K, Toriyama K, Matsumoto K, Saeki A, Yanagi K, Abiru S, Nakagawa Y, Shigeno M, Miyazoe S, Ichikawa T, Ishikawa H, Nakao K, Eguchi K. Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection Cancer. *2003 Jun 15; 97(12):3036-43*.

Ono K, Tanaka T, Tsunoda T, Kitahara O, Kihara C, Okamoto A, Ochiai K, Takagi T, Nakamura Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res. 2000, 60, 5007-5011*

Okey AB. An aryl hydrocarbon receptor odyssey to the shores of toxicology: the Deichmann Lecture*, International Congress of Toxicology-XI.* Toxicol Sci. *2007; 98:5–38*.

Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO, Botstein D. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. U. S. A. 1999, 96, 9212-9217.115.*

Pessione F, Ramond MJ, Njapoum C, Duchatelle V, Degott C, Erlinger S, et al. Cigarette smoking and hepatic lesions in patients with chronic hepatitis C. *Hepatology. 2001;**34**:121–125*.

Poon TC, Wong N, Lai PB, Rattray M, Johnson PJ and Sung JJ. A tumor progression model for hepatocellular carcinoma: bioinformatic analysis of genomic data. *Gastroenterology 2006, 131: 1262–1270.*

Puga A, Ma C, Marlowe JL. The aryl hydrocarbon receptor cross-talks with multiple signal transduction pathways. *Biochem Pharmacol.* 2009;**77**:713–722

Pusztai L. Chips to bedside: incorporation of microarray data into clinical practice. *Clin Cancer Res 2006 Dec, VOL: 12(24), P:7209*.

Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet 2000, 24:227-235.*

Rosenberg W: Mechanisms of immune escape in viral hepatitis. *Gut 1999, 44: 759-764.*

Rui Xue, Jianying Li, Dennis J. Streveler. Microarray Gene Expression Profile Data Mining Model for Clinical Cancer Research *Proceedings of the 37th Hawaii International Conference on System Sciences 2004*

Safe S. Molecular biology of the Ah receptor and its role in carcinogenesis. *Toxicol Lett.* 2001; 120:1–7.

Shah JV, Cleveland DW. Waiting for anaphase: Mad2 and the spindle assembly checkpoint. *Cell. 2000;103:997–1000*.

Simon RM, Korn EL, McShane LM, Radmacher MD, Wright GW, Zhao Y.: Design and analysis of DNA microarray investigations. *New York.Springer,2003.*

Simon R. Challenges of microarray data and the evaluation of gene expression profile signatures. *Cancer Invest 2008 May, VOL:26(4), P:327*.

Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol 2005 Oct, VOL: 23(29), P:7332.*

Simon R., Lam A.P.: BRB-ArrayTools Users Guide (version 3.4). Bethesda, MD: *Biometric Research Branch National Cancer Institute, 2005. Technical Report 46, hhtp:/linus.nci.nih.gov/brb.*

Singh D, Febbo PG, Ross K, Jackson DG, Manola J, Ladd C, Tamayo P, Renshaw AA, D'Amico AV, Richie JP, Lander ES, Loda M, Kantoff PW, Golub TR, Sellers WR. Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell 2002, 1, 203-209.*

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A. 2001, 98, 10869-10874.*

Sun S, Lee NP, Poon RT, Fan ST, He QY, Lau GK and Luk JM. Oncoproteomics of hepatocellular carcinoma: from cancer markers'discovery to functional pathways. *Liver Int 2007, 27: 1021–1038.*

Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES, Golub TR. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U. S. A. 1999, 96, 2907-2912. 52*

Tellinghuisen TL, Rice CM. Interaction between hepatitis C virus proteins and host cell factors Curr Opin Microbiol. *2002 Aug;5(4):419-27*.

Theise ND, Park YN, Kojiro M: Dysplastic nodules and hepatocarcinogenesis. *Clin Liver Dis 2002, 6: 497-512*
Thorgeirsson SS, Grisham JW: Molecular pathogenesis of human hepatocellular carcinoma. *Nat Genet 2002, 31: 339-346.*

Thorgeirsson SS, Lee JS and Grisham JW. Functional genomics of hepatocellular carcinoma. *Hepatology 2006, 43: S145–150.*

Tsunedomi R, Iizuka N, Hamamoto Y, Uchimura S, Miyamoto T, Tamesa T, et al. Patterns of expression of cytochrome P450 genes in progression of hepatitis C virus-associated hepatocellular carcinoma. *Int J Oncol. 2005;27:661–667.*

Tsukuma H, Hiyama T, Tanaka S, Nakao M, Yabuuchi T, Kitamura T,Nakanishi K. Risk factors for hepatocellular carcinoma among patients with chronic liver disease. *N Engl J Med 1993;328:1797-1801*

Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A. 2001, 98, 5116-5121*

van't Veer L.J., Bernards R.: Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature 2008 Apr, VOL:452(7187),P:564.*

van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature 2002, 415, 530-536* .

Xu XQ, Leow CK, Lu X, Zhang X, Liu JS, Wong WH and Asperger A. Molecular classification of liver cirrhosis in a rat model by proteomics and bioinformatics. *Proteomics 2004, 4: 3235–3245*.

Wang X, Jin DY, Wong YC, Cheung AL, Chun AC, Lo AK, et al. Correlation of defective mitotic checkpoint with aberrantly reduced expression of MAD2 protein in nasopharyngeal carcinoma cells. *Carcinogenesis. 2000;21:2293–2297*

Zhang X, Lu X, Shi Q, Xu XQ, Leung HC, Harris LN and Iglehart JD. Recursive SVM feature selection and sample classification for mass-spectrometry and microarray data. *BMC Bioinformatics 2006, 7: 197*.

**ABBREVIATIONS**

| | |
|---|---|
| 2DE | 2D electrophoresis |
| AFB1 | Aflatoxin B1 |
| AFP | Alpha-fetoprotein |
| AHR | Aryl Hydrocarbon receptor |
| ANOVA | Analysis of Variance |
| aRNA | Amplified RNA |
| cDNA | Complementary DNA |
| CGH | Comparative genomic hybridization |
| CTR | Normal control liver samples |
| DAVID | Database for Annotation, Visualization and Integrated Discovery |
| DNAs | Synthetic complementary DNA |
| dsDNA | Double strand DNA |
| FDR | False Discovery Rate |
| HBsAg | Hepatitis B surface Antigen |
| HBV | Hepatitis B virus |
| HCC | Hepatocellular Carcinoma |
| HCV | Hepatitis C Virus |
| HCV-Ab | Hepatitis C Virus antibodies |
| IPA | Ingenuity Pathways Analysis |
| MHC | Major Histocompatibility Complex |
| mRNA | Messenger RNA |
| MS | Mass Spectrometry |
| PC | Prostate Cancer |
| PCA | Principal Component Analysis |
| PT | Permutation Test |
| SOM | Self-Organizing Maps |
| SVD | Singular Value Decomposition |

**ACKNOWLEDGEMENTS**

I thank Dr. Franco M. Buonaguro to allow me to take on this interesting and promising project and for his supervision and help during the course of this thesis. He encouraged me at all times and the present work is a result of many constructive discussions we have led over the last three years.

I thank Dr. Franco Marincola for giving me the opportunity to work in his laboratory in the Department of Transfusion Medicine, Clinical Center, NIH, Bethesda, MD for ten months as a visiting student. I also thank for his supervision and support additional to Dr. Franco M. Buonaguro during my time at NIH.

I thank Dr. Ena Wang from the Department of Transfusion Medicine, Clinical Center, NIH, Bethesda, MD for her support and help in teaching me how to perform microarray experiments and statistical analysis.

I wish to thank Prof. Piero Pucci, Prof. Aldo Vitagliano, Dr. Luigi Buonaguro, Dr. Maria Lina Tornesello. I could always count on their help and advice.

I thank Dr. Francesco Izzo for providing Human Liver Samples.

Last but not least, I would like to thank the most important and beloved people in my life because without them none of this would have been possible:
I thank my husband Diego, my parents and my brother with all my heart for their support and love.

Journal of Translational Medicine

BioMed Central

Research

Open Access

# Gene profiling, biomarkers and pathways characterizing HCV-related hepatocellular carcinoma

Valeria De Giorgi[1,2], Alessandro Monaco[3], Andrea Worchech[3,4,5], MariaLina Tornesello[1], Francesco Izzo[6], Luigi Buonaguro[1], Francesco M Marincola[3], Ena Wang[3] and Franco M Buonaguro*[1]

Address: 1Molecular Biology and Viral Oncogenesis & AIDS Refer. Center, Ist. Naz. Tumori "Fond. G. Pascale", Naples - Italy, 2Department of Chemistry, University of Naples "Federico II", Naples, Italy, 3Infectious Disease and Immunogenetics Section (IDIS), Department of Transfusion Medicine, Clinical Center and Trans-NIH Center for Human Immunology (CHI), National Institutes of Health, Bethesda, MD -USA, 4Genelux Corporation, Research and Development, San Diego Science Center, San Diego, CA, USA, 5Department of Biochemistry, Biocenter, University of Wuerzburg, Am Hubland, Wuerzburg, Germany and 6Div. of Surgery "D", Ist. Naz. Tumori "Fond. G. Pascale", Naples - Italy

Email: Valeria De Giorgi -valeriadegiorgi@tin.it; Alessandro Monaco -monacoal@cc.nih.gov; Andrea Worchech -worschecha@cc.nih.gov; MariaLina Tornesello -mltornesello@alice.it; Francesco Izzo -izzo@connect.it; Luigi Buonaguro -lbuonaguro@tin.it; Francesco M Marincola -FMarincola@mail.cc.nih.gov; Ena Wang -ewang@mail.cc.nih.gov; Franco M Buonaguro* -irccsvir@unina.it

* Corresponding author

Published: 12 October 2009 Received: 2 July 2009 Accepted: 12 October 2009

*Journal of Translational Medicine* 2009, **7**:85 doi:10.1186/1479-5876-7-85

This article is available from: http://www.translational-medicine.com/content/7/1/85



**Abstract Background:** Hepatitis C virus (HCV) infection is a major cause of hepatocellular carcinoma (HCC) worldwide. The molecular mechanisms of HCV-induced hepatocarcinogenesis are not yet fully elucidated. Besides indirect effects as tissue inflammation and regeneration, a more direct oncogenic activity of HCV can be postulated leading to an altered expression of cellular genes by early HCV viral proteins. In the present study, a comparison of gene expression patterns has been performed by microarray analysis on liver biopsies from HCV-positive HCC patients and HCV-negative controls.

**Methods:** Gene expression profiling of liver tissues has been performed using a high-density microarray containing 36'000 oligos, representing 90% of the human genes. Samples were obtained from 14 patients affected by HCV-related HCC and 7 HCV-negative non-liver-cancer patients, enrolled at INT in Naples. Transcriptional profiles identified in liver biopsies from HCC nodules and paired non-adjacent non-HCC liver tissue of the same HCV-positive patients were compared to those from HCV-negative controls by the Cluster program. The pathway analysis was performed using the BRB-Array- Tools based on the "Ingenuity System Database". Significance threshold of *t*-test was set at 0.001.

**Results:** Significant differences were found between the expression patterns of several genes falling into different metabolic and inflammation/immunity pathways in HCV-related HCC tissues as well as the non-HCC counterpart compared to normal liver tissues. Only few genes were found differentially expressed between HCV-related HCC tissues and paired non-HCC counterpart.

**Conclusion:** In this study, informative data on the global gene expression pattern of HCV-related HCC and non-HCC counterpart, as well as on their difference with the one observed in normal liver tissues have been obtained. These results may lead to the identification of specific biomarkers relevant to develop tools for detection, diagnosis, and classification of HCV-related HCC.

## Introduction

Hepatocellular carcinoma (HCC) is the most common liver malignancy as well as the third and the fifth cause of cancer death in the world in men and women, respectively [1-3]. As for other types of cancer, the etiology and pathogenesis of HCC is multifactorial and multistep [4]. The main risk factor for development of HCC are the hepatitis B and C virus (HBV and HCV) infection [5-8]. Non viral causes, such as toxins and drugs (i.e., alcohol, aflatoxins, microcystin, anabolic steroids), metabolic liver diseases (i.e., hereditary haemochromatosis, $\alpha$1-antitrypsin deficiency), steatosis and non-alcoholic fatty liver diseases as well as diabetes, play a role in a minor number of cases [9-11]. The prevalence of HCC in Italy, and in Southern Italy in particular, is significantly higher compared to other Western countries. Hepatitis virus infection, long-term alcohol and tobacco consumption account for 87% of HCC cases in Italian population and, among these, 61% of HCC are attributable to HCV. In particular, a recent seroprevalence surveillance study conducted in the general population of Southern Italy Campania Region reported a 7.5% positivity for HCV infection which peaked at 23.2% positivity in the 65 years or older age group [12]. The multistep progression to HCC, in particular the one associated to hepatitis virus, is characterized by a process including chronic liver injury, tissue inflammation, cell death, cirrhosis, regeneration, DNA damage, dysplasia and finally, HCC. In this multistep process, the cirrhosis represents the preneoplastic stage showing regenerative, dysplastic as well as HCC nodules [13].

The precise molecular mechanism underlying the progression of chronic hepatitis viral infections to HCC is currently unknown. Activation of cellular oncogenes, inactivation of tumor suppressor genes, overexpression of growth factors, telomerase activation and defects in DNA mismatch repair may contribute to the development of HCC [14-16]. In this framework, differential gene expression patterns accompanying different stages of growth, disease initiation, cell cycle progression, and responses to environmental stimuli provide important clues to this complex process.

DNA microarray enables investigators to study expression profile and activation of thousands of genes simultaneously. In particular, the identification of cancer-related stereotyped expression patterns might allow the elucidation of molecular mechanisms underlying cancer progression and provides important molecular markers for diagnostic purposes. This strategy has been recently used to profile global changes in gene expression in liver samples obtained from patients with HCV-related HCC [17-19]. Several of these studies identified gene sets that may be useful as potential microarray-based diagnostic tools. However, the direct or indirect HCV role in HCC pathogenesis is still a controversial issue and additional efforts need to be made aimed to specifically dissect the relationship between stages of HCV chronic infection and progression to HCC.

The present study has been focused on investigating genes and pathways involved in viral carcinogenesis and progression to HCC in HCV-chronically infected patients.

## Materials and methods

### *Patient and Tissue Samples*

Liver biopsies from fourteen HCV-positive HCC patients and seven HCV-negative non-liver cancer control patients (during laparoscopic cholecystectomy) were obtained with informed consent at the liver unit of the INT "Pascale" in Naples. In particular, from each of the HCV-positive HCC patients, a pair of liver biopsies from HCC nodule and non-adjacent non-HCC counterpart were surgically excised. All liver biopsies were stored in RNA Later at -80°C (Ambion, Austin, TX). Confirmation of the histopathological nature of the biopsies was performed by the Pathology lab at INT before the processing for RNA extraction. The non-HCC tissue from HCV-positive patient were an heterogeneous sample representing the prevalent liver condition of each subject (ranging from persistent HCV-infection to cirrhotic lesions). Furthermore, laboratory analysis confirmed that the 7 controls were seronegative for hepatitis C virus antibodies (HCV Ab).

### *Preparation of RNA, probe preparation, and microarray hybridization*

Samples were homogenized in disposable tissue grinders (Kendall, Precision). Total RNA was extracted by TRIzol solution (Life Technologies, Rockville, MD), and purity of the RNA preparation was verified by the 260:280 nm ratio (range, 1.8-2.0) at spectrophotometric reading with NanoDrop (Thermo Fisher Scientific, Waltham, MA). Integrity of extracted RNA was evaluated by Agilent 2100 Bioana-

lyzer (Agilent Technologies, Palo Alto, CA), analyzing the presence of 28S and 18S ribosomal RNA bands as well as the 28S/18S rRNA intensity ratio equal or close to 1.5. In addition, phenol contamination was checked and a 260:230 nm ratio (range, 2.0-2.2) was considered acceptable.

Double-stranded cDNA was prepared from 3 μg of total RNA (T-RNA) in 9 μl DEPC-treated $H_2O$ using the Super script II Kit (Invitrogen) with a T7-(dT15) oligonucleotide primer. cDNA synthesis was completed at 42°C for 1 h. Full-length dsDNA was synthesized incubating the produced cDNA with 2 U of RNase-H (Promega) and 3 μl of Advantage cDNA Polymerase Mix (Clontech), in Advantage PCR buffer (Clontech), in presence of 10 mM dNTP and DNase-free water. dsDNA was extracted with phenol-chloroform-isoamyl, precipitated with ethanol in the presence of 1 μl linear acrylamide (0.1 μg/μl, Ambion, Austin, TX) and aRNA (amplified-RNA) was synthesized using Ambion's T7 MegaScript in Vitro Transcription Kit (Ambion, Austin, TX). aRNA recovery and removal of template dsDNA was achieved by TRIzol purification. For the second round of amplification, aliquots of 1 μg of the aRNA were reverse transcribed into cDNA using 1 μl of random hexamer under the conditions used in the first round. Second-strand cDNA synthesis was initiated by 1 μg oligo-dT-T7 primer and the resulting dsDNA was used as template for in vitro transcription of aRNA in the same experimental conditions as for the first round [20]. 6 μg of this aRNA was used for probe preparation, in particular test samples were labeled with USL-Cy5 (Kreatech) and pooled with the same amount of reference sample (control donor peripheral blood mononuclear cells, PBMC, seronegative for hepatitis C virus antibodies (HCV Ab)) labeled with USL-Cy3 (Kreatech). The two labeled aRNA probes were separated from unincorporated nucleotides by filtration, fragmented, mixed and co-hybridized to a custom-made 36 K oligoarrays at 42°C for 24 h. The oligo-chips were printed at the Immunogenetics Section Department of Transfusion Medicine, Clinical Center, National Institutes of Health (Bethesda, MD). After hybridization the slides were washed with 2 × SSC/0.1%SDS for 1 min, 1 × SSC for 1 min, 0.2 × SSC for 1 min, 0.05 × SSC for 10 sec., and dried by centrifugation at 800 g for 3 minutes at RT.

### *Data Analysis*

Hybridized arrays were scanned at 10-μm resolution with a GenePix 4000 scanner (Axon Instruments) at variable photomultiplier tube (PMT) voltage to obtain maximal signal intensities with less than 1% probe saturation. Image and data files were deposited at microarray data base (mAdb) at http://nciarray.nci.nih.gov and retrieved after median centered, filtering of intensity (>200) and spot elimination (bad and no signal). Data were further analyzed using Cluster and TreeView software (Stanford University, Stanford, CA).

### *Statistical Analysis*

#### *Unsupervised Analysis*

For this analysis, a low-stringency filtering was applied, selecting the genes differentially expressed in 80% of all experiments with a >3 fold change ratio in at least one experiment. 7'760 genes were selected for the analysis including the three groups of analyzed samples (the HCV-related HCC, their non-HCC counterpart, as well as samples from the controls); 5'473 genes were selected for the analysis including the HCV-related HCC and normal control samples; 6'069 genes were selected for the analysis including the HCV-related non-HCC paired tissue and normal control samples. Hierarchical cluster analysis was conducted on these genes according to Eisen et al. [21]; differential expressed genes were visualized by Treeview and displayed according to the central method [22].

#### *Supervised Analysis*

Supervised class comparison was performed using the BRB ArrayTool developed at NCI, Biometric Research Branch, Division of Cancer Treatment and Diagnosis. Three subsets of genes were explored. The first subset included genes upregulated in HCV-related HCC compared to normal control samples, the second subset included genes upregulated in the HCV-related non-HCC counterpart compared with normal control samples, the third subset included genes upregulated in HCV-related HCC compared to the non-HCC paired liver tissue samples. Paired samples were analyzed using a two-tailed paired Student's *t*-test. Unpaired samples were tested with a two-tailed unpaired Student's *t*-test assuming unequal variance or with an *F*-test as appropriate. All analyses were tested for an univariate significance threshold set at a *p*-value < 0.01 for the first subset of genes and at a *p*-value < 0.001 for the second subset. Gene clusters identified by the univariate *t*-test were challenged with two alternative additional tests, an univariate permutation test (PT) and a global multivariate PT. The multivariate PT was calibrated to restrict the false discovery rate to 10%. Genes identified by univariate *t*-test as differentially expressed (*p*-value < 0.001 and *p*-value < 0.01) and a PT significance <0.05 were considered truly differentially expressed. Gene function was assigned based on Database for Annotation, Visualization and Integrated Discovery (DAVID) and Gene Ontology http://www.geneontology.org/.

#### *Ingenuity pathway analysis*

The pathway analysis was performed using the gene set expression comparison kit implemented in BRB-ArrayTools. The human pathway lists determined by "Ingenuity

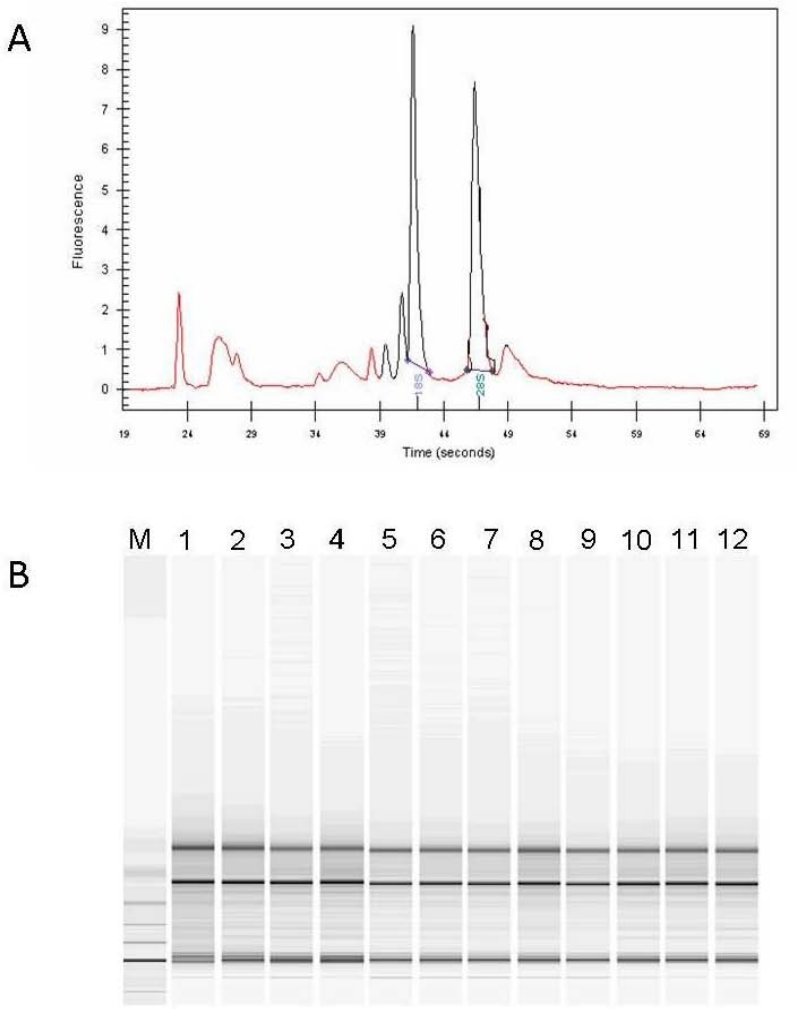


**Figure 1**
**Purity and integrity quality control of total extracted RNA**. (A) Representative Electropherogram of total RNA extracted from samples included in the analysis. (B) Representative Gel image evaluation of RNA integrity and 28S/18S rRNA ratio.

System Database" was selected. Significance threshold of *t*-test was set at 0.001. The Ingenuity Pathways Analysis (IPA) is a system that transforms large data sets into a group of relevant networks containing direct and indirect relationships between genes based on known interactions in the literature.

## Results

### *Quality Control*

The quality of extracted total RNA was verified by Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA), showing discrete 28S and 18S rRNA bands (Figure 1A) as well as a 28S/18S rRNA intensity ratio equal or close to 1.5 which is considered appropriate for total RNA extracted from liver tissue biopsies ("Assessing RNA Quality", http://www.ambion.com/techlib/tn/111/8.html). Based on this parameter, all extracted total RNA samples met the quality control criteria (Figure 1B).

### *Unsupervised analysis is concordant with Pathological Classification*

The gene expression profiles of tissue samples from the three groups of analyzed samples (the HCV-related HCC, their non-HCC counterpart, as well as samples from con-

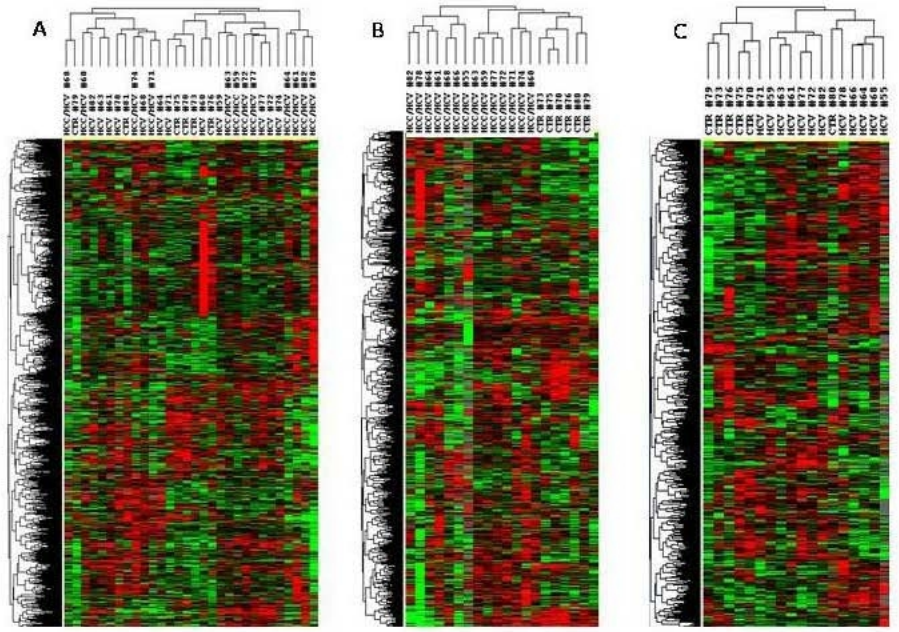


**Figure 2**

**Unsupervised hierarchical clustering**. Overall patterns of expression of genes across the 14 HCV-related HCC and non-HCC counterpart, as well as 7 HCV-negative control patients. Red indicates over-expression; green indicates under-expression; black indicates unchanged expression; gray indicates no detection of expression (intensity of both Cy3 and Cy5 below the cutoff value). Each row represents a single gene; each column represents a single sample. The dendrogram at the left of matrix indicates the degree of similarity among the genes examined by expression patterns. The dendrogram at the top of the matrix indicates the degree of similarity between samples. Panel A, unsupervised analysis including all three set of samples; Panel B, unsupervised analysis including HCV-related HCC and normal control liver samples; Panel C, unsupervised analysis including HCV-related non-HCC counterpart and normal control liver samples.

trol patients) were compared by an unsupervised analysis. No clear separation of the 3 different groups was observed, although control samples clustered mainly with samples from HCV-related non-HCC paired tissue, which includes dysplastic lesion in cirrhotic liver, representing a pre-neoplastic step (Figure 2A).

In order to identify genes differentially modulated in HCV-related lesions compared to normal liver tissue samples, an unsupervised analysis was then performed including only paired samples from HCV-related HCC and normal control samples or from the HCV-related non-HCC counterpart and control samples (Figures 2B and 2C). According to filtering described in Material and Methods, HCV-related HCC and normal control samples showed 5'473 genes differentially expressed, with a perfect clustering according to histological characteristics (Figure 2B). Similarly, HCV-related non-HCC tissue and normal control samples showed 6'069 genes differentially expressed with a perfect clustering according to histological characteristics also in this case (Figure 2C). The only exception to this pattern is represented by the normal control sample (CTR#80) which did not fall in the control cluster (CTR).

### *Supervised analysis*

The supervised analysis was performed comparing pairs of gene sets using an unpaired Student's *t*-test with a cut-off set at $p < 0.01$.

The analysis comparing gene sets in liver tissues from HCV-related HCC and normal controls identified 825

**Table 1: The first 40 up-regulated genes in HCV-related HCC**

| N° | Gene Name | Description |
|---|---|---|
| 1 | RYBP | RING1 and YY1 binding protein (RYBP) |
| 2 | ATP1B3 | ATPase, Na+/K+ transporting, beta 3 polypeptide |
| 3 | TMC | transmembrane channel-like 7 (TMC7) |
| 4 | ZNF567 | zinc finger protein 567 (ZNF567 |
| 5 | GPR108 | G protein-coupled receptor 108 (GPR108), transcript variant 1 |
| 6 | CD19 | CD19 molecule |
| 7 | SPINK1 | **serine peptidase inhibitor, Kazal type 1** |
| 8 | CDC2L6 | cell division cycle 2-like 6 (CDK8-like) |
| 9 | RSRC1 | arginine/serine-rich coiled-coil 1 (RSRC1) |
| 10 | METAP | methionyl aminopeptidase 1 |
| 11 | GPC3 | glypican 3 |
| 12 | SNHG11 | Small nucleolar RNA host gene (non-protein coding) 11 |
| 13 | RY1 | putative nucleic acid binding protein RY-1 (RY1) |
| 14 | CRELD2 | cysteine-rich with EGF-like domains 2 (CRELD2) |
| 15 | GLUL | glutamate-ammonia ligase (glutamine synthetase) |
| 16 | SERPINB1 | serpin peptidase inhibitor, clade B (ovalbumin), member 1 (SERPINB1) |
| 17 | TRMT6 | tRNA methyltransferase 6 homolog (S. cerevisiae) |
| 18 | UNC13D | unc-13 homolog D (C. elegans) (UNC13D) |
| 19 | E4F1--E4F | E4F transcription factor 1 (E4F1) |
| 20 | SLC22A2 | solute carrier family 22 (organic cation transporter), member 2 (SLC22A2) |
| 21 | CNIH4 | cornichon homolog 4 (Drosophila) (CNIH4) |
| 22 | TK1 | thymidine kinase 1, soluble (TK1) |
| 23 | MAFB | v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian) |
| 24 | PPP1CB | protein phosphatase 1, catalytic subunit, beta isoform (PPP1CB), transcript variant 3 |
| 25 | DNTTIP2 | deoxynucleotidyltransferase, terminal, interacting protein 2 (DNTTIP2) |
| 26 | ARID4B | AT rich interactive domain 4B (RBP1-like) (ARID4B), transcript variant 1 |
| 27 | SMARCC2 | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, |
| 28 | PRO1386 | PRO1386 protein |
| 29 | TRIOBP | TRIO and F-actin binding protein (TRIOBP), transcript variant 1 |
| 30 | VARS | valyl-tRNA synthetase |
| 31 | ITGA5 | integrin, alpha 5 (fibronectin receptor, alpha polypeptide) |
| 32 | TERF1 | telomeric repeat binding factor (NIMA-interacting) 1 (TERF1), transcript variant 2 |
| 33 | PURA | purine-rich element binding protein A (PURA) |
| 34 | TUBA1B | tubulin, alpha 1b |
| 35 | SNRPE | small nuclear ribonucleoprotein polypeptide E |
| 36 | RRAGD | Ras-related GTP binding D |
| 37 | VWF | von Willebrand factor |
| 39 | GLRX3 | glutaredoxin 3 (GLRX3) |
| 40 | ILF2 | interleukin enhancer binding factor 2, 45 kDa |

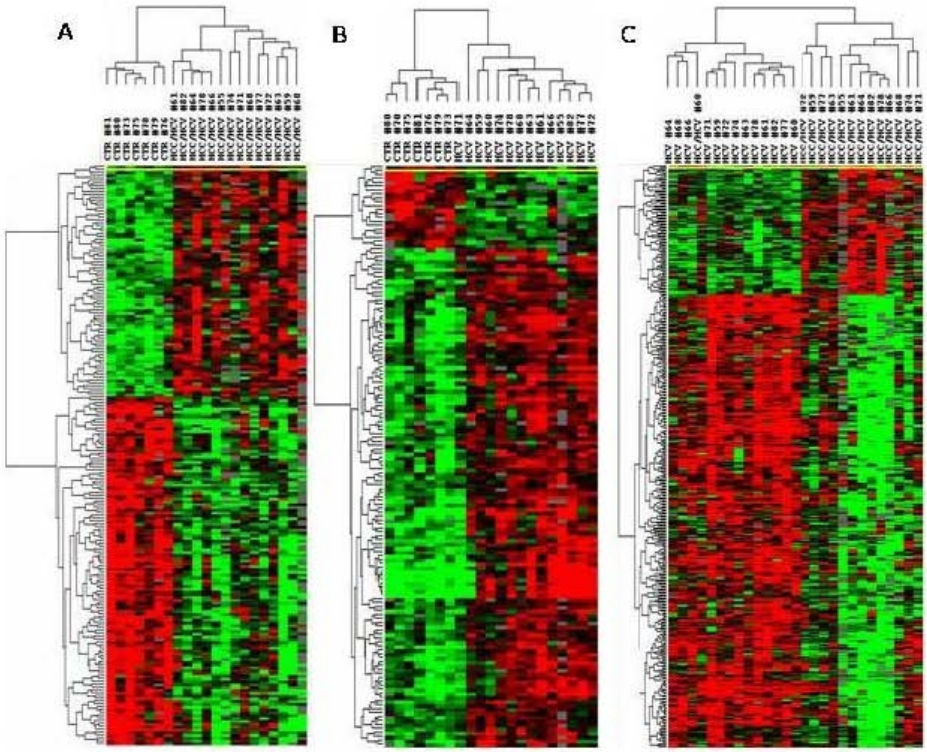


**Figure 3**
**Heat map of the gene signature, identified by Class Comparison Analysis**. Panel A, analysis including HCV-related HCC and normal control liver samples; Panel B, analysis including HCV-related non-HCC liver tissues and control liver samples; Panel C, analysis including HCV-related HCC and HCV-related non-HCC counterpart liver samples. The expression pattern of the genes is shown each row represents a single gene.

genes differentially expressed. Among them, 465 were shown to be up-regulated and 360 down-regulated in HCV-related HCC liver tissues (Figure 3A). The first 40 genes showing the highest fold of up-regulation are listed in Table 1.

The analysis comparing gene sets in liver tissues from HCV-related non-HCC tissue and controls identified 151 genes differentially expressed. Among them, 127 were shown to be up-regulated and 24 down-regulated in HCV-related non-HCC liver tissues (Figure 3B). The first 40 genes showing the highest fold of up-regulation are listed in Table 2.

The analysis comparing gene sets in liver tissues from HCV-related HCC and HCV-related non-HCC counterpart identified 383 genes differentially expressed. Among them, 83 were shown to be up-regulated and 300 down-regulated in HCV-related HCC liver tissues (Figure 3C). The first 40 genes showing the highest fold of up-regulation are listed in Table 3.

### *Ingenuity pathway analysis*

The pathway analysis was performed including the genes found up-regulated in the supervised comparisons, using the gene set expression comparison kit implemented in BRB-Array- Tools. The human pathway lists determined

**Table 2: The first 40 up-regulated genes in HCV-related non-HCC counterpart**

| N° | Gene Name | Description |
|---|---|---|
| 1 | NMNAT3 | nicotinamide nucleotide adenylyltransferase 3 (NMNAT3). |
| 2 | OASL | 2'-5'-oligoadenylate synthetase-like (OASL), transcript variant 2 |
| 3 | TMPRSS3 | transmembrane protease, serine 3 (TMPRSS3), transcript variant C |
| 4 | MFSD7 | major facilitator superfamily domain containing 7 (MFSD7) |
| 5 | AEBP1 | AE binding protein 1 (AEBP1), mRNA. |
| 6 | UBD | ubiquitin D (UBD) |
| 7 | S100A4 | S100 calcium binding protein A4 (S100A4), transcript variant 1 |
| 8 | C1orf151 | chromosome 1 open reading frame 151 (C1orf151) |
| 9 | CRIP1 | Cysteine-rich protein 1 (intestinal) |
| 10 | ASCC3 | activating signal cointegrator 1 complex subunit 3 |
| 11 | ZNF271 | zinc finger protein 271 (ZNF271), transcript variant 2 |
| 12 | ANXA4 | annexin A4 (ANXA4) |
| 13 | NMI | N-myc (and STAT) interactor (NMI) |
| 14 | UBE2L6 | ubiquitin-conjugating enzyme E2L 6 (UBE2L6), transcript variant 1 |
| 15 | B2 M | beta-2-microglobulin (B2 M) |
| 16 | HLA-F | Major histocompatibility complex, class I, F |
| 17 | PSMB9 | Proteasome (prosome, macropain) subunit, beta type, 9 |
| 18 | TAP1 | transporter 1, ATP-binding cassette, sub-family B (MDR/TAP) |
| 19 | PSME2 | proteasome (prosome, macropain) activator subunit 2 (PA28 beta) |
| 20 | IFI16 | interferon, gamma-inducible protein 16 |
| 21 | IFI27 | interferon, alpha-inducible protein 27 |
| 22 | ARHGAP9 | Rho GTPase activating protein 9 |
| 23 | RABGAP1L | RAB GTPase activating protein 1-like |
| 24 | TNK1 | tyrosine kinase, non-receptor |
| 25 | DEF6 | differentially expressed in FDCP 6 homolog (mouse) |
| 26 | BTN3A3 | butyrophilin, subfamily 3, member A3 |
| 27 | RPS6KA1 | ribosomal protein S6 kinase, 90 kDa, polypeptide 1 |
| 28 | CD24 | CD24 molecule |
| 29 | PARP10 | poly (ADP-ribose) polymerase family, member 10 |
| 30 | APOL3 | apolipoprotein L, 3 (APOL3), transcript variant alpha/d |
| 31 | STAT | signal transducer and activator of transcription 1, 91 kDa |
| 32 | ANKRD10 | Ankyrin repeat domain 10 |
| 33 | CKB | creatine kinase, brain (CKB) |
| 34 | H2AFZ | H2A histone family, member Z |
| 35 | PSMB9 | proteasome (prosome, macropain) subunit, beta type, 9 |
| 36 | RARRES3 | retinoic acid receptor responder (tazarotene induced) 3 |
| 37 | RGS10 | regulator of G-protein signaling 10 (RGS10), transcript variant 2 |
| 38 | TUBB | tubulin, beta |
| 39 | NOL3 | nucleolar protein 3 (apoptosis repressor with CARD domain) |
| 40 | CD7 | CD74 molecule, major histocompatibility complex, class II invariant chain |

by "Ingenuity System Database" was selected. Significance threshold of *t*-test was set at 0.001. Samples from HCV-related non-HCC liver tissue showed strong up-regulation of genes involved in Antigen Presentation, Protein Ubiquitination, Interferon signaling, IL-4 signaling, Bacteria and Viruses cell cycle and chemokine signaling pathways. Samples from HCV-related HCC showed strong up-regulation of genes involved in Metabolism, Aryl Hydrocarbon receptor signaling, 14-3-3 mediated signaling and protein Ubiquitination pathways. Significant pathways were listed respectively in Figures 4, 5, 6 and 7.

## Discussion

The pathogenetic mechanisms leading to HCC development in HCV chronic infection are not yet fully elucidated. In particular, besides inducing liver tissue inflammation and regeneration, which ultimately may result in cellular transformation and HCC development, HCV may play a more direct oncogenic activity inducing an altered expression of cellular genes. To this aim, global gene expression profile can identify specific genes differentially expressed and provide powerful insights into mechanisms regulating the transition from pre-neoplastic to fully blown neoplastic proliferation [23,24].

**Table 3: The first 40 up-regulated genes in HCV-related HCC**

| N° | Gene Name | Description |
|---|---|---|
| 1 | CAPG | capping protein (actin filament), gelsolin-like |
| 2 | OCC-1 | PREDICTED: misc_RNA (OCC-1) |
| 3 | EED | embryonic ectoderm development (EED), transcript variant 1 |
| 4 | RPLP0 | ribosomal protein, large, P0 (RPLP0), transcript variant 1 |
| 5 | RPLP0P2 | ribosomal protein, large, P0 pseudogene 2 |
| 6 | AP1S2 | adaptor-related protein complex 1, sigma 2 subunit |
| 7 | RRAGD | Ras-related GTP binding D (RRAGD) |
| 8 | PFDN4 | prefoldin subunit 4 (PFDN4) |
| 9 | CCDC104 | coiled-coil domain containing 104 (CCDC104) |
| 10 | C7orf28B | chromosome 7 open reading frame 28B |
| 11 | PSIP1 | PC4 and SFRS1 interacting protein 1 (PSIP1), transcript variant 2. |
| 12 | LPCAT1 | lysophosphatidylcholine acyltransferase 1 |
| 13 | FSCN3 | fascin homolog 3, actin-bundling protein, testicular |
| 14 | RAB24 | RAB24, member RAS oncogene family |
| 15 | ZNF446 | zinc finger protein 446 (ZNF446) |
| 16 | SEC11B | PREDICTED: SEC11 homolog B (S. cerevisiae) |
| 17 | ZNF586 | zinc finger protein 586 (ZNF586) |
| 18 | SCNM1 | sodium channel modifier 1 |
| 19 | SF3A1 | splicing factor 3a, subunit 1, 120 kDa |
| 20 | RUFY1 | RUN and FYVE domain containing 1 |
| 21 | TRIM55 | tripartite motif-containing 55 |
| 22 | GOLGA4 | golgi autoantigen, golgin subfamily a |
| 23 | GPATCH4 | G patch domain containing 4 (GPATCH4), transcript variant 1 |
| 24 | THOP1 | thimet oligopeptidase 1 |
| 25 | TUBB2C | tubulin, beta 2C (TUBB2C) |
| 26 | PHLDB3 | Pleckstrin homology-like domain, family B |
| 27 | FAM104A | family with sequence similarity 104, member A |
| 28 | FASTK | Fas-activated serine/threonine kinase |
| 29 | EIF2AK4 | eukaryotic translation initiation factor 2 alpha kinase 4 |
| 30 | ZFP41 | ZFP41--zinc finger protein 41 homolog (mouse) |
| 31 | PRKRIP1 | PRKR interacting protein 1 (IL11 inducible) |
| 32 | DSTN | destrin (actin depolymerizing factor) |
| 33 | PHIP | pleckstrin homology domain interacting protein (PHIP) |
| 34 | NUCKS1 | nuclear casein kinase and cyclin-dependent kinase substrate 1 |
| 35 | TNRC8 | Trinucleotide repeat containing 8 |
| 36 | CCDC132 | coiled-coil domain containing 132 |
| 37 | EPRS | glutamyl-prolyl-tRNA synthetase |
| 39 | HIST1H4C | histone cluster 1, H4c |
| 40 | CDCA8 | cell division cycle associated 8 |

In the present study, the differential gene expression was evaluated by microarray analysis on liver tissues obtained from fourteen HCV-positive HCC patients and seven HCV-negative control patients. In particular, from each of the HCV-positive HCC patients, a pair of liver biopsies from HCC nodule and non-HCC non adjacent counterpart were surgically excised.

The unsupervised analysis didn't show a clear separation of samples from the 3 different groups (HCV-related HCC, their non-HCC counterpart, as well as control patients), suggesting the lack of a clear-cut distinct gene signature pattern. Nevertheless, normal control samples, with the exception of CTR#76 sample, grouped in a single cluster close to samples from HCV-related paired non-HCC samples. The latter, in fact, comprise several non-HCC pathological stages including dysplastic, not fully transformed lesions, representing pre-neoplastic step in the progression to HCC and should still retain a gene signature pattern closer to normal than to transformed cell physiology. On the contrary, the unsupervised analysis including only one of the HCV-related liver tissues (HCC or non-HCC counterpart) and normal controls showed a clear-cut segregation of the pathological from the control cluster, indicating the identification of specific gene signature patterns peculiar to the HCV-related pre-neoplastic (non-HCC) and neoplastic (HCC) tissues compared to normal controls.

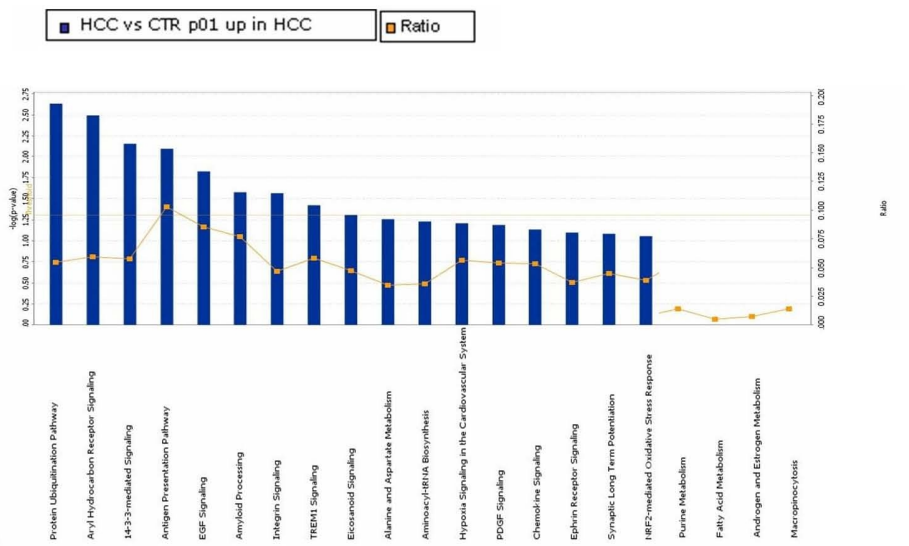


**Figure 4**
**Significant pathways at the nominal 0.01 level of the unpaired Student's *t*-test**. The human pathway lists determined by "Ingenuity System Database" in HCV-related HCC samples.

A supervised analysis was performed by pairwise comparison between samples of the three groups analyzed in the present study. The results indicated that the HCV-related HCC liver tissues showed 825 genes differentially expressed compared to controls, of which 465 were up-regulated and 360 down-regulated. The HCV-related non-HCC liver tissues showed 151 genes differentially expressed compared to controls, of which 127 were up-regulated and 24 down-regulated. The HCV-related HCC liver tissues showed 383 genes differentially expressed compared to HCV-related non-HCC counterpart, of which 83 were up-regulated and 300 down-regulated. In each of these independent class comparison analysis, the differentially expressed genes were selected based on a 3-fold difference at a significance $p$-value < 0.01.

The up-regulated genes identified within the individual class comparison analysis were further evaluated and classified by a pathway analysis, according to the "Ingenuity System Database".

The genes up-regulated in samples from HCV-related HCC are classified in metabolic pathways, and the most represented are the Aryl Hydrocarbon receptor signaling (AHR) and, protein Ubiquitination pathways, which have been previously reported to be involved in cancer, and in particular in HCC, progression.

The Aryl Hydrocarbon receptor signal transduction Pathway (AHR) is involved in the activation of the cytosolic aryl hydrocarbon receptor by structurally diverse xenobiotic ligands (including dioxin, and polycyclic or halogenated aromatic hydrocarbons) and mediating their toxic and carcinogenic effects [25,26]. More recently AHR pathway has been shown to be involved in apoptosis, cell cycle regulation, mitogen-activated protein kinase cascades [27]. In particular, studies on liver tumor promotion have shown that dioxin-induced AHR activation mediates clonal expansion of initiated cells by inhibiting apoptosis and bypassing AHR-dependent cell cycle arrest [28]. Furthermore, it has been shown that changes in mRNA expression of specific genes in the AHR pathway are linked to progression of HCV-associated hepatocellular carcinoma [29]. Moreover, the HCV-induced AHR signal transduction pathway, could be directly involved in the

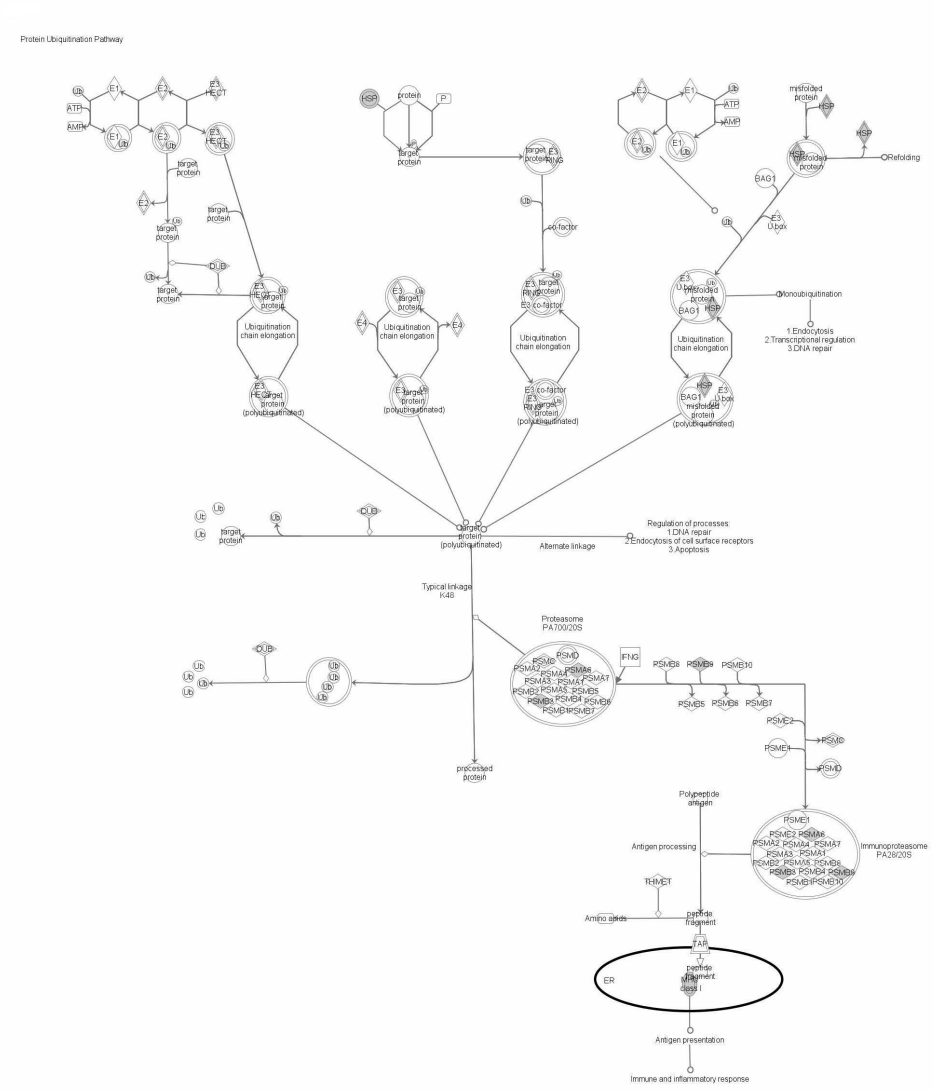


**Figure 5**
**Significant pathways at the nominal 0.01 level of the unpaired Student's *t*-test**. The 1 top-scoring pathway of genes upregulated IPA image.

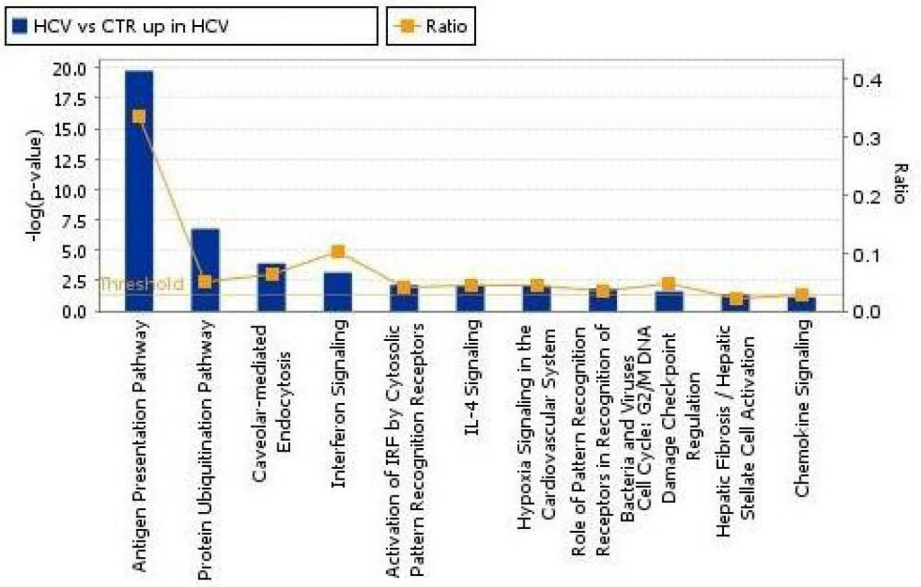


**Figure 6**
**Significant pathways at the nominal 0.001 level of the unpaired Student's *t*-test**. The human pathway lists determined by "Ingenuity System Database" in HCV-related non-HCC samples.

increased severity of hepatic lesions in patients with chronic hepatitis C induced by smoking [30,31].

The ubiquitin and ubiquitin-related proteins of the ubiquitination pathway play instrumental roles in cell-cycle regulation [32] as well as cell death/apoptosis [33] through modification of target proteins. In particular, ubiquitin-like proteins, i.e. FAT10, has been reported to bind non-covalently to the human spindle assembly checkpoint protein, MAD2 [34], which is responsible for maintaining spindle integrity during mitosis [35] and whose inhibited function has been associated with chromosomal instability [36,37]. Moreover, FAT10 overexpression has been previously shown in hepatocellular carcinoma [38].

The genes up-regulated in samples from HCV-related non-HCC tissue are classified in several pathways prevalently associated to inflammation and native/adaptive immunity and most of the overexpressed genes belong to the Antigen Presentation pathway. Considering the chronic HCV infection, these result could be unexpected and contradictory, since a reduced native and/or adaptive specific immune response would represent a very much favorable environment for the virus. Nevertheless, these findings, which confirm also a recent report by others [39], could explain the generic massive inflammation and immunopathological tissue damage characteristic of HCV-related cirrhosis [40].

In this study, informative data on the global gene expression pattern in HCV-related HCC as well as HCV-related non-HCC counterpart liver tissues have been obtained compared to normal controls. These data, which need further confirmation studies on a larger set of samples and also at protein level, may be extremely helpful for the identification of exclusive activation markers to characterize gene expression programs associated with progression of HCV-related lesions to HCC.

## Competing interests
The authors declare that they have no competing interests.

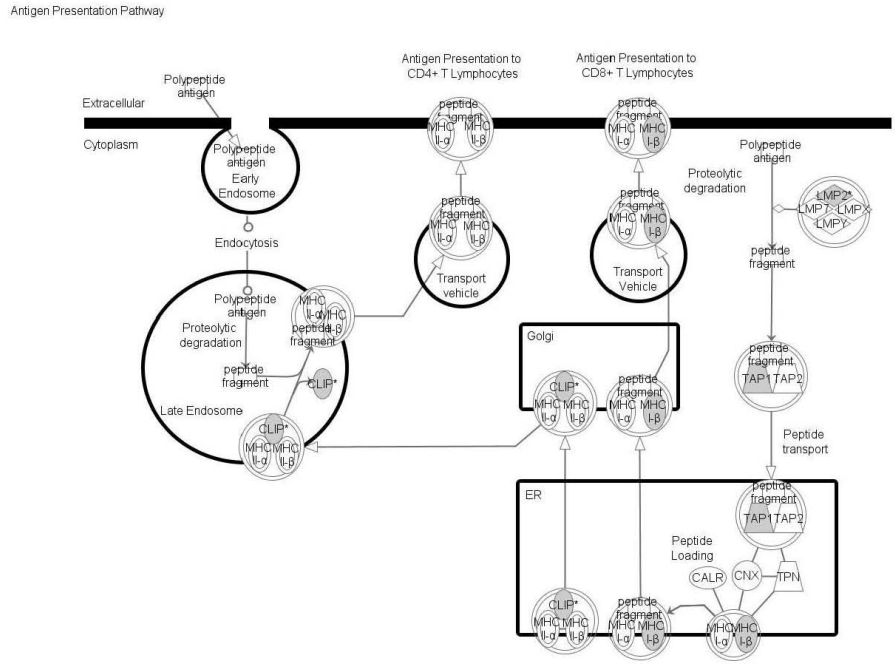


**Figure 7**
**Significant pathways at the nominal 0.001 level of the unpaired Student's *t*-test.** The 1 top-scoring pathway of genes upregulated IPA image.

## Authors' contributions

FMB, FI, MLT and FMM were responsible for the overall planning and coordination of the study. AW and LB were involved in the data analysis; VDG and EW were involved in genetic analyses. FI was involved in the patients enrollment and liver sample collection. VDG and AM were responsible for specimen processing and RNA analysis. VDG and FMB compiled and finalized the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We are indebted to Dr. Marianna Sabatino for her invaluable technical support and fruitful discussions. This study was supported by grants from the Italian Ministry of Health - Ministero Italiano Salute (Ricerca Corrente 2008-9 and FSN 2005 Cnv 89).

## References

1. El-Serag HB, Mason AC: **Rising incidence of hepatocellular carcinoma in the United States.** *N Engl J Med* 1999, **340:**745-750.
2. Davila JA, Petersena NJ, Nelson HA, El-Serag HB: **Geographic variation within the United States in the incidence of hepatocellular carcinoma.** *J Clin Epidemiol* 2003, **56:**487-493.
3. El-Serag HB: **Hepatocellular carcinoma and hepatitis C in the United States.** *Hepatology* 2002, **36:**S74-S83.
4. Romeo R, Colombo M: **The natural history of hepatocellular carcinoma.** *Toxicology* 2002, **181-182:**39-42.
5. Block TM, Mehta AS, Fimmel CJ, Jordan R: **Molecular viral oncology of hepatocellular carcinoma.** *Oncogene* 2003, **22:**5093-5107.
6. Buendia MA: **Hepatitis B viruses and cancerogenesis.** *Biomed Pharmacother* 1998, **52:**34-43.
7. Davis GL, Albright JE, Cook SF, Rosenberg DM: **Projecting future complications of chronic hepatitis C in the United States.** *Liver Transpl* 2003, **9:**331-338.
8. Colombo M: **The role of hepatitis C virus in hepatocellular carcinoma.** *Recent Results Cancer Res* 1998, **154:**337-344.
9. Ohata K, Hamasaki K, Toriyama K, Matsumoto K, Saeki A, Yanagi K, *et al.*: **Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection.** *Cancer* 2003, **97:**3036-3043.
10. Brunt EM: **Nonalcoholic steatohepatitis.** *Semin Liver Dis* 2004, **24:**3-20.
11. Davila JA, Morgan RO, Shaib Y, McGlynn KA, El-Serag HB: **Diabetes increases the risk of hepatocellular carcinoma in the United**

**States: a population based case control study.** *Gut* 2005, **54**:533-539.
12. Fusco M, Girardi E, Piselli P, Palombino R, Polesel J, Maione C, *et al.*: **Epidemiology of viral hepatitis infections in an area of southern Italy with high incidence rates of liver cancer.** *Eur J Cancer* 2008, **44**:847-853.
13. Schafer DF, Sorrell MF: **Hepatocellular carcinoma.** *Lancet* 1999, **353**:1253-1257.
14. Levrero M: **Viral hepatitis and liver cancer: the case of hepatitis C.** *Oncogene* 2006, **25**:3834-3847.
15. Macdonald GA, Greenson JK, Saito K, Cherian SP, Appelman HD, Boland CR: **Microsatellite instability and loss of heterozygosity at DNA mismatch repair gene loci occurs during hepatic carcinogenesis.** *Hepatology* 1998, **28**:90-97.
16. Blum HE, Moradpour D: **Viral pathogenesis of hepatocellular carcinoma.** *J Gastroenterol Hepatol* 2002, **17(Suppl 3)**:S413-S420.
17. Iizuka N, Oka M, Yamada-Okabe H, Mori N, Tamesa T, Okada T, *et al.*: **Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method.** *Cancer Res* 2002, **62**:3939-3944.
18. Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, *et al.*: **Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression.** *Cancer Res* 2001, **61**:2129-2137.
19. Shirota Y, Kaneko S, Honda M, Kawai HF, Kobayashi K: **Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays.** *Hepatology* 2001, **33**:832-840.
20. Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM: **High-fidelity mRNA amplification for gene profiling.** *Nat Biotechnol* 2000, **18**:457-459.
21. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci USA* 1998, **95**:14863-14868.
22. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, *et al.*: **Systematic variation in gene expression patterns in human cancer cell lines.** *Nat Genet* 2000, **24**:227-235.
23. Nacht M, Ferguson AT, Zhang W, Petroziello JM, Cook BP, Gao YH, *et al.*: **Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer.** *Cancer Res* 1999, **59**:5464-5470.
24. Sanchez-Carbayo M, Socci ND, Lozano JJ, Li W, Charytonowicz E, Belbin TJ, *et al.*: **Gene discovery in bladder cancer progression using cDNA microarrays.** *Am J Pathol* 2003, **163**:505-516.
25. Safe S: **Molecular biology of the Ah receptor and its role in carcinogenesis.** *Toxicol Lett* 2001, **120**:1-7.
26. Okey AB: **An aryl hydrocarbon receptor odyssey to the shores of toxicology: the Deichmann Lecture, International Congress of Toxicology-XI.** *Toxicol Sci* 2007, **98**:5-38.
27. Puga A, Ma C, Marlowe JL: **The aryl hydrocarbon receptor cross-talks with multiple signal transduction pathways.** *Biochem Pharmacol* 2009, **77**:713-722.
28. Bock KW, Kohle C: **Ah receptor- and TCDD-mediated liver tumor promotion: clonal selection and expansion of cells evading growth arrest and apoptosis.** *Biochem Pharmacol* 2005, **69**:1403-1408.
29. Tsunedomi R, Iizuka N, Hamamoto Y, Uchimura S, Miyamoto T, Tamesa T, *et al.*: **Patterns of expression of cytochrome P450 genes in progression of hepatitis C virus-associated hepatocellular carcinoma.** *Int J Oncol* 2005, **27**:661-667.
30. Pessione F, Ramond MJ, Njapoum C, Duchatelle V, Degott C, Erlinger S, *et al.*: **Cigarette smoking and hepatic lesions in patients with chronic hepatitis C.** *Hepatology* 2001, **34**:121-125.
31. Hezode C, Lonjon I, Roudot-Thoraval F, Mavier JP, Pawlotsky JM, Zafrani ES, *et al.*: **Impact of smoking on histological liver lesions in chronic hepatitis C.** *Gut* 2003, **52**:126-129.
32. Jentsch S, Pyrowolakis G: **Ubiquitin and its kin: how close are the family ties?** *Trends Cell Biol* 2000, **10**:335-342.
33. Jesenberger V, Jentsch S: **Deadly encounter: ubiquitin meets apoptosis.** *Nat Rev Mol Cell Biol* 2002, **3**:112-121.
34. Liu YC, Pan J, Zhang C, Fan W, Collinge M, Bender JR, *et al.*: **A MHC-encoded ubiquitin-like protein (FAT10) binds noncovalently to the spindle assembly checkpoint protein MAD2.** *Proc Natl Acad Sci USA* 1999, **96**:4313-4318.
35. Shah JV, Cleveland DW: **Waiting for anaphase: Mad2 and the spindle assembly checkpoint.** *Cell* 2000, **103**:997-1000.
36. Wang X, Jin DY, Wong YC, Cheung AL, Chun AC, Lo AK, *et al.*: **Correlation of defective mitotic checkpoint with aberrantly reduced expression of MAD2 protein in nasopharyngeal carcinoma cells.** *Carcinogenesis* 2000, **21**:2293-2297.
37. Gemma A, Hosoya Y, Seike M, Uematsu K, Kurimoto F, Hibino S, *et al.*: **Genomic structure of the human MAD2 gene and mutation analysis in human lung and breast cancers.** *Lung Cancer* 2001, **32**:289-295.
38. Lee CG, Ren J, Cheong IS, Ban KH, Ooi LL, Yong TS, *et al.*: **Expression of the FAT10 gene is highly upregulated in hepatocellular carcinoma and other gastrointestinal and gynecological cancers.** *Oncogene* 2003, **22**:2592-2603.
39. Mas VR, Maluf DG, Archer KJ, Yanek K, Kong X, Kulik L, *et al.*: **Genes involved in viral carcinogenesis and tumor initiation in hepatitis C virus-induced hepatocellular carcinoma.** *Mol Med* 2009, **15**:85-94.
40. Schuppan D, Krebs A, Bauer M, Hahn EG: **Hepatitis C and liver fibrosis.** *Cell Death Differ* 2003, **10(Suppl 1)**:S59-S67.

# Molecular Signature Associated To Human Liver Cancer

**V. De Giorgi[1,2], L. Buonaguro[1], A. Worschech[3,4,5], M.L. Tornesello[1], F. Izzo[6], , F.M. Marincola[3], E. Wang[3] and F.M. Buonaguro[1]**

[1] Molecular Biology and Viral Oncogenesis & AIDS Refer. Center, Ist. Naz. Tumori "Fond. G. Pascale", Naples – Italy

[2]Department of Chemistry, University of Naples "Federico II", Naples, Italy

[3]Infectious Disease and Immunogenetics Section (IDIS), Department of Transfusion Medicine, Clinical Center and Trans-NIH Center for Human Immunology (CHI), National Institutes of Health, Bethesda ,MD -USA;

[4] Genelux Corporation, Research and Development, San Diego Science Center, San Diego, CA, USA.

[5] Department of Biochemistry, Biocenter, University of Wuerzburg, Am Hubland, Wuerzburg, Germany

[6] Div. of Surgery "D", Ist. Naz. Tumori "Fond. G. Pascale", Naples – Italy;.

*Corresponding author:
F.M. Buonaguro, M.D.
Director
Unit of Molecular Biology and Viral Oncogenesis & AIDS Reference Center
Istituto Nazionale Tumori "Fond. G. Pascale"
Via Mariano Semmola, 142
80131 NAPLES - ITALY
Tel. +39-081-5903.830
Fax +39-081-5451276
E-mail: irccsvir@unina.it

## ABSTRACT

Hepatocellular carcinomas (HCCs) are a heterogeneous group of tumors that differ in risk factors and genetic alterations. In Italy, , and Southern Italy in particular HCV infection is the main etiological cause of this cancer and one of the more relevant driving factors. Using nanotechnology-based high-density oligoarrays to characterize patterns of gene expression in HCCs, consistent differences between gene-expression patterns in HCC and normal liver tissue have been identified. Expression patterns in HCC were also readily distinguished from those associated with liver metastasis. To characterize the molecular events of the hepatocarcinogenic process and to identify new biomarkers for early HCC diagnosis, gene expression profiles of 71 tissue biopsies sampled from HCV-related primary HCC and their corresponding HCV-positive non-HCC counterpart, as well as gastrointestinal liver metastasis paired with the apparently normal peritumoral tissue, were analyzed and compared with gene profiles of 6 liver biopsies from control patients.

A set of genes sufficient for the molecular signature and able to complement or substitute conventional pathology has been identified. In addition using a time course analysis , several molecular markers for early HCC diagnosis have been recognized. These findings provide a comprehensive molecular portrait of genomic changes in HCC progression.

JOURNAL OF CLINICAL MICROBIOLOGY, June 2010, p. 2298–2300 Vol. 48, No. 6
0095-1137/10/$12.00 doi:10.1128/JCM.00418-10


# Cardiac Tamponade and Heart Failure Due to Myopericarditis as a Presentation of Infection with the Pandemic H1N1 2009 Influenza A Virus▽

Simona Puzelli,[1]* Franco M. Buonaguro,[2] Marzia Facchini,[1] Annapina Palmieri,[1] Laura Calzoletti,[1] Maria A. De Marco,[1] Pasquale Arace,[3] Enrico de Campora,[3] Ciro Esposito,[4] Antonio Cassone,[1] Giovanni Rezza,[1] Isabella Donatelli,[1] the Surveillance Group for Pandemic H1N1 2009 Influenza Virus in Italy,† and the Campania H1N1 Task Force‡

*Istituto Superiore di Sanità (ISS), National Institute of Health, Rome, Italy[1]; Fondazione Pascale Institute, Naples, Italy[2]; Santobono-Pausilipon Hospital, Naples, Italy[3]; and Cotugno Hospital, Naples, Italy[4]*

Received 1 March 2010/Returned for modification 12 March 2010/Accepted 2 April 2010

**We describe a fatal case of myopericarditis presenting with cardiac tamponade in a previously healthy 11-year-old child. Pandemic H1N1 2009 influenza A virus sequences were identified in throat and myocardial tissues and pericardial fluid, suggesting damage of myocardial cells directly caused by the virus.**

## CASE REPORT

Here, we report a fatal case of myopericarditis presenting with cardiac tamponade associated with infection with the pandemic A/H1N1 2009 influenza virus in a previously healthy 11-year-old girl with no known risk factors, including obesity, for severe, complicated pandemic influenza.

The patient developed a high fever (39.2°C) with coughing and vomiting on the evening of 28 October. The next morning, she was visited by a general practitioner, who prescribed antipyretic, antibiotic, and cortisone treatments. About 36 h after the onset of symptoms, the girl presented to the emergency unit of the district hospital with asthenia and dizziness. She was pale and afebrile at 35°C, with tachycardia of 140 beats/min and tachypnea of 32 breaths/min. Routine blood samples were analyzed. The white blood cell count was elevated at $23.8 \times 10^9$ cells/liter, and C-reactive protein was elevated at 21.6 mg/liter. Her creatine kinase (CK) level was elevated at 5,814 IU/liter (normal range, 26 to 140 IU/liter), the troponin level was 4.24 ng/ml (normal range, 0 to 0.060 ng/ml), the myoglobin level was 1,849.4 ng/ml (normal range, 12 to 76 ng/ml), and the lactate dehydrogenase level was 1,408 IU/liter (normal range, 266 to 500 IU/liter). Her chest X-ray was reported to be clear. An electrocardiogram (ECG) showed sinus tachycardia with diffuse ST segment elevation. An urgent echocardiogram was performed, which demonstrated normal ventricle size, severely compromised pump function (ejection fraction, 20 to 30%), diffuse ipokinesia, and a small pericardial effusion, consistent with myopericarditis with a predominant myocarditic component. Metabolic lactic acidosis was also detected and treated with intravenous administration of $Na^+$-$HCO_3$.

Because of the need for specialized care and treatment, the patient was transferred to a regional pediatric hospital, but during ambulance transport she had loss of consciousness and needed respiratory assistance. Clinical conditions worsened during transfer to the pediatric intensive care unit (ICU): a peripheral pulse was absent, the central pulse was weak, and the capillary refill time was >5 s. Cardiac frequency (CF) detected by the monitor was 110 beats/min with a wide QRS interval. CF rapidly decreased. Atropine was administered, and cardiopulmonary resuscitation (CPR) maneuvers were performed. The central pulse became absent, and the echocardiogram showed no contractile activity.

The pathological exam showed macroscopic changes such as increased ventricular size (left ventricle thickness, 2.2 cm; right ventricle thickness, 0.35 cm); the left/right ventricle size ratio was 6.28, versus a normal value of 3.1. Pericardial effusion (150 ml) was detected. Pulmonary congestion leading to pneumonic hepatization was also found.

Microscopic changes revealed mild inflammation: modest infiltration of histiocytes (CD68 positive) and myocellular necrosis were detected (Fig. 1 and 2).

PCR analysis of a nasopharyngeal swab collected at the ICU was positive for the pandemic H1N1 2009 influenza virus. The sample was grown in MDCK-SIAT1 cells (hemagglutinating units [HAU], 64). Molecular analyses of several samples of tissues and fluids collected during the autopsy were performed. The pandemic influenza virus was detected in bronchi, in the myocardial tissue, and in both the pericardium and pericardial fluid, whereas it was not found in the lungs or in other tissues or fluids, such as the pharynx, tonsils, and spleen (Table 1). The virus was isolated from myocardial tissue but not from

* Corresponding author. Mailing address: National Influenza Centre, Department of Infectious Diseases—Istituto Superiore di Sanità, Viale Regina Elena, 299-00161 Rome, Italy. Phone: (39) 06 49903243. Fax: (39) 06 49902082. E-mail: simona.puzelli@iss.it.

† Contributing members of the Surveillance Group for Pandemic H1N1 2009 Influenza Virus in Italy, ISS, are Gabriele Vaccari, Domenico Spagnolo, and Concetta Fabiani.

‡ Contributing members of the Campania H1N1 Task Force are Mario Santangelo (Regione Campania Health Affairs); Gaetano De Rosa and Claudio Buccelli (Federico II University, Naples); Valeria De Giorgi and Renato Franco (Fondazione Pascale Institute, Naples); Oreste Perrella (Cotugno Hospital, Naples); Fabiana Cerrato, Anna Dolcini, Rodolfo Paladini, Raffaele Testa, and Maria Erennia Vitullo (Santobono-Pausilipon Hospital, Naples); and Umberto Ferbo (Moscati Hospital, Avellino).

▽ Published ahead of print on 14 April 2010.

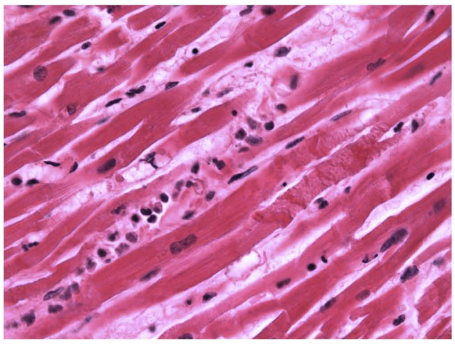

FIG. 1. Myocardium sample showing moderate myocarditis and the presence of histiocyte infiltration and focal necrosis. The sample was stained with hematoxylin and eosin. Original magnification, ×250.

other tissues (the pharynx, tonsils, spleen, mesenteric lymph nodes, and liver).

The patient was also examined for other common bacterial and viral respiratory agents (e.g., pneumococcus, mycoplasma, and respiratory syncytial virus [RSV], etc.), and the results were all negative.

Sequence analysis of the hemagglutinin (HA) showed a D222E mutation (H1 numbering), which has been detected in many countries at frequencies of <0.01 to 31.7% (13). It is unclear whether and to what extent change at amino acid position 222, which is located in the receptor binding domain of the HA, influences receptor binding specificity and, ultimately, virus pathogenicity. With regard to neuraminidase (NA) genes, no significant mutations were detected, and the virus was found to be susceptible to NA inhibitors.

The study findings suggest that cardiac tamponade and heart failure following myopericarditis in this young patient were

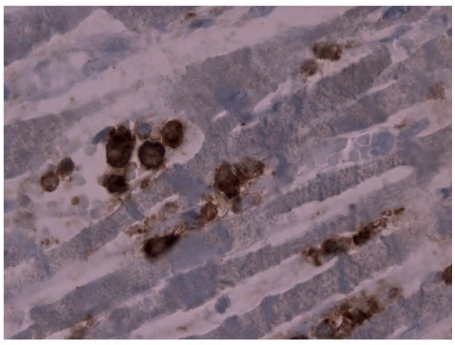

FIG. 2. Cluster of histiocytic cells (CD68 positive) in the neighborhood of damaged myocytes. Cells were stained with peroxidase. Original magnification, ×250.

TABLE 1. Laboratory results

| Clinical sample or tissue type | Result[a] from: | |
|---|---|---|
| | Real-time RT-PCR analysis[b] for pandemic A/H1N1 | Isolation in MDCK-SIAT1 cells (HAU) |
| Rhinopharyngeal swab | **Positive** | **Positive (64)** |
| Pharynx | Negative | Negative |
| Tonsil | Negative | Negative |
| Bronchus | **Positive** | Negative |
| Lung | Negative | Negative |
| Spleen | Negative | Negative |
| Heart | **Positive** | **Positive (16)** |
| Pericardium | **Positive** | Negative |
| Pericardial fluid | **Positive** | Negative |

[a] Positive results are listed in bold.
[b] Real-time reverse transcriptase PCR (RT-PCR) analysis was performed using the CDC pandemic kit.

due to direct tropism and damage caused by the 2009 H1N1 influenza virus. Cardiac involvement is not an uncommon complication of seasonal influenza. Elevated CK levels in 12% of patients affected by influenza but without cardiac symptoms and in up to 15% of patients with abnormal ECG patterns suggestive of myocarditis have been reported previously (4, 7). Myocardial damage or clinical myocarditis has also been reported in studies from Japan (6) and Canada (9). In molecular studies conducted in the United States and in Italy, influenza virus accounted for 2 to 10% of all viral agents isolated from myocardial tissues (1, 3).

---

Viruses are the most important infectious cause of myocarditis. Molecular studies of cardiac samples obtained through cardiac catheterization of patients presenting with acute viral myocarditis resulted in the identification of viral genomes in 38 to 53% of the cases (1, 3). Enteroviruses, especially group B coxsackieviruses, appear to be the major agents implicated (11), but other viruses may also be involved. Influenza viruses A and B may play a role in myocarditis and/or pericarditis, and sporadic cases of myocarditis due to the influenza virus have been reported over the past few decades (8). Overall, there is consistent evidence that influenza viruses may trigger cardiovascular death and that influenza vaccines reduce the risk of cardiac events in subjects with established cardiovascular disease. However, most of the evidence comes from observational and interventional studies of adults and the elderly in whom a cardiac event generally consists of a clinical sequela of primary respiratory infection (8, 12).

Cardiac involvement in influenza is usually reported to occur between 4 and 9 days after the onset of influenza symptoms and is characterized by worsening dyspnea, ECG abnormalities (i.e., ST elevation and Q waves), elevation of cardiac enzymes (e.g., the M [muscle type] and B [brain type] subunits of CK [CK-MB]), and impaired left ventricular function (10). Fulminant myocarditis is characterized by profound left ventricular dysfunction and cardiogenic shock, which requires inotropic support. In some cases, pericardial effusions may result in cardiac tamponade. Myocyte damage may be due to a direct cytolytic effect of the virus or to the host immune response (5); the former mechanism usually plays the major role in cases

with early myocardial involvement, whereas immune response-mediated damage is most likely to occur in later phases of infection (8). In our case, cardiac involvement occurred early in the course of infection, on the third day after the onset of fever, and was associated with fever decrease by crisis. The timing of cardiac involvement, in combination with virus detection in myocardial and pericardial tissues and fluid, strongly suggests a direct effect of the 2009 pandemic influenza virus on the myocardium and pericardium. Importantly, increased virus tropism for myocardial tissue may be favored by the cytokine cascade, enhancing or modifying virus receptor exposure on endothelial cells lining the myocardial tissue (12). A recent study has already reported a severe form of acute myocarditis in four children confirmed to be positive for the 2009 pandemic virus (2). Our work further demonstrates that cardiac tamponade resulting from myopericarditis may occur during pandemic H1N1 influenza virus infection in young subjects, without known predisposing factors. All these observations suggest the possibility that the novel H1N1 virus is more commonly associated with severe forms of myocarditis than previously circulating influenza virus strains. Furthermore, the above data strongly highlight the importance of early diagnosis and treatment in these cases to reduce the risk of severe cardiac events in children.

We thank Tiziana Grisetti for editing the manuscript.

This work has been partially supported by grants from the Ministero della Salute-Istituto Superiore di Sanità within the research project "Come Contrastare la Pandemia Influenzale: Individuazione di Nuovi Farmaci Efficaci."

## REFERENCES

1. **Bowles, N. E., N. Ji, D. L. Kearney, et al.** 2003. Detection of viruses in myocardial tissues by polymerase chain reaction: evidence of adenovirus as a common cause of myocarditis in children and adults. J. Am. Coll. Cardiol. **42:**466–472.
2. **Bratincsàk, A., H. G. El-Said, J. S. Bradley, K. Shayan, P. D. Grossfeld, and C. R. Cannavino.** 2010. Fulminant myocarditis associated with pandemic H1N1 influenza A virus in children. J. Am. Coll. Cardiol. **55:**928–929.
3. **Calabrese, F., E. Carturan, C. Chimenti, et al.** 2004. Overexpression of tumor necrosis factor (TNF)α and TNFα receptor I in human viral myocarditis: clinicopathologic correlations. Mod. Pathol. **17:**1108–1118.
4. **Greaves, K., J. S. Oxford, C. P. Price, et al.** 2003. The prevalence of myocarditis and skeletal muscle injury during acute viral infection in adults: measurement of cardiac troponins I and T in 152 patients with acute influenza infection. Arch. Intern. Med. **163:**165–168.
5. **Hanasaki, K., A. Varki, I. Stamenkovic, and M. P. Bevilacqua.** 1994. Cytokine-induced beta-galactoside alpha-2,6-sialyltransferase in human endothelial cells mediates alpha 2,6-sialylation of adhesion molecules and CD22 ligands. J. Biol. Chem. **269:**10637–10643.
6. **Kaji, M., H. Kuno, T. Turu, et al.** 2001. Elevated serum myosin light chain I in influenza patients. Intern. Med. **40:**594–597.
7. **Karjalainen, J., M. S. Nieminen, and J. Heikkila.** 1980. Influenza A1 myocarditis in conscripts. Acta Med. Scand. **207:**27–30.
8. **Mamas, A. M., D. Fraser, and L. Neyses.** 2008. Cardiovascular manifestations associated with influenza virus infection. Intern. J. Cardiol. **130:**304–309.
9. **Moore, D. L., W. Vaudry, D. W. Scheifele, et al.** 2006. Surveillance for influenza admissions among children hospitalized in Canadian immunization monitoring program active centers, 2003–2004. Pediatrics **118:**610–619.
10. **Onitsuka, H., T. Imamura, N. Miyamoto, et al.** 2001. Clinical manifestations of influenza A myocarditis during the influenza epidemic of winter 1998–1999. J. Cardiol. **37:**315–323.
11. **Savoia, M. C., and M. N. Oxman.** 1995. Myocarditis and pericarditis, p. 799–813. *In* G. L. Mandell, J. E. Douglas, and R. Dolin (ed.), Principles and practice of infectious diseases, 4th ed. Churchill Livingstone Inc., New York, NY.
12. **Warren-Gash, C., L. Smeeth, and A. C. Hayward.** 2009. Influenza as a trigger for acute myocardial infarction or death from cardiovascular disease: a systematic review. Lancet Infect. Dis. **9:**601–610.
13. **World Health Organization.** 28 December 2009. Preliminary review of D222G amino acid substitution in the haemagglutinin of pandemic influenza A (H1N1) 2009 viruses. World Health Organization, Geneva, Switzerland. http://www.who.int/csr/resources/publications/swineflu/h1n1_d222g/en/index.html.