

Genes Associated with Recurrence of Hepatocellular Carcinoma: Integrated Analysis by Gene Expression and Methylation Profiling

Ju Dong Yang¹, So-Young Seol²,
Sun-Hee Leem², Yong Hoon Kim³,
Zhifu Sun⁴, Ju-Seog Lee⁵,
Snorri S. Thorgeirsson⁶, In-Sun Chu⁷,
Lewis R. Roberts¹ and Koo Jeong Kang^{1,3}

¹Miles and Shirley Fiterman Center for Digestive Diseases, College of Medicine, Mayo Clinic, Rochester, MN, USA; ²Department of Biological Sciences, Dong-A University, Busan; ³Department of Surgery, Keimyung University School of Medicine, Daegu, Korea; ⁴Division of Biomedical Statistics and Informatics, College of Medicine, Mayo Clinic, Rochester, MN; ⁵Department of Systems Biology, The University of Texas M. D. Anderson Cancer Center, Houston, TX; ⁶Laboratory of Experimental Carcinogenesis, National Cancer Institute, Bethesda, MD, USA; ⁷Korean Bioinformation Center, Korea Research Institute of Bioscience & Biotechnology, Daejeon, Korea

Received: 23 April 2011
Accepted: 6 September 2011

Address for Correspondence:

Koo Jeong Kang, MD
Department of Surgery, Keimyung University Dong-San Medical Center, 56 Dalseong-ro, Jung-gu, Daegu 700-712, Korea
Tel: +82.53-250-7655, Fax: +82.53-250-7322
E-mail: kjkang@dsmc.or.kr

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST), Republic of Korea (2009-0063260, S.-H.L.) and the intramural program of KRIBB (I.-S.C.). Microarray study was supported by Shared Research Equipment Assistance Program (S.-H.L.) by Korea Basic Science Institute, MEST, and NIH grants CA100882 and CA128633 (to LRR).

INTRODUCTION

Hepatocellular carcinoma (HCC) is the seventh most common cancer and the third leading cause of cancer related mortality in the world (1). However, the mechanisms of liver cancer initiation and progression are incompletely elucidated. DNA methylation is one of the major mechanisms for the epigenetic suppression of gene activity during carcinogenesis. Methylation of tumor suppressor genes has the potential to be used as a biomarker for cancer detection as well as a prognostic marker in cancer (2). An increasing number of putative tumor suppressor genes have been reported to be down regulated by promoter CpG island methylation in HCC. Hypomethylation can also be implicated in tumorigenesis through the activation of oncogenes

and local increases in chromosomal instability. Gene expression is suppressed by DNA methylation. The goal of this study was to identify genes whose CpG site methylation and mRNA expression are associated with recurrence after surgical resection for hepatocellular carcinoma (HCC). Sixty-two HCCs were examined by both whole genome DNA methylation and transcriptome analysis. The Cox model was used to select genes associated with recurrence. A validation was performed in an independent cohort of 66 HCC patients. Among fifty-nine common genes, increased CpG site methylation and decreased mRNA expression were associated with recurrence for 12 genes (Group A), whereas decreased CpG site methylation and increased mRNA expression were associated with recurrence for 25 genes (Group B). The remaining 22 genes were defined as Group C. Complement factor H (*CFH*) and myosin VIIA and Rab interacting protein (*MYRIP*) in Group A; proline/serine-rich coiled-coil 1 (*PSRC1*), meiotic recombination 11 homolog A (*MRE11A*), and myosin 1E (*MYO1E*) in Group B; and autophagy-related protein LC3 A (*MAP1LC3A*), and NADH dehydrogenase 1 alpha subcomplex assembly factor 1 (*NDUFAF1*) in Group C were validated. In conclusion, potential tumor suppressor (*CFH*, *MYRIP*) and oncogenes (*PSRC1*, *MRE11A*, *MYO1E*) in HCC are reported. The regulation of individual genes by methylation in hepatocarcinogenesis needs to be validated.

Key Words: Carcinoma, Hepatocellular; Gene Expression Profiling; Microarray analysis; DNA methylation; Survival

and local increases in chromosomal instability.

Comparative and integrative functional genomics using genome wide microarray analysis methods have improved the molecular classification and prognostic prediction of HCC (3, 4). Recently, array based technology assessing genome-wide DNA methylation has been used to identify genes that are silenced by methylation in HCC (5, 6). Recent studies have also shown that particular methylation patterns revealed through genome wide DNA methylation array analysis are associated with cancer risk and poor prognosis in HCC (5).

In this study, we hypothesized that epigenetic changes in individual genes, (CpG site hypermethylation of tumor suppressor genes and hypomethylation of oncogenes), are associated with HCC recurrence after surgical resection with curative in-

tent. To test our hypothesis, we obtained gene expression and methylation profiling data using surgically-resected HCC tissues. Genes whose methylation and expression are associated with recurrence were identified in the training data set and validated in an independent data set. In addition, we aimed to identify pathways enriched by recurrence predicting genes using mRNA expression and CpG site methylation.

MATERIALS AND METHODS

Patients

Patients included in this study were treated with surgical resection for HCC between December, 2001 and December 2007 at a single institution, Keimyung University Dong-San Medical Center, Daegu, Korea. A total of 169 patients had resection with curative intent during the period and gene expression and methylation profiling were available from 62 HCC tissues for the analyses.

Clinical and laboratory data including histopathology at HCC diagnosis and information on recurrence and survival after resection were collected. After surgery, the patients were seen every three to six months for assessment of tumor recurrence using serum alpha-fetoprotein, PIVKA-II, and high-resolution multi-detector CT. Validation of recurrence predicting genes was performed using data previously published by Lee et al. (4) which includes gene expression profiles of 139 HCCs. Data on tumor recurrence was available for 66 of the 139 patients and was used for the validation. As methylation profiling data was not available in the validation data set, only the association between mRNA expression and recurrence could be validated.

Transcriptome profiling

The process of transcriptome profiling was described in our previous study (7). Briefly, total RNA from liver cancer samples was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). Labeled, amplified material (1,500 ng per array) was hybridized to Version 3 of the Illumina Human-6 BeadChip (48 K) (Illumina Inc., San Diego, CA, USA). Array signals were developed using Amersham fluorolink streptavidin-Cy3 (GE Healthcare Biosciences, Little Chalfont, UK). Arrays were scanned with an Illumina BeadArray Reader confocal scanner (BeadStation 500-GXDW; Illumina, Inc.). Gene expression profiles of 48,803 probes from 62 individual HCC tissue samples were obtained. Transcriptome profiling for the validation set has also been previously reported (4). RNA from HCC and adjacent benign tissue from the 139 HCCs was analyzed at the US National Cancer Institute using the Qiagen Human Array-Ready Oligo Set (version 2.0), which contains 70-mer probes for 21,329 genes. CsCl density-gradient centrifugation was used to isolate total RNAs from frozen liver tissue. Total RNA from 19 normal livers was used as the reference for all microarray experiments.

DNA samples for whole genome methylation analysis

DNA was extracted using the QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA). The quality of genomic DNA was checked by gel electrophoresis and DNA concentrations were determined using the Ultrospec 3100 pro spectrophotometer (Amersham Bioscience, Buckinghamshire, UK).

Bisulfite conversion of DNA and BeadArray technology

The EZ DNA methylation kit (Zymo Research, Orange, CA, USA) was used for bisulfite conversion of 2 µg of genomic DNA. After bisulfite treatment, the remaining assay steps were identical to the GoldenGate genotyping assay using Illumina-supplied reagents and conditions (8). 5 µL of the converted DNA was used with the Infinium HumanMethylation27 v1.0 BeadChip array from Illumina to genotype bisulfite-converted and unconverted CpGs.

The DNA was biotinylated and purified from excess biotin. Assay oligonucleotides were added and hybridized to resuspended DNA. Allele-specific extension and ligation of the hybridized oligos was performed; fluor-labeled strands were then hybridized to the Illumina HumanMethylation27 BeadChip. Arrays were scanned with an Illumina BeadArray Reader confocal scanner. The chip interrogates 27,578 CpG sites of over 14,000 genes.

Statistics for survival

Gene expression microarray data pre-processing was performed using Illumina BeadStudio software version 3. The un-normalized and non-background subtracted data was then normalized using the quantile normalization method in the Linear Models for Microarray Data (LIMMA) package in the R language environment. The Cox proportional hazards model was used to select genes that were significantly associated with recurrence. Patients without an event at the time of last follow-up were treated as censored. The hazard ratio and nominal *P* value were obtained for each gene. As the gene expression data was in log₂ scale, the hazard ratio (HR) represented an increased or decreased risk as gene expression value doubled.

As methylation status per tested locus was reported as a ratio from 0 to 1, it was transformed into a 1 to 10 scale so that every one unit of increase or decrease represented a 10% methylation difference for the estimated hazard ratio from the Cox proportional hazards model. Fig. 1 schematically illustrates the experimental design and analytical steps. Individual genes for which the expression or methylation status in the training set was associated with significant differences in recurrence-free survival were identified. The independent data set was used to validate the association of mRNA expression of the individual genes selected in the training set with recurrence. All analyses were conducted using the R statistical package (<http://www.r-project.org/>) or Partek[®] genomic software suite.

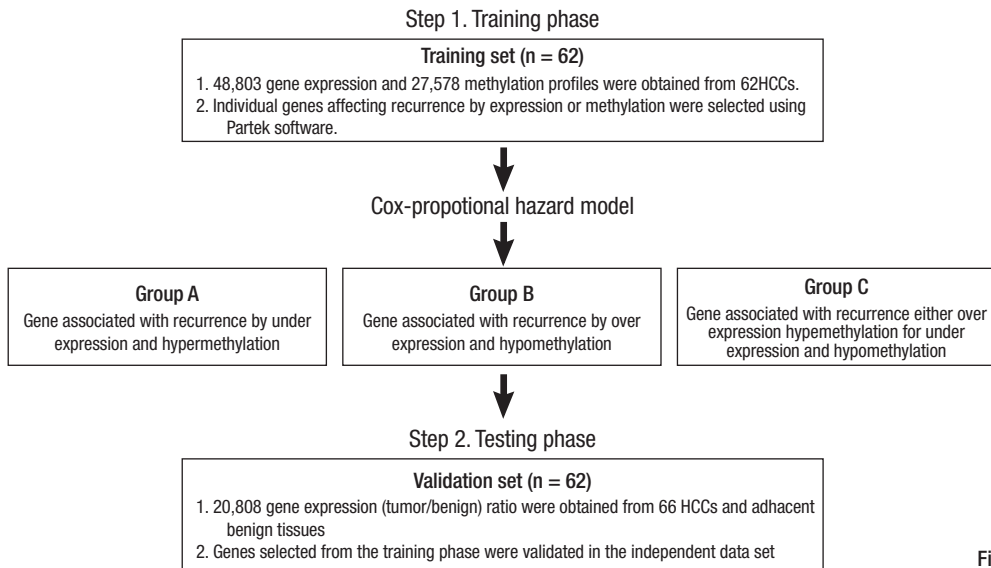


Fig. 1. Flow chart of the experimental protocol.

Table 1. Demographic and clinicopathologic characteristics of the 62 HCC training set patients

Characteristics	Data	Characteristics	Data
Age (yr)	56.0 ± 10.2 (31-77)	Size of tumor (cm)	5.9 ± 4.0
Sex		Number of tumor	
Male	46 (74.2%)	1	53 (85.5%)
Female	16 (25.8%)	2	5 (8.1%)
Causes of underlying liver disease		3	2 (3.2%)
HBV	47 (75%)	4	1 (1.6%)
HCV	3 (4.8%)	5	1 (1.6%)
HBV/HCV combined	1 (1.6%)	Edmonson-Steiner Grade	
Alcohol	3 (4.8%)	1	2 (3.4%)
Unknown	8 (12.9%)	2	36 (61.0%)
AFP (ng/mL)	3756 ± 13226	3	19 (33.9%)
< 10	17 (27.4%)	4	2 (3.4%)
10 >=, < 400	28 (45.2%)	Stage (TNM)	
>= 400	17 (27.4%)	1	35 (56.5%)
AST (IU/mL)	52.1 ± 39.4	2	19 (30.6%)
ALT (IU/mL)	42.2 ± 22.5	3	8 (12.8%)
ICG R15(%)	11.3 ± 6.1	Recurrence	
< 10	27 (43.5%)	Yes	37 (59.7%)
=> 10	35 (56.5%)	No	25 (40.5%)
Cirrhosis		Recurrence free survival (months)	23.5 [7.0-74.1]*
Yes	40 (64.5%)	Overall survival (months)	66.2 [22.7-]*
No	22 (35.4%)		

Data shown as mean ± standard deviation (range) or median [inter quartile range] for continuous variable and proportion (%) for categorical variables *Median [interquartile range]. HBV, hepatitis B virus; HCV, hepatitis C virus; AFP, alpha-fetoprotein; AST, aspartate aminotransferase; ALT, alanine transaminase; ICG, indocyanine green test.

Statistics for pathway analysis

MetaCore software (<http://www.genego.com/metacore.php>) was used to investigate the pathways associated with recurrence predicting gene sets by mRNA expression and CpG site methylation using the default Metacore reference database. Pathways with which the recurrence predicting genes were associated more highly than would be expected by chance were identified. The significance of the association with pathway was estimated by hypergeometric test *P* value.

Ethics statement

The study protocol was approved by the institutional review board for the use of human subjects at the Keimyung University School of Medicine (IRB No: 08-122).

All participants provided written informed consent.

RESULTS

Clinical characteristics of the patients

Table 1 summarizes the clinical characteristics of the patients in this study. The median age was 56 yr and there was a strong male

Table 2. Top ten signaling pathways enriched by recurrence predicting gene by mRNA expression or CpG site methylation

Recurrence predicting gene by mRNA expression		Recurrence predicting gene by CpG site Methylation	
Pathways	P value	Pathways	P value
Cytoskeleton remodeling: TGF, WNT, and cytoskeletal remodeling	3.2E-07	Development: IGF1 receptor signaling	2.8E-05
Immune response: alternative complement pathways	3.9E-07	Development: FGFR signaling pathways	3.7E-05
Immune response: lectin induced complement pathways	1.2E-06	Signal transduction: AKT signaling	1.0E-04
Immune response: classical complement pathways	2.6E-06	Cytoskeleton remodeling: TGF,WNT, and cytoskeletal remodeling	1.5E-04
Blood coagulation: Blood coagulation	1.9E-05	Development: PIP3 signaling in cardiac myocyte	1.7E-04
Cell cycle: metaphase check point	5.3E-05	Development: WNT signaling pathways	3.7E-04
Cytoskeleton remodeling: cytoskeleton remodeling	6.0E-05	Proteolysis: Putative ubiquitin pathways	6.5E-04
Cell cycle: spindle assembly and chromosome separation	1.5E-04	Immune response: PIP3 signaling in B lymphocytes	8.2E-04
Development: VEGF-family signaling	1.8E-04	Transcription: CREB pathways	1.0E-03
Transcription: androgen receptor nuclear signaling	4.0E-04	Signal transduction: PTEN pathways	1.3E-03

Table 3. Twelve genes (Group A) for which both increased CpG site methylation and decreased mRNA expression were associated with recurrence of HCC

Gene symbol	Annotation	Training set				Testing set	
		Methylation		Expression		Expression	
		HR	P value	HR	P value	HR	P value
<i>MYRIP</i>	Myosin VIIA and Rab interacting protein	1.84	0.01	0.15	0.02	0.65	< 0.01
<i>CFH</i>	Complement factor H isoform a precursor	1.44	0.01	0.65	0.001	0.64	< 0.01
<i>MRPL4</i>	Mitochondrial ribosomal protein L4 isoform a	39.9	0.04	0.3	0.01	None	
<i>THBS1</i>	Thrombospondin 1 precursor	11.7	0.02	0.76	0.04	1.83	0.09
<i>GPSN2</i>	Glycoprotein; synaptic 2	5.05	0.01	0.47	0.03	None	
<i>CHKA</i>	Choline kinase alpha isoform a	3.09	0.02	0.09	0.03	0.82	0.4
<i>TIMP3</i>	Tissue inhibitor of metalloproteinase 3 precursor	2.33	0.04	0.62	< 0.01	1.4	0.2
<i>PCDHB11</i>	Protocadherin beta 11 precursor	2.25	< 0.01	0.16	0.01	None	
<i>HSD17B6</i>	3-hydroxysteroid epimerase	1.39	0.04	0.84	0.02	None	
<i>KLK10</i>	Kallikrein 10 precursor	1.28	0.04	0.1	0.04	0.96	0.93
<i>CYGB</i>	Cytoglobin	1.24	0.03	0.6	0.03	None	
<i>FXSD1</i>	Phospholemman precursor	1.19	0.02	0.69	0.04	0.92	0.62

HR, hazard ratio.

preponderance. All patients were Asian and HBV was the most common etiology of HCC. More than 85% of the patients had a single lesion and stage 1 or stage 2 HCC. Recurrence developed in 60% of patients with a median follow up of 20.2 months. The median overall survival was 66.2 months and the median recurrence-free survival was 23.5 months. Clinical characteristics of the patients in the validation data set were reported in the previous study (4). Briefly, the median age was 57 yr and 73% were male. HBV (56%) was the most common etiology. In contrast to the training cohort, only 44% were Asian. Recurrence developed in 65% of patients with a median follow up of 24.6 months.

Genes whose mRNA expression and CpG site methylation associated with recurrence and pathway enriched by recurrence predicting genes

At a P value < 0.05, 3,876 genes were associated with recurrence by mRNA expression. These genes were enriched in cytoskeleton remodeling processes through TGF, Wnt, and cytoskeletal remodeling; immune response through classic, alternative, and lectin induced complement pathways; the cell cycle through metaphase check point, and spindle assembly and chromosome separation; and blood coagulation, VEGF, and androgen recep-

tor nuclear signaling pathways.

Similarly, genes that were significantly associated with recurrence by methylation were selected. At a P value < 0.05, 541 CpG sites were included in this category. Pathways enriched in genes regulated by the 541 recurrence predicting CpG sites were development through IGF-1R, FGFR, PIP3, and Wnt; signal transduction through AKT and PTEN; cytoskeleton remodeling process through TGF, Wnt, and cytoskeletal remodeling; immune response through PIP3 signaling in B cells; proteolysis through ubiquitin; and transcription through CREB Pathways. Of note, the cytoskeleton remodeling, TGF, Wnt, and cytoskeletal remodeling pathway was enriched in both recurrence predicting gene sets (Table 2).

Common genes for which both methylation and mRNA expression were associated with recurrence

Fifty-nine genes were associated with recurrence by both CpG site methylation and mRNA expression (Fig. 2). These common genes were classified into three groups. Group A includes 12 genes for which both increased CpG site methylation and decreased mRNA expression were associated with recurrence (Table 3). Group B includes 25 genes for which both decreased gene meth-

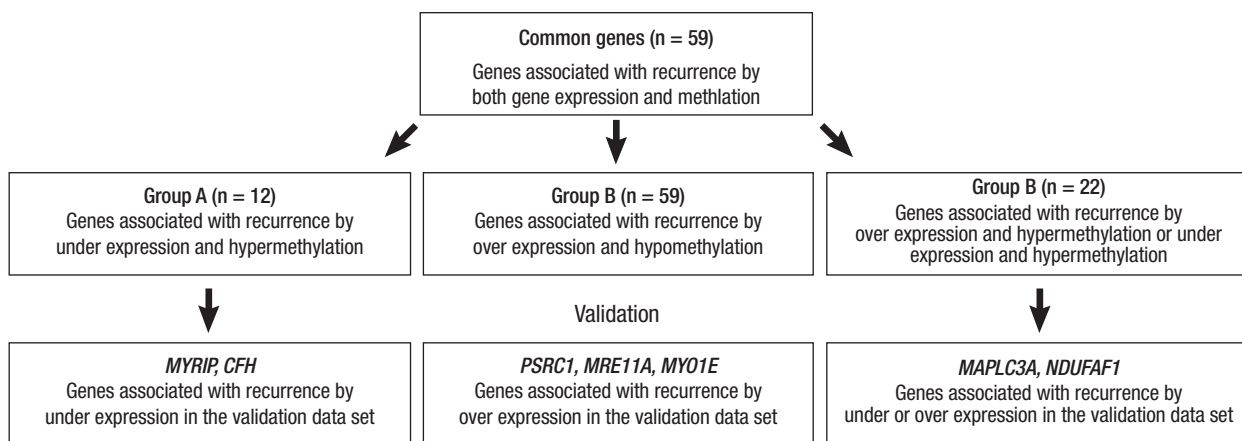


Fig. 2. Selection of recurrence associated genes in the training and validation data set. *MYRIP*, Myosin VIIA and Rab interacting protein; *CFH*, Complement factor H; *PSRC1*, Proline/serine-rich coiled-coil 1; *MRE11A*, Meiotic recombination 11 homolog A; *MYO1E*, Myosin IE; *MAP1LC3A*, autophagy-related protein LC3 A; *NDUFAF1*, NADH dehydrogenase 1 alpha subcomplex assembly factor 1.

Table 4. Twenty-five genes (Group B) for which both decreased CpG site methylation and increased mRNA expression were associated with recurrence of HCC

Gene symbol	Annotation	Training Set				Testing Set	
		Methylation		Expression		Expression	
		HR	P value	HR	P value	HR	P value
<i>PSRC1</i>	p53-regulated DDA3 isoform a	4.19E-07	0.04	4.59	0.01	3.16	0.03
<i>MRE11A</i>	Meiotic recombination 11 homolog A isoform 1	6.61E-04	0.03	4.59	0.01	2.10	0.02
<i>MYO1E</i>	Myosin IE	0.01	0.02	11.4	0.01	2.49	< 0.01
<i>NR2C1</i>	Nuclear receptor subfamily 2; group C; member 1 isoform a	1.61E-10	0.01	2.82	0.04	2.25	0.09
<i>NSUN4</i>	NOL1/NOP2/Sun domain family 4 protein	1.13E-07	0.01	2.01	0.03	1.92	0.17
<i>C6orf192</i>	Hypothetical protein LOC116843	2.30E-06	0.02	1.84	0.01	0.92	0.90
<i>C1orf156</i>	Hypothetical protein MGC9084	6.03E-05	0.01	5.23	0.02	0.89	0.85
<i>NUP43</i>	Nucleoporin 43 kDa	7.11E-05	0.02	2.68	0.01	1.43	0.35
<i>RDH14</i>	Retinol dehydrogenase 14 (all-trans and 9-cis)	1.38E-04	< 0.01	0.006	< 0.01	0.46	0.19
<i>COPS8</i>	COP9 signalosome subunit 8 isoform 1	2.99E-04	0.01	10.7	0.02	0.67	0.03
<i>METTL2B</i>	Methyltransferase like 2	3.27E-04	0.02	14.6	< 0.01	0.92	0.89
<i>MED18</i>	Mediator of RNA polymerase II transcription; subunit 18 homolog	4.86E-04	0.03	3.65	0.03	1.00	0.99
<i>CBX8</i>	Chromobox homolog 8	4.90E-04	0.02	2.26	0.01	0.51	0.18
<i>EIF2B5</i>	Eukaryotic translation initiation factor 2B; subunit 5 epsilon; 82 kDa	5.17E-04	0.01	2.74	0.01	0.81	0.72
<i>EYA3</i>	Eyes absent 3	6.94E-04	0.04	6.43	< 0.01	3.21	0.07
<i>CHD6</i>	Chromodomain helicase DNA binding protein 6	0.001	0.04	2.48	0.02	1.38	0.50
<i>QARS</i>	Glutaminyl-tRNA synthetase	0.001	0.04	2.01	0.04	0.72	0.31
<i>AMPD2</i>	Adenosine monophosphate deaminase 2 (isoform L) isoform 3	0.01	0.03	15.3	< 0.01	1.34	0.33
<i>LRIG2</i>	Leucine-rich repeats and immunoglobulin-like domains 2	0.02	0.02	6.65	< 0.01	1.36	0.17
<i>SERAC1</i>	Serine active site containing 1	0.41	0.02	9.01	0.01	1.65	0.34
<i>SLC45A2</i>	Membrane-associated transporter protein isoform a	0.61	0.03	2.07	0.01	0.94	0.9
<i>DNMT3B</i>	DNA cytosine-5 methyltransferase 3 beta isoform 1	0.74	0.01	2.41	0.04	2.21	0.17
<i>NMB</i>	Neuromedin B isoform 1	0.80	0.03	1.83	0.01	None	
<i>FANCG</i>	Fanconi anemia; complementation group G	0.83	0.02	2.40	0.01	2.59	0.05
<i>SMAD2</i>	Sma- and Mad-related protein 2	0.87	0.03	5.65	0.04	0.54	0.11

HR, hazard ratio.

ylation and increased mRNA expression were associated with recurrence (Table 4). We also defined Group C genes for which

both decreased CpG site methylation and mRNA expression or both increased CpG site methylation and mRNA expression were

Table 5. Twenty-two genes (Group C) for which both decreased CpG site methylation and decreased mRNA expression or both increased CpG site methylation and increased mRNA expression were associated with recurrence of HCC

Gene symbol	Annotation	Training Set				Testing Set	
		Methylation		Expression		Expression	
		HR	P value	HR	P value	HR	P value
<i>MAP1LC3A</i>	Autophagy-related protein LC3 A	0.78	0.04	0.06	0.01	0.63	0.04
<i>NDUFAF1</i>	NADH dehydrogenase 1 alpha subcomplex, assembly factor 1	0.83	0.03	0.50	0.03	0.27	< 0.01
<i>BMF</i>	Bcl2 modifying factor	4.77E-05	0.05	0.13	0.04	1.14	0.52
<i>MRFAP1L1</i>	Morf4 family associated protein 1-like 1	3.01E-04	0.02	0.27	0.02	None	
<i>CLPP</i>	Endopeptidase Clp	3.02E-03	0.05	0.44	0.03	0.81	0.72
<i>PECI</i>	Dodecenoyl-CoA isomerase	0.20	0.03	0.72	0.02	1.99	0.24
<i>CT45-2</i>	Cancer/testis antigen 45-2	0.61	0.02	0.05	< 0.01	None	
<i>MS4A3</i>	CD20 antigen-like protein	0.63	0.02	0.10	0.03	None	
<i>PAPSS2</i>	PAPS synthase 2	0.66	0.04	0.66	0.03	0.90	0.52
<i>MAPKAP1</i>	Mitogen-activated protein kinase associated protein 1	0.68	0.04	0.45	0.04	0.84	0.66
<i>MAGEA11</i>	Melanoma antigen family A, 11	0.75	0.03	0.05	0.04	1.64	0.18
<i>LCT</i>	Lactase-glycosylceramidase	0.76	0.02	0.06	0.02	0.94	0.87
<i>THSD4</i>	Thrombospondin, type I, domain containing 4	0.78	0.03	0.27	0.03	2.58	0.07
<i>CLEC4A</i>	C-type lectin domain family 4, member A	0.80	0.03	0.15	0.01	2.69	0.03
<i>SEMA3B</i>	Semaphorin-3B	1.15	0.04	2.56	0.01	0.77	0.54
<i>HNT</i>	Neurotrimin	1.16	0.02	2.27	0.03	None	
<i>NEUROD2</i>	Neurogenic differentiation 2	1.19	0.04	6.84	0.05	1.50	0.43
<i>PCDHB14</i>	Protocadherin beta 14	1.19	0.05	12.85	0.02	1.22	0.78
<i>BDNF</i>	Brain-derived neurotrophic factor	1.20	< 0.01	9.29	0.03	1.31	0.75
<i>TSSC1</i>	Tumor suppressing subtransferable candidate 1	1.80	0.04	3.05	0.01	0.98	0.96
<i>ZNF267</i>	Zinc finger protein 267	1.81	0.05	9.40	0.04	1.46	0.26
<i>STARD3NL</i>	STARD3 N-terminal-like protein	2.35	< 0.01	2.24	0.03	0.64	0.30

HR, hazard ratio.

associated with recurrence (Table 5).

Kaplan-Meier curves were generated using the training data set and recurrence-free survival was compared between groups classified above or below the median (the 31 patients with above median mRNA expression vs the other 31 patients with below median gene expression; and the 31 patients with above median CpG site methylation vs the other 31 patients with below median CpG site methylation) using the log rank test (Fig. 3). Some of the genes show significant differences in recurrence free survival (*MRE11A* & *CFH*) when expression/methylation patterns were dichotomized.

Validation

The validation data set were used to examine whether the mRNA expression of the 59 common genes were associated with recurrence of HCC in an independent data set. Down-regulation of myosin VIIA and Rab interacting protein (*MYRIP*) and complement factor H (*CFH*) in Group A; up-regulation of proline/serine-rich coiled-coil 1 (*PSRC1*), meiotic recombination 11 homolog A (*MRE11A*), and myosin IE (*MYOIE*) in Group B; and down regulation of autophagy-related protein LC3 A (*MAP1LC3A*), and NADH dehydrogenase 1 alpha subcomplex assembly factor 1 (*NDUFAF1*) in Group C were associated with recurrence

in the validation data set. Mitochondrial ribosomal protein L4 (*MRPL4*), protocadherin beta 11 (*PCDHB11*), synaptic glycoprotein SC2 (*GPSN2*), cytoglobin (*CYGB*), hydroxysteroid 17-beta dehydrogenase 6 (*HSD17B6*) in Group A and neuromedin B (*NMB*) in Group B, morf4 family associated protein 1-like 1 (*MRFAP1L1*), cancer/testis antigen 45-2 (*CT45-2*), CD20 antigen-like protein (*MS4A3*), and neurotrimin (*HNT*) in Group C could not be validated due to the absence of a corresponding oligonucleotide in the validation data set.

DISCUSSION

In this study, we have first identified two sets of recurrence predicting genes by 1) mRNA expression, and 2) promoter CpG site methylation. We identified pathways enriched in the two sets of recurrence predicting genes and also identified genes common to both the mRNA and CpG methylation analyses. Among 59 common genes, Genes in Group A were considered to be potential tumor suppressors since both CpG hypermethylation and down regulation of mRNA expression were associated with recurrence. In contrast, genes in Group B were considered to be potential oncogenes because CpG hypomethylation and increased mRNA expression were associated with recurrence. We will first discuss the relevant pathways enriched by recurrence

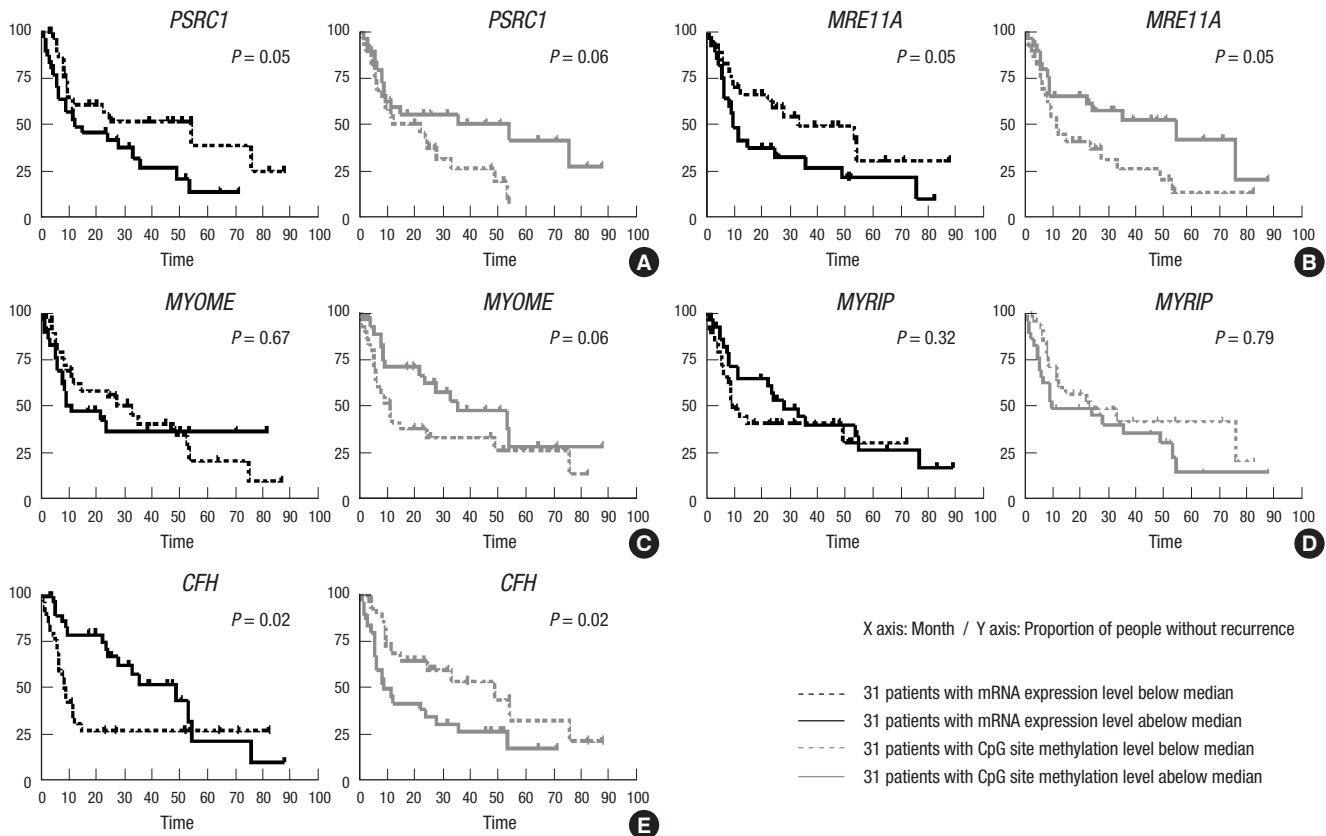


Fig. 3. Effect of individual gene expression and methylation on recurrence free survival. Patients were divided into two groups based on their gene expression and methylation status (over expression vs under expression; hypermethylation vs hypomethylation). Recurrence-free survival was compared between two groups in the training data set. Group A and B genes associated with recurrence both in the training and validation data sets were included in the figure. (A) 31-patients with *PSRC1* mRNA expression level above median is more likely develop recurrence than the other 31 patients with *PSRC1* mRNA expression level below median ($P = 0.05$). On the other hand 31 patients with *PSRC1* CpG site methylation level above median is less likely to develop recurrence than the other 31 patients with *PSRC1* CpG site methylation level below median ($P = 0.06$). (B) 31-patients with *MRE11A* mRNA expression level above median is more likely develop recurrence than the other 31 patients with *MRE11A* mRNA expression level below median ($P = 0.05$). On the other hand 31 patients with *MRE11A* CpG site methylation level above median is less likely to develop recurrence than the other 31 patients with *MRE11A* CpG site methylation level below median ($P = 0.05$). (C) *MYO1E* mRNA expression level was not associated with recurrence when patients were dichotomized ($P = 0.67$). On the other hand 31 patients with *MYO1E* CpG site methylation level above median is less likely to develop recurrence than the other 31 patients with *MYO1E* CpG site methylation level below median ($P = 0.06$). (D) *MYRIP* mRNA expression level ($P = 0.32$) and CpG site methylation level ($P = 0.79$) were not associated with recurrence when patients were dichotomized. (E) 31-patients with *CFH* mRNA expression level above median is less likely develop recurrence than the other 31 patients with *CFH* mRNA expression level below median ($P = 0.02$). On the other hand 31 patients with *CFH* CpG site methylation level above median is more likely to develop recurrence than the other 31 patients with *PSRC1* CpG site methylation level below median ($P = 0.02$).

predicting genes. Specific genes that have been reported to function as tumor suppressors or oncogenes in the scientific literature will be discussed next, followed by discussion of genes without existing proof of tumor suppressor or oncogenic function.

Pathways enriched in recurrence predicting genes by mRNA expression include pathways critical in tumorigenesis and cancer progression such as the cytoskeleton remodeling, immune response, cell cycle, VEGF and androgen receptor signaling pathways (9). Some pathways enriched by recurrence predicting genes by CpG site methylation are major signaling pathways in HCC development, progression, and metastasis. These include IGF-1R, FGFR, PIP3, Wnt, signal transduction through AKT and PTEN, cytoskeleton remodeling process through TGF, Wnt, and cytoskeletal remodeling pathways (9). Interestingly, the cytoskeleton remodeling TGF, Wnt, and cytoskeletal remodeling pathway was commonly enriched in both recurrence predicting gene sets.

The Wnt and TGF β pathways are major signaling pathways in HCC (3, 9). It is currently unknown whether CpG site methylation of genes involved in these pathways regulates the activity of these signaling pathways. This should be further investigated in future studies.

Proline/serine-rich coiled-coil 1 (*PSRC1*) is a micro-tubule-associated oncoprotein that is transcriptionally down regulated by p53. Its promoter contains a putative p53 binding motif that is responsible for the p53 mediated repression of the gene (10). *PSRC1* protein promotes cell growth by several different mechanisms. It enhances β -catenin dependent transactivation and cyclin D1 production by binding to adenomatous polyposis coli 2 (*APC2*), a homolog of adenomatous polyposis coli (*APC*), which is a component of the β -catenin destruction complex (11). It also inhibits p53-binding protein 2 (*ASPP2*) induced apoptotic signaling of p53 (12). Over expression of *PSRC1* was observed in

hepatoma tissues compared to the adjacent benign tissue (10). The mechanism of *PSRC1* over-expression in HCC tissues is yet to be understood. Mutation of p53 may partly explain the up-regulation of *PSRC1* in hepatoma tissues. Our study has shown that hypomethylation of the *PSRC1* gene is associated with an increased risk for recurrence. Furthermore, high expression of the *PSRC1* gene was associated with shorter recurrence-free survival. This is the first report that increased expression of the *PSRC1* gene is associated with recurrence free survival in patients who underwent curative resection for HCC. The result also indirectly supports the concept that hypomethylation of *PSRC1* may contribute to the over-expression of *PSRC1* and tumor progression in HCC.

DNA methyltransferase 3 β (*DNMT3B*) is a de novo methyltransferase that catalyzes the methylation of the CpG islands of genes. *DNMT3B* is highly expressed in HCC, but rarely expressed in normal liver and increased expression of *DNMT3B* is associated with poorer overall survival in HCC patients (13). Hypomethylation of *DNMT3B* was associated with recurrence of HCC in our study. Our result also confirmed that increased expression of *DNMT3B* is associated with poor prognosis in HCC patients, although the validation analysis showed a similar trend without statistical significance ($P = 0.17$). The mechanistic role of *DNMT3B* promoter hypomethylation and increased expression of *DNMT3B* in liver carcinogenesis remains to be determined.

Tissue inhibitor of metalloproteinase 3 (*TIMP3*) is a well-known tumor suppressor gene in HCC. Tumor growth, invasion and metastatic tendency are restrained by the metalloproteinase inhibitory effect of *TIMP3* and aberrant methylation of *TIMP3* is observed in 25% of HCCs (14). In the training data set, hypermethylation and downregulation of *TIMP3* were associated with recurrence, but the association of gene expression with recurrence was not validated in the independent data set.

Cytoglobin (*CYGB*) is a recently discovered globin protein and its function is not yet completely established. Shivapurkar et al. (15) have shown that CpG site methylation of the *CYGB* promoter region is associated with down regulation of gene expression and methylation of the gene had an excellent ability to differentiate cancers of the breast, lung, bladder, and leukemia cells from benign tissues. Knockdown of the *CYGB* gene in a lung cancer cell line increased colony formation and transfection of *CYGB* led to decreased colony formation, suggesting a tumor suppressor role of *CYGB* (15). In the liver, *CYGB* is highly expressed in hepatic stellate cells and its expression is associated with hepatic fibrosis (16). *CYGB* appears to act as a scavenger of reactive oxygen species, preventing liver injury and resultant liver fibrosis by inhibiting free radical-induced hepatic stellate cell activation. Over expression of *CYGB* inhibits the differentiation of rat hepatic stellate cells into the myofibroblast-like phenotype, and protects against progressive liver damage and fibrosis (17). There has been no report of a role for *CYGB* in HCC pathogenesis as

yet. It is plausible that down regulation of *CYGB* enhances liver fibrosis and cirrhosis in patients with chronic hepatitis and allows the action of genotoxic reactive oxygen species, eventually contributing to the development of HCC. Although suppression of *CYGB* expression in HCC could not be validated due to the lack of a gene probe in the validation data set, this is the first study that suggests that CpG site methylation of *CYGB* and down regulation of *CYGB* are associated with the risk of recurrence after surgical resection for HCC.

Kallikrein-related peptidase 10 (*KLK10*) is a member of the human tissue kallikrein family of secreted serine proteases. *KLK10* is known to function as a tumor suppressor and loss of *KLK10* gene expression by CpG island methylation has been reported in breast, ovarian, and prostate cancer (18). *KLK10* methylation is an independent predictor of disease-free survival and overall survival in early-stage breast cancer patients (19). Lu et al. (20) recently reported that *KLK10* is frequently down-regulated by methylation in HCC. In their experiments, 20 to 40% of CpG sites in the *KLK10* gene were methylated in HCC cell lines compared to 2.1% of CpG sites in normal liver DNA. Hypermethylation of *KLK10* was observed in 55% of 49 HCC samples and there was an inverse correlation between *KLK10* mRNA level and methylation ($r = -0.435$, $P < 0.05$). Interestingly, transfection of the J7 cell line with *KLK10* decreased anchorage-independent growth, and transfection of Huh7 cells with *KLK10* increased the sensitivity of Huh7 cells to cytotoxic agents, suggesting a tumor suppressor effect of *KLK10* in HCC tissues. In our analysis, hypermethylation and down regulation of *KLK10* was associated with HCC recurrence in the training data set. However, expression of *KLK10* was not associated with recurrence in the validation data set. Determination of the true prognostic value of *KLK10* methylation in HCCs requires further validation.

Complement factor H (*CFH*) regulates alternate pathways of complement activation. A *CFH* polymorphism has a strong association with age-related macular degeneration. Mutation of *CFH* and resultant dysregulation of the complement system is also linked to atypical hemolytic uremic syndrome (21). The expression of *CFH* in adipose tissue correlates with insulin resistance (21). Although the role of *CFH* in cancer is not well understood, over-expression of *CFH* was associated with reduction of the metastatic and/or mortality risk ($P = 0.036$, HR 0.44) in soft tissue sarcomas (22). Our results suggest that *CFH* acts as a tumor suppressor and its expression may be down regulated by promoter methylation. The mechanism of action of *CFH* in HCC and other cancers remains to be understood.

Meiotic recombination 11 homolog A (*MRE11A*) is a nuclear protein involved in homologous recombination, maintenance of telomere length, and DNA double-strand break repair. *MRE11A* is an important component of the *MRE11* complex, which is involved in checkpoint activation and DNA repair to prevent chromosomal instability. Null mutation of *MRE11A* is known to be

lethal in higher eukaryotes and a truncating mutation of *MRE11A* is associated with an ataxia-telangiectasia-like disorder syndrome (23). In breast cancer, 7% of tumors showed aberrantly reduced protein expression for *MRE11A* and it has been suggested as a candidate familial cancer predisposing gene (24). Our result suggests that this gene acts as an oncogene which is contradictory to its presumed role in breast cancer.

Autophagy-related protein LC3 A (*MAP1LC3A*) is a protein located in the autophagosomal/autolysosomal membranes and antibodies directed against *MAP1LC3A* have been used to detect autophagic activity. There are three different patterns of autophagic expression. Of these, high expression of stone like structure *MAP1LC3A* was associated with poor prognosis in breast and endometrial cancers (25, 26). Our study could not provide specific information on the subcellular expression pattern of *MAP1LC3A*. Roles for *MRE11A* and *MAP1LC3A* in hepatocarcinogenesis have not been reported before and their potential regulation by methylation and functional roles in liver carcinogenesis remain to be more fully characterized.

Several genes that were selected in our study have not previously been reported as cancer-associated genes. Myosin 1E (*MYO1E*) is a class I myosin that is comprised of a single motor domain, a neck domain and a cargo-binding tail domain. *MYO1E* appears to be involved in receptor mediated endocytosis through interactions with synaptojanin-1 and dynamin (27). Myosin VIIA and Rab interacting protein (*MYRIP*) is a protein linking Rab27a and Myosin VIIa and involved in melanosome transport in the retinal pigment epithelium (28). *MYO1E* links protein kinase A to the exocytosis machinery for the coordination of cellular events and also plays an important role in hemostasis by regulating the release of vWF from endothelial cells (29). NADH dehydrogenase 1 alpha subcomplex assembly factor 1 (*NDUFA1*) is a mitochondrial chaperone protein involved in the assembly of the mitochondrial NADH (30). Functional roles for *MYRIP*, *MYO1E*, and *NDUFA1* have not been reported in cancer and potential mechanisms for their roles in HCC remain to be determined.

Our study has several weaknesses. Although we selected genes that are associated with recurrence by both mRNA expression and CpG site methylation in order to focus on genes showing methylation-mediated regulation of gene expression (concomitant hypomethylation and increased expression, or concomitant hypermethylation and decreased expression), the results do not provide direct evidence that promoter methylation of the selected genes regulates gene expression in HCC. The association between gene expression and recurrence was investigated in the validation data set, but due to a lack of corresponding information on genome-wide methylation in the validation data set, the association between CpG site methylation and recurrence could not be validated. Nevertheless the analysis identified a number of genes that have been previously reported as regulated by methylation in HCC, confirming its utility for iden-

tifying methylation regulated genes, and also identified potential novel oncogenes or tumor suppressor genes in HCC for further evaluation by functional studies. Some of the genes selected by regression analysis show significant differences in recurrence free survival (*MRE11A* & *CFH*), however others (*PSRC1*, *MYO1E*, *MYRIP*) show a trend towards a difference in survival without significance when expression/methylation patterns were dichotomized (Fig. 3). This may be due to loss of statistical power from the conversion of continuous predictors to binary predictors. Lastly few genes that were known to be oncogenes or tumor suppressor genes in the previous literature were not completely validated in our analysis. This might have been attributed to chance from multiple comparisons. However, it is still meaningful to discuss them as 1) they were shown to be associated with HCC recurrence both by gene expression and methylation that have minimized the false positive association: 2) these are biologically plausible.

In conclusion, our study suggests potential novel oncogenes (*PSRC1*, *MRE11A*, *MYO1E*), tumor suppressor genes (*CFH*, *MYRIP*), and their enriched pathways (cytoskeleton remodeling_TGF, Wnt, and cytoskeletal remodeling) implicated in hepatocarcinogenesis that may be regulated by CpG site methylation, and affect prognosis after resection for HCC. Several of the identified genes and pathways have been reported in other cancers but are novel in HCC and other genes and pathways associated with HCC were confirmed in our study. The regulation of individual genes by promoter methylation, their enriched pathways, and prognostic roles in HCC will need to be further validated using functional studies.

REFERENCES

1. Yang JD, Roberts LR. *Hepatocellular carcinoma: a global view. Nat Rev Gastroenterol Hepatol* 2010; 7: 448-58.
2. Jones PA, Baylin SB. *The fundamental role of epigenetic events in cancer. Nat Rev Genet* 2002; 3: 415-28.
3. Couluouan C, Factor VM, Thorgeirsson SS. *Transforming growth factor-beta gene expression signature in mouse hepatocytes predicts clinical outcome in human cancer. Hepatology* 2008; 47: 2059-67.
4. Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF, Mikaelian A, Roberts LR, Demetris AJ, Sun Z, Nevens F, Roskams T, Thorgeirsson SS. *A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med* 2006; 12: 410-6.
5. Arai E, Ushijima S, Gotoh M, Ojima H, Kosuge T, Hosoda F, Shibata T, Kondo T, Yokoi S, Imoto I, Inazawa J, Hirohashi S, Kanai Y. *Genome-wide DNA methylation profiles in liver tissue at the precancerous stage and in hepatocellular carcinoma. Int J Cancer* 2009; 125: 2854-62.
6. Shin SH, Kim BH, Jang JJ, Suh KS, Kang GH. *Identification of novel methylation markers in hepatocellular carcinoma using a methylation array. J Korean Med Sci* 2010; 25: 1152-9.
7. Yang JD, Sun Z, Hu C, Lai J, Dove R, Nakamura I, Lee JS, Thorgeirsson SS, Kang KJ, Chu IS, Roberts LR. *Sulfatase 1 and sulfatase 2 in hepato-*

- cellular carcinoma: associated signaling pathways, tumor phenotypes, and survival. *Genes Chromosomes Cancer* 2011; 50: 122-35.
8. Fan JB, Gunderson KL, Bibikova M, Yeakley JM, Chen J, Wickham Garcia E, Lebruska LL, Laurent M, Shen R, Barker D. *Illumina universal bead arrays. Methods Enzymol* 2006; 410: 57-73.
 9. Yang JD, Nakamura I, Roberts LR. *The tumor microenvironment in hepatocellular carcinoma: current status and therapeutic targets. Semin Cancer Biol* 2011; 21: 35-43.
 10. Hsieh WJ, Hsieh SC, Chen CC, Wang FF. *Human DDA3 is an oncoprotein down-regulated by p53 and DNA damage. Biochem Biophys Res Commun* 2008; 369: 567-72.
 11. Hsieh PC, Chang JC, Sun WT, Hsieh SC, Wang MC, Wang FF. *p53 downstream target DDA3 is a novel microtubule-associated protein that interacts with end-binding protein EB3 and activates beta-catenin pathway. Oncogene* 2007; 26: 4928-40.
 12. Sun WT, Hsieh PC, Chiang ML, Wang MC, Wang FF. *p53 target DDA3 binds ASP2 and inhibits its stimulation on p53-mediated BAX activation. Biochem Biophys Res Commun* 2008; 376: 395-8.
 13. Oh BK, Kim H, Park HJ, Shim YH, Choi J, Park C, Park YN. *DNA methyltransferase expression and DNA methylation in human hepatocellular carcinoma and their clinicopathological correlation. Int J Mol Med* 2007; 20: 65-73.
 14. Yuan Y, Wang J, Li J, Wang L, Li M, Yang Z, Zhang C, Dai JL. *Frequent epigenetic inactivation of spleen tyrosine kinase gene in human hepatocellular carcinoma. Clin Cancer Res* 2006; 12: 6687-95.
 15. Shivapurkar N, Stastny V, Okumura N, Girard L, Xie Y, Prinsen C, Thunnissen FB, Wistuba II, Czerniak B, Frenkel E, Roth JA, Liloglou T, Xinarianos G, Field JK, Minna JD, Gazdar AF. *Cytoglobin, the newest member of the globin family, functions as a tumor suppressor gene. Cancer Res* 2008; 68: 7448-56.
 16. Man KN, Philipsen S, Tan-Un KC. *Localization and expression pattern of cytoglobin in carbon tetrachloride-induced liver fibrosis. Toxicol Lett* 2008; 183: 36-44.
 17. Xu R, Harrison PM, Chen M, Li L, Tsui TY, Fung PC, Cheung PT, Wang G, Li H, Diao Y, Krissansen GW, Xu S, Farzaneh F. *Cytoglobin overexpression protects against damage-induced fibrosis. Mol Ther* 2006; 13: 1093-100.
 18. Sidiropoulos M, Pampalakis G, Sotiropoulou G, Katsaros D, Diamandis EP. *Downregulation of human kallikrein 10 (KLK10/NES1) by CpG island hypermethylation in breast, ovarian and prostate cancers. Tumour Biol* 2005; 26: 324-36.
 19. Kioulafa M, Kaklamanis L, Stathopoulos E, Mavroudis D, Georgoulas V, Lianidou ES. *Kallikrein 10 (KLK10) methylation as a novel prognostic biomarker in early breast cancer. Ann Oncol* 2009; 20: 1020-5.
 20. Lu CY, Hsieh SY, Lu YJ, Wu CS, Chen LC, Lo SJ, Wu CT, Chou MY, Huang TH, Chang YS. *Aberrant DNA methylation profile and frequent methylation of KLK10 and OXGR1 genes in hepatocellular carcinoma. Genes Chromosomes Cancer* 2009; 48: 1057-68.
 21. Moreno-Navarrete JM, Martínez-Barricarte R, Catalán V, Sabater M, Gómez-Ambrosi J, Ortega FJ, Ricart W, Blüher M, Frühbeck G, Rodríguez de Córdoba S, Fernández-Real JM. *Complement factor H is expressed in adipose tissue in association with insulin resistance. Diabetes* 2010; 59: 200-9.
 22. Linton KM, Hey Y, Saunders E, Jeziorska M, Denton J, Wilson CL, Swindell R, Dibben S, Miller CJ, Pepper SD, Radford JA, Freemont AJ. *Acquisition of biologically relevant gene expression data by Affymetrix microarray analysis of archival formalin-fixed paraffin-embedded tumours. Br J Cancer* 2008; 98: 1403-14.
 23. Stewart GS, Maser RS, Stankovic T, Bressan DA, Kaplan MI, Jaspers NG, Raams A, Byrd PJ, Petrini JH, Taylor AM. *The DNA double-strand break repair gene hMRE11 is mutated in individuals with an ataxia-telangiectasia-like disorder. Cell* 1999; 99: 577-87.
 24. Bartkova J, Tommiska J, Oplustilova L, Aaltonen K, Tamminen A, Heikinen T, Mistrik M, Aittomäki K, Blomqvist C, Heikkilä P, Lukas J, Nevanlinna H, Bartek J. *Aberrations of the MRE11-RAD50-NBS1 DNA damage sensor complex in human breast cancer: MRE11 as a candidate familial cancer-predisposing gene. Mol Oncol* 2008; 2: 296-316.
 25. Sivridis E, Koukourakis MI, Zois CE, Ledaki I, Ferguson DJ, Harris AL, Gatter KC, Giatromanolaki A. *LC3A-positive light microscopy detected patterns of autophagy and prognosis in operable breast carcinomas. Am J Pathol* 2010; 176: 2477-89.
 26. Sivridis E, Giatromanolaki A, Liberis V, Koukourakis MI. *Autophagy in endometrial carcinomas and prognostic relevance of 'stone-like' structures (SLS): what is destined for the atypical endometrial hyperplasia? Autophagy* 2011; 7: 74-82.
 27. Krendel M, Osterweil EK, Mooseker MS. *Myosin 1E interacts with synaptojanin-1 and dynamin and is involved in endocytosis. FEBS Lett* 2007; 581: 644-50.
 28. Klomp AE, Teofilo K, Legacki E, Williams DS. *Analysis of the linkage of MYRIP and MYO7A to melanosomes by RAB27A in retinal pigment epithelial cells. Cell Motil Cytoskeleton* 2007; 64: 474-87.
 29. Nightingale TD, Pattni K, Hume AN, Seabra MC, Cutler DF. *Rab27a and MyrIP regulate the amount and multimeric state of VWF released from endothelial cells. Blood* 2009; 113: 5010-8.
 30. Vogel RO, Janssen RJ, Ugalde C, Grovenstein M, Huijbens RJ, Visch HJ, van den Heuvel LP, Willems PH, Zeviani M, Smeitink JA, Nijtmans LG. *Human mitochondrial complex I assembly is mediated by NDUFAF1. FEBS J* 2005; 272: 5317-26.

AUTHOR SUMMARY**Genes Associated with Recurrence of Hepatocellular Carcinoma: Integrated Analysis by Gene Expression and Methylation Profiling**

Ju Dong Yang, So-Young Seol, Sun-Hee Leem, Yong Hoon Kim, Zhifu Sun, Ju-Seog Lee, Snorri S. Thorgeirsson, In-Sun Chu, Lewis R. Roberts and Koo Jeong Kang

DNA methylation is one of mechanisms for suppression of gene expression. Hypermethylation of tumor suppressor gene or hypomethylation of oncogene can contribute to cancer development. In this study, we identified genes whose methylation and mRNA expression are associated with recurrence after resection for hepatocellular carcinoma (HCC) using integrated analysis by whole genome DNA methylation and mRNA expression array data. TGF, Wnt, and cytoskeletal remodeling were major pathways enriched with the recurrence predicting genes.