Clinical and functional analysis of RAC GTPASE activating protein (RACGAP1) as predictive biomarker for recurrence of hepatocellular carcinoma

Wang, Suk Mei

2010

Wang, S. M. (2010). Clinical and functional analysis of RAC GTPASE activating protein (RACGAP1) as predictive biomarker for recurrence of hepatocellular carcinoma. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/42459>

<https://doi.org/10.32657/10356/42459>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**CLINICAL AND FUNCTIONAL ANALYSIS OF RAC
GTPASE ACTIVATING PROTEIN (RACGAP1) AS
PREDICTIVE BIOMARKER FOR RECURRENCE OF
HEPATOCELLULAR CARCINOMA**

WANG SUK MEI

WANG SUK MEI

SCHOOL OF CHEMICAL AND BIOMEDICAL ENGINEERING

2010

CLINICAL AND FUNCTIONAL ANALYSIS OF RACGAP1 AS
PREDICTIVE BIOMARKER FOR RECURRENCE OF
HEPATOCELLULAR CARCINOMA

2010

**CLINICAL AND FUNCTIONAL ANALYSIS OF RAC
GTPASE ACTIVATING PROTEIN (RACGAP1) AS
PREDICTIVE BIOMARKER FOR RECURRENCE OF
HEPATOCELLULAR CARCINOMA**

WANG SUK MEI

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2010

Acknowledgements

I would like to thank my technical advisor, Prof. Hui Kam Man for giving me the opportunity to work on this project, and his advice throughout my research study in National Cancer Centre of Singapore. I would like to express my heartiest appreciation and gratitude to my main supervisor, Associate Prof William Chen Wei Ning for his patience, generous guidance and full support in overcoming the difficulties throughout the period of preparing the thesis. I would also like to thank him for his invaluable advice and constant encouragement during the very hard time of my project. I am also thankful to my academic mentor, Associate Prof Phillip Cheang Hong Ning, for offering me the opportunity to be a part time postgraduate in Nanyang Technological University of Singapore, and his timely technical advice.

My deepest gratitudes are also accorded to Dr Mac Ho Meng Fatt, Miss Kam Siok Yuen and Miss Chua Hui Min for their invaluable advices and generous technical assistance in conquering some of the laborious techniques particularly immunohistochemistry during the hard time of my project. My sincere gratitudes are also extended to Dr Yu Kun for his helpful discussion on bioinformatics, Dr Puan Kia Joo for his technical guidance on flow cytometry analysis, Ms Bai Jing for her help in proteomic work and Dr Tan Poi Kiang for her critical comments on my project.

I would also like to thank Miss Shih Yao for her kind efforts in sourcing journals that are not subscribed, and her understanding and trust that allow me to stay late in the

library during my thesis writing-up. Special gratitudes to friends and colleagues (Ms Annie, Dr Lau Wen Min, Mr Sia Kian Chuan, Mr Derrick Low Thiah Hian and Mr Teoh Eng Seng) for their encouragement, and for providing comic relief especially during the frustrated moment when much experimental failures were encountered. I would like to thank National Cancer Centre of Singapore for providing me the infrastructure and facilities as required throughout my project, and also Nanyang Technological University, School of Chemical and Biomedical Engineering, without which the completion of this thesis would not be possible.

Last but not least, I would like to express my heartiest gratitudes to my family especially my husband (Chong Shii), parents and son for their understanding and invaluable constant support and encouragement, and without whom I would most certainly have sustained Permanent Head Damage.

Wang Suk Mei

Dec 2010

Table of Contents

Acknowledgements	I
Table of Contents	III
List of Tables	VIII
List of Figures	X
List of Abbreviations	XII
Publications	XIV
Summary	XV

CHAPTER 1	Introduction and Literature Review	1
1.1	The Liver	1
1.1.1	Microstructure and vascular supply	1
1.1.2	Segmental anatomy	2
1.1.3	Cellular features	4
1.2	Hepatocellular carcinoma	5
1.2.1	Pathology of Precancerous changes	5
1.2.1.1	Chronic hepatitis, fibrosis and cirrhosis	5
1.2.1.2	Dysplasia and cancer	8
1.2.2	Epidemiology and Etiology	10
1.2.3	Early detection and diagnosis	13
1.2.4	Staging and treatment	14
1.3	Recurrence of HCC after curative resection	17
1.4	Risk factors for HCC recurrence	19
1.4.1	Viral Hepatocarcinogenesis	19
1.4.2	Cirrhosis	20
1.4.3	Vascular invasion	21
1.5	Gene Expression Profiling in cancer	22
1.6	Support vector machine	24
1.7	Identification of Rac GTPase Activating Protein 1 (RACGAP1) as a HCC recurrence biomarker	26
1.7.1	RACGAP1 plays a role in cell differentiation	26
1.7.2	RACGAP1 plays a role in regulating cytokinesis	26
1.7.3	RACGAP1 in other cancer databases or other aspects	28

1.8	Proteomics	28
1.8.1	Principle of mass spectrometry-based proteomics	29
1.8.2	Protein profiling studies with proteomics approach	32
1.9	Objective and approaches for identifying molecular markers for HCC recurrence	33
CHAPTER 2	Materials and Methods	34
2.1	Methods	34
2.1.1	Patient samples procurement	34
2.1.2	Total RNA isolation and quantitation	35
2.1.3	Affymetrix GeneChip® hybridization	36
2.1.4	Microarray data pre-processing and analysis	37
2.1.5	Statistical and computational analysis	38
2.1.6	Reverse transcription	40
2.1.7	Polymerase chain reaction (PCR)	41
2.1.8	Real-time PCR quantitation	42
2.1.9	Immunostaining of RACGAP1	43
2.1.10	Generation of pIRES2-RACGAP1-AcGFP1 plasmid construct	44
2.1.11	Bacterial transformation and colony screening	46
2.1.12	Plasmid DNA isolation	47
2.1.13	Cell culture	48
2.1.14	DNA transfection	49
2.1.15	siRNA transfection	49
2.1.16	Western blot analysis	50
2.1.17	Cell migration and invasion assays	51
2.1.18	Cdc42, Rac1 and Rho A activation assay	52
2.1.19	DNA fragmentation (TUNEL) assay by FACS and microscopy	53
2.1.20	FACS analysis of Ki-67 proliferation assay	54
2.1.21	Network Analysis of gene expression in cells depleting RACGAP1	54
2.1.22	Protein profiling via LC-MS/MS analysis of cells depleting RACGAP1	55
2.2	Primers sequences for PCR and Realtime PCR	58
2.3	siRNA sequences	59
2.4	Reagents and antibodies	60
2.4.1	General buffers	60
2.4.2	Reagents and materials for Affymetrix GeneChip experiment	61
2.4.3	Buffers for Immunohistochemistry	63
2.4.4	Reagents and materials for cloning and bacteria transformation	64
2.4.5	Buffers for protein analysis	65
2.4.6	Buffers for Flow Cytometry (FACS)	67
2.4.7	Antibodies	67

CHAPTER 3	Identification of molecular markers associated with recurrence of human hepatocellular carcinoma by microarray analysis	70
3.1	Results	70
3.1.1	Clinical characteristics of HCC samples	70
3.1.2	Vascular invasion and cirrhosis are associated with HCC recurrence	71
3.1.3	Stratification of HCC patients to recurrence and non-recurrence groups using either vascular invasion or cirrhosis at a time	74
3.1.4	Stratification of HCC patients using the combined vascular invasion and cirrhosis reveals various recurrence risk groups	75
3.1.5	Prediction of recurrence is insufficient using gene signature derived from a high risk clinical factor analyzed at a time	78
3.1.6	Prediction of recurrence in group 2 and 3 using gene signature derived from the same group by combining high risk clinical factors	80
3.1.7	Kaplan Meier analysis of 57-gene set prediction reveals poorer prognosis in predicted recurrent than non-recurrent patients from both group 2 and 3	82
3.1.8	The 57-gene set is specifically applicable in predicting HCC recurrence of patients from both group 2 and 3	83
3.1.9	Performance comparison of 65-gene set predictor of group 1 and 4 versus the 57-gene set predictor of group 2 and 3	85
3.1.10	Comparison analysis of two HCC recurrence-related gene sets	87
3.1.11	Biological and functional comparison analysis of two HCC recurrence-related gene sets	90
3.1.12	Hierarchical clustering and functional annotation of two HCC recurrence gene set predictors	93
3.1.13	Verification of genes selected from two gene set predictors by quantitative real-time PCR	99

3.2	Discussion	102
3.2.1	Cross assessment of prediction ability of public HCC recurrence-related gene expression data using clinical samples in this study	102
3.2.2	Comparison of public HCC recurrence-related gene expression data with the prediction signatures in this study	103
3.2.3	Differential gene expression to predict recurrence of various risk groups	105
3.2.4	Comparative functional annotation of different gene sets	107
CHAPTER 4	Rac GTPase activating protein 1 (RACGAP1) up-regulation as a prognostic indicator for the recurrence of hepatocellular carcinoma	111
4.1	Results	111
4.1.1	Significant overexpression of RACGAP1 in HCC recurrence within 2 years	111
4.1.2	Immunostaining validated RACGAP1 upregulation in HCC recurrent samples	112
4.1.3	RACGAP1 overexpression is a significant independent risk factor for HCC recurrence	115
4.1.4	RACGAP1 expression has prognostic indication for disease recurrence	116
4.1.5	Endogenous RACGAP1 expression in human hepatocellular carcinoma and other cancer cell lines	118
4.1.6	siRNA-mediated silencing of RACGAP1 in metastatic HCC cell lines	120
4.1.7	Silencing of RACGAP1 results in suppressed migration and invasion in HCC cells	121
4.1.8	Silencing of RACGAP1 promotes cell death in HCC cells	123
4.1.9	Silencing of RACGAP1 represses Cdc42 and Rac1 rather than RhoA activation	125
4.1.10	Exogenous overexpression of RACGAP1 results in increased cell proliferation in PLC/PRF 5 cells	128
4.1.11	Exogenous overexpression of RACGAP1 enhances cell migration and invasion	129

4.1.12	Exogenous overexpression of RACGAP1 activates Cdc42, Rac1 and RhoA	131
4.1.13	Genome-wide analysis of siRNA-mediated silencing of RACGAP1 in MHCC97-H cells	133
4.1.14	Functional analysis of differential genes after knockdown of RACGAP1 in MHCC97-H cells	134
4.1.15	Comparative functional analysis of cell-based depletion and clinically overexpression of RACGAP1	136
4.1.16	Interaction of the RACGAP1 signaling networks and pathways that are closely associated with cell motility	137
4.1.17	Validation of deregulated molecules in response to RACGAP1 silencing	139
4.1.18	LC-MS/MS analysis of RACGAP1-depleting MHCC97-H cells	140
4.2	Discussion	142
4.2.1	Clinical significance of RACGAP1 up-regulation in HCC Recurrence	142
4.2.2	Up-regulation of RACGAP1 has prognostic indication for HCC recurrence within 2 years	143
4.2.3	Functional analysis of RACGAP1 via knockdown and overexpression experiments using cell-based approach	145
4.2.4	Pathway and network analysis of RACGAP1 loss-of-function in vitro reveals mitotic polo-like-kinase is the most affected signaling pathway	148
4.2.5	Validation of perturbed molecules in response to RACGAP1 suppression by western blot and LC-MS/MS analysis	150
CHAPTER 5	Conclusion and Future Perspectives	152
REFERENCES		156

List of Tables

Table 2.1. Protocol for preparing SDS-PAGE gels	66
Table 3-1. Univariate analyses to evaluate the correlation between clinicopathological factors and disease recurrence of 80 HCC patients	72
Table 3-2. Multivariate analyses to assess the association of clinicopathological parameters with HCC recurrence occurred at different duration after Resection	73
Table 3-3. Performance of various gene sets, derived by considering different clinicopathological factors in predicting recurrence of HCC	79
Table 3-4. Performance of the 57-member gene signature in predicting disease recurrence for HCC patients in combined group 2 and 3	82
Table 3-5. Multivariate analyses to determine if prediction performance of the 57-gene signature is confounded by any clinical factors	84
Table 3-6. Performance of the 57-member gene signature in predicting recurrence for HCC patients in group 1 and group 4 respectively	85
Table 3-7. Performance comparison of 2 gene set predictors in their corresponding training and test samples using different classifiers	87
Table 3-8. Fourteen commonly overexpressed genes in both gene list A and B	89
Table 3-9. Thirteen commonly underexpressed genes in both gene list A and B	89
Table 3-10. Eight genes with inversed expression between gene list A and B	90
Table 3-11. The top 30 significant functional annotation of 57 genes	95
Table 3-12. The top 30 significant functional annotation of 65 genes	98
Table 4-1. Immunohistochemistry staining of RACGAP1 in all individual HCC recurrence and non-recurrence samples	114
Table 4-2. Univariate and multivariate analyses to examine the association of clinicopathological factors and RACGAP1 expression with HCC recurrence	116

Table 4-3. Top 10 deregulated canonical pathways in RACGAP1-deficient MHCC97-H cells	135
Table 4-4. Differential molecules consistently detected by microarray and LC-MS/MS approaches in RACGAP1-depleting MHCC97-H cells	141

List of Figures

Figure 1-1. Anatomy and microstructure of the liver	2
Figure 1-2. Functional segmentation of liver according to Couinaud classification	3
Figure 1-3. Diagram of the chief cellular features in the liver	4
Figure 1-4. Histopathological progression of HCC	8
Figure 1-5. Barcelona Clinic Liver Cancer (BCLC) staging and treatment schedule	15
Figure 2-1. Restriction map and multiple cloning site of pIRES2-AcGFP1 vector	65
Figure 3-1. Stratification of recurrence and recurrence-free cases according to one clinical factor, invasion or cirrhosis at a time	75
Figure 3-2. Stratification of HCC patients to recurrence and non-recurrence guided by combined clinical factors	77
Figure 3-3. Kaplan-Meier analysis of the prediction outcomes of HCC recurrence	83
Figure 3-4. Multi-step filtering approach generates 2 optimized gene sets to predict 2 combination groups of HCC patients	86
Figure 3-5. Venn diagram showing comparison of the two gene sets	88
Figure 3-6. Functional classification of 2 sets of HCC recurrence-related gene signature	92
Figure 3-7. Gene expression pattern of 57 gene set	94
Figure 3-8. Gene expression pattern of 65 gene set	97
Figure 3-9. Validation of 7 genes selected from 57 gene set predictor	100
Figure 3-10. Validation of 6 genes selected from 65 gene set predictor	101
Figure 3-11. Overview of the findings of two differential gene set predictors	107
Figure 4-1. Significant up-regulation of RACGAP1 in HCC recurrence within 2 years	112

Figure 4-2. Immunostaining validates clinically significant overexpression of RACGAP1 in HCC recurrence	113
Figure 4-3. Kaplan-Meier recurrence-free survival analysis of HCC patients according to RACGAP1 expression	117
Figure 4-4. RACGAP1 endogenous expression in HCC and other cancer cell lines	119
Figure 4-5. Time course analysis of siRNA-mediated silencing effect on RACGAP1 protein level in HCCLM3 and MHCC97-H cells	121
Figure 4-6. Suppressed cell migration and invasion after siRNA-mediated depletion of RACGAP1 <i>in vitro</i>	122
Figure 4-7. TUNEL assay after siRNA-mediated silencing of RACGAP1	124
Figure 4-8. Deactivation of small GTPase pull-down assay after silencing of RACGAP1	127
Figure 4-9. Overexpression of RACGAP1 promotes cell proliferation in PLC/PRF 5 cells	129
Figure 4-10. Overexpression of RACGAP1 promotes cell migration and invasion in PLC/PRF 5 cells	130
Figure 4-11. Overexpression of RACGAP1 activates small GTPases in PLC/PRF 5 cells	132
Figure 4-12. Gene expression profiling of siRNA-mediated RACGAP1-depleting cells	134
Figure 4-13. Comparison of aberrant signaling networks resulted from cell based-depletion of RACGAP1 and clinical up-regulation of RACGAP1 in HCC recurrence	137
Figure 4-14. Interaction of the altered RACGAP1 signaling network and the Intermediate regulators or pathways closely related to cell motility	138
Figure 4-15. Verification of the altered target molecules upon silencing of RACGAP1 by western blot	140
Figure 4-16. Summary of the significant findings of RACGAP1 in both clinical and cell-based assays	147

List of Abbreviations

µg	microgram
µl	microlitre
µM	micromolar
AFP	alpha-fetoprotein
AKT	v-akt murine thymoma viral oncogene homolog
AURKB (ARK-2)	aurora kinase B
BRCA1	breast cancer 1, early onset
BSA	bovine serum albumin
Cdc42	Cell division control protein 42 homolog
cDNA	complementary deoxyribonucleic acid
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CO ₂	carbon dioxide
cRNA	complementary ribonucleic acid
CT	computerised tomography
DAB	3,3' Diaminobenzidine
DAPI	4',6-Diamidino-2-phenylindole
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagle's Medium
DNA /RNA	deoxyribonucleic / ribonucleic acid
dNTPs	deoxyribonucleotide triphosphates
ECT2	Epithelial cell transforming sequence 2 oncogene
EDTA	ethylenediaminetetraacetate
Erk1/Erk2	p44/42 MAP kinase
FACS	fluorescence-activated cell sorting
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GDP	Guanosine diphosphate
GTP	Guanosine-5'-triphosphate
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HRP	Horseradish Peroxidase
ICAT	Isotopo-coded affinity tags
KIF23 (MKLP-1)	kinesin family member 23
KNN	K Nearest Neighbour
LB	Luria Bertani / lysogeny broth
LC-MS/MS	High-performance Liquid Chromatography with tandem Mass Spectrometry
LFT	liver function tests
MES	2-(N-morpholino)ethanesulfonic acid
MgCl ₂	Magnesium chloride
min	minutes
ml	millilitre

mM	millimolar
nm	nanometer
PARP	Poly (ADP-ribose) polymerase
PBS	phosphate buffer saline
PCR	Polymerase chain reaction
PI	propidium iodide
PLK1	polo-like kinase 1 (<i>Drosophila</i>)
PLK4	polo-like kinase 4 (<i>Drosophila</i>)
PRC1	protein regulator of cytokinesis 1
Rac1	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)
RACGAP1	Rac GTPase activating protein 1
Rho(A, B, C)	ras homolog gene family, member A, B, C
RIN	RNA integrity number
RNAi	RNA interference
rpm	revolutions per minute
SAPE	streptavidin phycoerythrin
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
sec	seconds
siRNA	short interfering RNA
SOC	Super Optimal Broth with Catabolite repression
SPLDA	Sparse Linear Discriminant Analyst
SPSS	STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES
SSPE	salt sodium phosphate EDTA
SVM	Support Vector Machine
TBE	tris borate EDTA
TdT	terminal deoxynucleotide transferase
TUNEL	terminal deoxynucleotide transferase dUTP nick end labeling

Publications

1. **Wang SM**, Ooi LL, Hui KM. Identification of RACGAP1 as a prognostic biomarker associated with recurrence of human hepatocellular carcinoma. *Manuscript in preparation for publication Oct 2010.*
2. Ong HT*, **Wang SM***, Tan PK, Low DTH, Ooi LL, Hui KM. The tumor suppressor function of LECT-2 in human hepatocellular carcinoma makes it a potential therapeutic target. *Cancer Gene Ther.*2010. *Being reviewed Oct 2010* (*, joint first author).
3. Tan MG, Kumarasinghe MP, **Wang SM**, Ooi LL, Aw SE, Hui KM. Modulation of iron-regulatory genes in human hepatocellular carcinoma and its physiological consequences. *Exp Biol Med (Maywood).*2009;234(6):693-702.
4. **Wang SM**, Ooi LL, Hui KM. Identification and validation of a novel gene signature Associated with the recurrence of human hepatocellular carcinoma. *Clin Cancer Res* 2007;13:6275-83. *(Publication arising from thesis work)*
5. Lam P, Sian Lim K, **Wang SM**, Hui KM. A microarray study to characterize the molecular mechanism of TIMP-3-mediated tumor rejection. *Mol Ther* 2005;12:144-52.
6. Linn YC, **Wang SM**, Hui KM. Comparative gene expression profiling of cytokine-induced killer cells in response to acute myloid leukemic and acute lymphoblastic leukemic stimulators using oligonucleotide arrays. *Exp Hematol* 2005;33:671-81.
7. Leo JC, **Wang SM**, Guo CH, Aw SE, Zhao Y, Li JM, Hui KM, Lin VC. Gene regulation profile reveals consistent anticancer properties of progesterone in hormone-independent breast cancer cells transfected with progesterone receptor. *Int J Cancer* 2005;117:561-8.

Summary

Hepatocellular carcinoma (HCC) is the sixth most common malignancy and the third most common cause of cancer-related death worldwide. Although early HCC is potentially curable by partial hepatectomy, postresection prognosis remains poor because of the high recurrence rate, as the overall 5-year survival rate is lower than 5%. The heterogeneous nature of human HCC has also limited the usefulness of conventional clinicopathological features available at diagnosis for both treatment and prediction of disease outcome. Hence, better understanding of the clinicopathological features and molecular aspects of HCC to assess and stratify patients with different risks of disease recurrence at diagnosis would be extremely beneficial for the clinical management of HCC. In this study, we first attempt to systematically analyze and combine some of the conventional clinicopathological features available at diagnosis to segregate HCC patients with different risk of developing recurrence. We also employ molecular expression profiling to identify novel differentially regulated genes closely associated with recurrent disease to improve the prediction of HCC recurrence at diagnosis especially for patients who have 50% risk of developing disease recurrence.

Rac GTPase activating protein 1, identified as one of the member in our prediction gene set, was originally characterized as a human chimaerin-like protein in male germ cells. *In vitro*, the GAP domain strongly stimulates Rac1 and Cdc42 GTPase but is almost inactive on RhoA in mammalian cells. However, inverse result is found in *Xenopus laevis* system in which it acts primarily on Rho rather than Rac or Cdc42 to assemble the

cytokinetic apparatus. The alteration of RACGAP1 together with their neighbourhood targets particularly PLK1, ANLN, ECT2, KIF23, and PRC1, were detected as part of the gene member with elevated expression in other cancer databases including epithelial ovarian cancer, cervical cancer, high grade breast cancer with transition from preinvasive to invasive stage as well as ER positive breast cancer. Although much has been reported about RACGAP1 in regulating furrow ingression for completion of cytokinesis via interaction with Rac1, Cdc42 and Rho and mitosis in hematopoietic cells, its role in solid tumor progression, particularly HCC recurrence is still not explored. Our data revealed for the first time that RACGAP1 is significantly and consistently overexpressed in HCC with aggressive tumor behaviour and its expression has distinct prognostic indication for disease recurrence within 2 years after resection. RACGAP1 contributes to cell invasion and proliferation that are prerequisite conditions favoring HCC tumor cell survival and metastases by modulation of Rho GTPase molecules such as Rac1 and Cdc42. We depicted that the mitotic roles of polo-like kinase signaling pathway, the most perturbed pathways attributed by RACGAP1 deregulation, contributed to HCC recurrence. We also integrated the phenotypic, genome wide and proteomic analysis, so as to combine strengths from individual approach to provide robust cross validation for the data and observation obtained. Although the proteomic analysis in this study is preliminary and has little impact on the overall findings achieved from the microarray experiments, further efforts are needed to fine-tune this technique to be equivalent mature as the microarray analysis.

CHAPTER 1 Introduction and Literature Review

1.1 The liver

1.1.1 Microstructure and vascular supply

The liver is the largest organ and the largest mass of glandular tissue, positioned superiorly in the abdominal cavity of the body. It is engaged in a wide range of metabolic activities necessary for homeostasis, nutrition, detoxification and immune defence. It receives major blood supply (about 75%) from the hepatic portal vein, which carries oxygen-depleted venous blood containing substrates absorbed from the digestive tube, pancreas, and spleen. The hepatic artery carries the remaining 25% of the blood supply that are oxygenated to the liver (1). The liver is important in removal and breakdown of toxic materials from the blood stream, due to the existence of phagocytic macrophages which form part of the mononuclear phagocyte system of the body. Some of these degraded waste products are carried in the exocrine secretion of the liver, called bile and drained to the bile canaliculi that join to form bile ducts, which later fuse to form hepatic duct before delivering back to the duodenum for disposal. The portal vein and hepatic artery which branch to supply the sinusoids, and along with the bile duct system, make up the portal triad, the same site at which the lymphatic vessels leave the liver. The sinusoids bathe the hepatocytes for the exchange of oxygen, metabolites, nutrients and toxic substances between the blood and the liver cells, before draining the portal and arterial blood mixture to the central vein (Figure 1-1B).

Figure 1-1. Anatomy and microstructure of the liver. Adapted from (1, 2) with permission from Johns Hopkins Medicine, Division of Gastroenterology and Romrell LJ.

1.1.2 Segmental anatomy

The liver has four lobes or eight segments, depending on whether its anatomy is classified by its gross morphological or functional definition. The traditional morphological definition, based on the gross macro structure of the liver, divides it into right, quadrate, caudate and left lobes. The Couinaud segmental classification divides the liver into eight independent functional segments, each with its self-contained bile duct, hepatic artery and portal vein in the centre and vascular outflow through the hepatic veins in the periphery. Hence, this classification is of obvious importance in hepatic surgery as each segment can be resected along with the vessels that define the periphery without damaging those remaining. Figure 1-2 illustrates the Couinaud classification of the liver which follows a clockwise manner.

Figure 1-2. Functional segmentation of liver according to Couinaud classification. Adapted from (3) with permission from Elsevier.

Segment I or caudate lobe is anatomically unique due to its direct connections to the inferior vena cava through hepatic veins though it is smallest in size and hardly visible on a frontal view because of its posterior location. The left and right lobes were functionally divided by the centre hepatic vein. The left lobe is further subdivided into a medial (segment IV) and lateral section (segments II and III) by the left hepatic vein, while the right lobe is subdivided into anterior (segments V and VIII) and posterior section (segments VI and VII) by the right hepatic vein. Segment IV is usually divided into segment IVa and IVb according to Bismuth classification (4). The left and right branches of portal vein further divide the liver transversely into superior and inferior segments (5). As a results, this give rise to left superior segments (segments II and IVA), left inferior segments (segments III and IVB), right superior segments (segments VII and VIII) and right inferior segments (segments V and VI).

1.1.3 Cellular features

Cellular components of the liver consist of hepatocytes, stellate cells, endothelial sinusoidal cells and Kupffer cells (3) as shown in Figure 1-3. Hepatocytes (parenchymal epithelial cells) constitute about 80% of the liver volume and 60% of its cell number. Hepatocytes are large polygonal cells typically featured by high metabolic activity including the presence of endoplasmic reticulum, many mitochondria, lysosomes, golgi apparatus and large preoxisomes in their cytoplasm. Glycogen granules and lipid vacuoles are usually prominent. Their role in iron metabolism is shown by the presence of ferritin and haemosiderin vacuoles. The lateral plasma membranes of adjacent hepatocytes form the bile canaliculi.

Figure 1-3. Diagram of the chief cellular features in the liver. Adapted from (3) with permission from Elsevier.

Hepatic stellate cells (HSCs), the mesenchymal in origin cells are also called Ito cells or perisinusoidal lipocytes. They are much less numerous than hepatocytes, and

during quiescent stage, situated at the space of Disse in the liver that is lined by hepatocytes and sinusoidal cells (6). They present numerous cytoplasmic lipid granules containing retinyl palmitate (7) and are significant source of growth factors for liver homeostasis and regeneration. The quiescent HSCs are activated and lose their vitamin A stores to become myofibroblast-like in response to liver injury. Hepatic venous sinusoids are lined by flattened sinusoidal endothelial cells. Kupffer cells are hepatic macrophages derived from circulating blood monocytes and originate in the bone marrow. They constitute the major part of the mononuclear phagocyte system, involved in removing cellular and microbial debris from circulation and secreting cytokines for defense. They also share a function with the spleen by removing aged and damaged red cells from the hepatic circulation.

1.2 Hepatocellular carcinoma

1.2.1 Pathology of liver precancerous changes

1.2.1.1 Chronic hepatitis, fibrosis and cirrhosis

In contrast to acute hepatitis with massive apoptosis of hepatocytes (8, 9), chronic hepatitis B presents varying degree of inflammation with necrotic rather than apoptotic hepatocytes prevails as the major mode of cell death (9-12). Inflammation is less prominent in the immune-tolerant or inactive phase, but become pronounced in the immune-mediated viral clearance phase (9). The chronic inflammation process, involved repeated cycles of necrosis and proliferation of hepatocytes (wound healing

reaction). To replace damaged hepatocytes, the hepatocyte stellate cells are activated to secrete extra-cellular matrix protein or collagen type III (reticular) fibers that form collagenous scar tissue which lead to hepatic fibrosis around central vein (13) as shown in Figure 1-4. The fibrosis stage increases with the degree of bringing necrosis that connects central to portal. The chronic hepatitis pattern of injury is also seen in chronic hepatitis C, autoimmune hepatitis, drug-induced liver injury, and chronic cholestasis. Steatosis, which may be caused by the metabolic factors or indirectly by HBV virus, develops more frequently in chronic hepatitis B than in the general population (14).

In both chronic hepatitis B and C, the inflammatory infiltrates compose of CD4-positive T cells predominating over CD8-positive T cells, 15% of which are B cells while 10% are natural killer (NK) cells (15). During liver pathological process, the HBV-specific CD8 response that normally helps to clear acute infection, not only becomes defective in controlling virus replication, but also recruits non-virus CD8 T cells (16) that may hinder antiviral effector function and exert the proinflammatory response (17). Cytokine production by the liver-infiltrative lymphocytes may functionally differ from those by peripheral blood lymphocytes, thus favor the HBV persistence in the absence of liver damage (18, 19). The remarkable decrease of NK cells and CD4/CD8 ratio during advanced fibrosis, and the interactions between hepatic stellate cells with specific lymphocyte subsets may contribute significantly to the pathogenesis of fibrosis (18) which eventually lead to cirrhosis.

Cirrhosis, an intermediate predisposing lesion for HCC, is typically featured by co-existing regenerative nodules, irreversible fibrosis and severe liver injury (20, 21) where liver parenchymal architecture and blood flow pattern are destroyed. Cirrhosis is a diffuse form of hepatic advanced fibrosis, which results from severe prolonged chronic hepatitis with extensive deposition of fibrous septa in the hepatocellular necrosis area and vascular disorder (22, 23) which eventually give rise to abnormal liver nodules (Figure 1-4). Cirrhotic hepatocytes are more resistant to TGF- β -induced apoptosis than normal hepatocytes (12). This anti-apoptotic phenotype renders the balance skewed to necrotic death (24, 25), thus explains the massive hepatic necrosis during chronic liver disease. The underlying mechanism may involve the reduction of caspase-8 function, MAP kinases such as PI-3/Akt kinase, ERK and p38 signaling, repression of NF- κ B-mediated genes signaling and bile acid importers. There is a low frequency of healthy carriers or patients with chronic hepatitis that develop HCC without cirrhosis

Figure 1-4. Histopathological progression of HCC. Adapted from (26) with permission from Nature Publishing Group.

1.2.1.2 Dysplasia and cancer

The continuous rounds of hepatocytes damage, regeneration in chronic liver injury provoke the production of monoclonal populations that give rise to dysplastic hepatocytes, and later evolve to premalignant dysplastic nodules (27). Liver cell dysplasia is defined as a group of cells displaying cytologic abnormalities found significantly in HBV-related cirrhotic nodules that harbour HCC (28). It can be categorized into “large cell change” and “small cell change”, based on its distinguished histological pattern. The former lesion is characterized by a foci of enlarged cell with pleomorphic nuclei including hyperchromasia and multinucleation, and is usually associated with chronic hepatitis, cirrhosis (9), and also a higher

cumulative probability of HCC development (29). The latter is identified as foci of dense small hepatocytes with high nucleus to cytoplasmic ratio, increased proliferative activity, and genetic resemblance to HCC (9). The large cell change lesion is referred as “Dysplasia foci” if the size measuring less than 1 mm, and termed “nodules” if the size becomes greater than 1 mm. Cirrhotic nodules in which dysplasia and HCC develop typically enlarge up to 8 to 10 mm in diameter, due to its high rate of cell proliferation. Similar nodules with or without dysplasia may arise in non-cirrhotic liver.

Early HCC contains small number of malignant or dysplastic cells that may not form nodules, but rather resemble foci (30) which proliferate slowly. At a more advanced time, the malignant phenotypes of the nodules including neovessel formation nourished by hepatic artery and trabeculae thickening are evident (26). After further molecular disruption for instance genomic instability or loss of p53 (Figure 1-4), these high grade dysplastic nodules become susceptible to HCC development in 30% of the cases within 1 to 5 years (31). At early stage, the well-differentiated tumors are usually actively proliferating until their size reach 1 to 1.5 cm, they become less differentiated and later undifferentiated (32). At this point, tumor cells are capable to invade the surrounding fibrous stroma and vessels, and occasionally spread to extrahepatic tissues, leading to angiogenesis, tissue invasion and metastasis which are typical characteristics of aggressive end-stage disease (33, 34), and these happen in 25% of the incidences.

1.2.2 Epidemiology and Etiology

Of all primary liver malignant tumors, hepatocellular carcinoma (HCC) constitutes over 95% of the cases followed by cholangiocarcinoma and other rare cases such as hepatoblastoma (childhood liver cancer), embryonal sarcoma, malignant hemangioendothelioma and angiosarcoma. Hepatocellular carcinoma (HCC) is a type of neoplasia arises from hepatocytes and is characterized by a trabecular growth pattern with vascular invasion (35, 36). HCC ranks the fifth among the frequent cancers and the third among the frequent cause of cancer mortality worldwide (37). In Singapore, between 2003 to 2007, HCC ranks the fourth most frequent cancers and the third most frequent cause of cancer mortality among male group, but not in either report among female group (38). The geographic distribution of HCC is highly variable. High incidence areas include Eastern Asia and sub-Saharan Africa (36). Japan, Hong Kong and Shanghai have always been the most prevalent regions in Asia, with incidence varies between 23 and 45 patients per 100,000 per year (Figure 1) (39). In low incidence areas include Europe and the USA, HCC cases have increased progressively in the past decade (35, 40). Between 1990-1991 and 2005, the death rates for liver cancer has largely increased in both men (47.2%) and women (27.1%) in the United States (41). This is frequently attributed to the rising incidence of hepatitis C virus infection within these two decades, probably rival the currently reported in Japan (42). In addition, the immigrants from endemic region, for instance Vietnamese men become the leading group, while women second only to Korean

having by far the incidence and death rates from liver cancer, due to the increased prevalence of HBV infection in Vietnam (43).

The etiology of HCC is multifactorial, explaining the widely varying geographic incidences of this cancer worldwide. Both chronic hepatitis B (HBV) or C virus (HCV) infection are traditionally known to be the extremely associated with HCC development (44). Over 85% of HCC patients in the world carry markers of HBV and/or HCV (45). HBV and HCV infections are the predominant cause of cirrhosis that ultimately leads to HCC (46). The endemic presence of HBV infection in Asia and Africa reflect the high occurrence rate of HCC in these regions. The rate of developing HCC in chronic HBV carriers is 100-fold higher than those without HBV (47). Although HCV is lower prevalent than HBV, it causes more HCC in economically developed regions including developed Asian countries (36). In Japan, the risk of HCC development in HCV carriers with liver cirrhosis (5-7% per year) is between 2 to 3 times higher than HBV carriers (2.5-3% per year) (48, 49). HCV infection was shown to have higher carcinogenic potential than HBV as the development of HCC occurs most frequently and is significantly higher in patients with HCV than those with HBV infection (49). However, other study proposes that HBV-related HCCs have a poorer prognosis or more aggressive than HCV-related tumors, and this becomes significant among patients with advanced HCC (50). Several studies in human immunodeficiency virus (HIV) infection that aggravates HCV infection may explain the increase of HCC cases especially in the developed countries (51, 52).

High exposure to food contaminated with aflatoxin B₁ (derived from fungi *Aspergillus flavus* and *Aspergillus parasiticus* which grow on improperly stored grains), a problem in China and Africa also serves as an aetiological factor in the origin of HCC. In addition, other frequent etiologies of cirrhosis associated with risk of HCC development are genetic disorders including type I tyrosinemia and hemochromatosis (HHC) that lead to chronic liver disease (CLD). These two genetic diseases have higher risk for developing HCC (17) than other CLD for instance alcoholic liver disease (ALD), primary biliary cirrhosis (PBC) and Wilson's disease (WD) (51-57), though they all share common lesions including fibrosis, inflammation and cirrhosis. Obesity and diabetes (26, 58) are also linked to HCC development, as higher HCC incidence has been reported in these patients who have no history of liver disease. They may share the common molecular mechanism as both are associated to lipid accumulation in the liver and insulin resistance. Lipid accumulation in the liver of obese patients activates inflammatory response which increases IL-6 and TNF production by adipocytes and Kupffer cells (59). These two molecules in turn activate STAT3 and NF- κ B which induce proliferation of the damaged hepatocytes and ultimately lead to HCC. Similarly, accumulation of lipid in the liver of diabetic patients also promotes cytokine production, generation of oxygen radical due to fatty acid oxidation, hepatocytes damage and apoptosis, and ultimately fibrosis (60). The generation of IL-6 in Kupffer cells is gender dependent as its level elevates in males who usually have higher association with HCC.

1.2.3 Early detection and diagnosis of HCC

Early detection is one of the most important approach to decrease morbidity and mortality of HCC, allowing effective therapy strategy to be planned. Clinical screening for the early development of HCC is routinely performed by using noninvasive techniques including ultrasonography (US) and detection of serum alpha-fetoprotein (AFP) levels. The former method is a regular clinical practice in cirrhotic patients every 6 months, which yield almost 50% of the HCC detected cases amendable to therapy that benefits outcome. The main limitations of this technique are firstly, the expensive cost that restricts its use to high risk group that are potentially curable at diagnosis. Secondly, its sensitivity is limited by the prevalence of obesity in western countries, for instance in the US, where the higher sensitive methods such as Magnetic Resonance Imaging (MRI) and dual-phase helical CT (61) are preferred. The use of elevated serum AFP levels lacks specificity for screening of early lesions, because of its high false positive rate in which patients with active hepatitis or vascular invasion also have high AFP level. However, the elevated AFP concentration and male sex are always the strongest predictors for HCC risk .

The progression of chronic hepatitis to HCC is a multistage process and liver biopsy plays a central role for evaluating the pathological changes in this disease progression. It remains the gold standard to identify precursor lesions of premalignant stage including cirrhosis, liver cell dysplasia, to differentiate macroregenerative nodules from well-differentiated HCC (9), and particularly when imaging techniques

fail to detect premalignant lesions due to the absence of vessel formation in earlier stages (62). The picrosirius red staining (63) of liver collagens including collagen type I and III, and the reticulin staining (64) of liver connective tissue including perisinusoidal and pericellular fibrosis are more specific than routine H&E staining used in detection of fibrosis, a typical lesion of CLDs. However, the use of liver biopsy for detecting liver disease is limited by sampling bias and high inter-examiner variation (65) and this has exerted the urgent need of methodological standardization (66-68). A basic critical measure for this standardization is the size of liver biopsy, if unacceptable small, may reflect the flawed disease condition of the whole liver organ. A recent study (68) has reported a marked increase of “mild” cases positive in necro-inflammatory and fibrosis lesions, from 49.7 to 86.6%, when the length of biopsy decreased from 3 to 1.5 cm. Other study by Guido and Rugge (69) has emphasized on the use of liver biopsy of 2 cm and more for reliable grading and staging.

1.2.4 Staging and treatment for HCC

The treatment for HCC should ideally have minimal influence in liver function as tumor arises in a diseased liver that is functionally deranged. If the liver function is severely impaired, the patient will still ultimately not benefit from the approach, due to the dismal outcome, irrespective of the treatment success. Figure 1-6 presents the classification of Barcelona Clinic Liver Cancer (BCLC) (34, 70) that links HCC staging with treatment strategy, which aims to guide clinical practice.

Figure 1-5. Barcelona Clinic Liver Cancer (BCLC) staging and treatment strategy for HCC. Adapted from (34, 70), with permission from Elsevier and Oxford University Press.

Patients usually receive curative treatments including transplantation, surgical operation, or local ablation, either with percutaneous ethanol injection (PEI) or radiofrequency (RF) if diagnosed early enough during initial stage. These treatments could yield 5-year survival rates ranging from 50% to 70% (71, 72) because candidates are pre-selected based on tumor status and liver function. The former is defined by size of the main or multiple nodules (single ≤ 2 cm, single 2-5 cm, 3 nodules ≤ 3 cm) and the latter is determined by acceptable serum bilirubin level and portal hypertension. In developed countries, 30% to 60% of HCC cases can be diagnosed at early stage because surveillance programs were implemented on patients with cirrhosis (71). This allows the curative treatments to benefit 30% to 40% of

cases between 1990 and 2010 (Figure 1-6), and later 40% to 60% of cases estimated in the next decade, between 2010 and 2020 (73).

Surgical resection and transplantation are the two top treatment options in well pre-selected patients having early tumors, with best outcomes achieving 5-years survival rate of 60-70%. Hepatic resection is the preferred treatment for HCC patients without cirrhosis, who constitute 5% of cases in the West and 40% in Asia (74). Patients with well-sustained liver function and portal pressure will frequently tolerate the major hepatic resection with low life-threatening complications as their remaining liver are less susceptible to cytokine-induced damage and ischemia-related injury post surgery resection. The efficacy of hepatic resection in cirrhotic patients which is traditionally low, has now increased, as the mortality rate and blood transfusions rate become less than 3% and 10%, respectively (72). These improvements are not only consequences of the advancement in the knowledge of the disease stage, risk factors of recurrence and survival that refine the selection of candidates, but also adequate surgical techniques and implementation of anatomic resections according to Couinaud classification. However, the main difficulty is high tumor recurrence rate after curative resection during followup (>50% at 3 years and 70% at 5 years) (71).

Majority of the recurrences (60 to 70%) correspond to intrahepatic metastases (true recurrence) which are undetectable during surgical resection, while 30 to 40% of cases are *de novo* tumors. Although adjuvant therapies such as internal radiation (75), adoptive immunotherapy (76) for prevention of undetected intrahepatic metastases,

and retinoids (77) and interferon (78) have reported some positive results for preventing *de novo* tumor, they are still not considered as the standard of management for patients after complete resection. Although presence of vascular invasion, poor differentiation degree and additional nodules are significantly associated with tumor dissemination, molecular analysis that helps to uncover the aberrant network of HCC, can provide a better prognosis to improve the treatment performance and a more specific molecular target-orientated management post resection.

1.3 Recurrence of HCC after curative resection

Intrahepatic recurrence is one of the major challenges in HCC treatment, as it arises in 30-50% of patients who undergo hepatic resection (79). Poor outcome is always correlated with a high recurrence rate as metastasis complicates 70% of resection cases at 5 years. Recurrence is categorized to true (intrahepatic) metastasis and multicentric occurrence (71, 74) which are clinically divided to “early” or “late” recurrence according to a cut-off time of 2 years after surgery (80). Early recurrence which accounts for 60% to 70% of recurrence, is related to tumor factors including presence of poor differentiation, vascular invasion, satellites, positive margin cut and high serum AFP level (81, 82) that may cause tumor cells dissemination from primary HCC before resection. Late recurrence, which appears after 2 years reflects a *de-novo* primary tumor in the preneoplastic liver remnant affected by continuous virus infection, inflammation and cirrhosis (83-85). These underlying factors generates “field effect” which is potentially caused by accumulated genetic hits such as cellular

transformation (86). The early recurrence is usually undetectable by imaging techniques within the first 2 years after surgery (87, 88) and in general carries poorer prognosis than late recurrence. Several cohort studies have recently exhibited that patients with elevated serum HBV DNA level have higher rate of developing disease recurrence (89-91) or metastasis (92) than those with low levels after surgical resection or transarterial chemolipidolization (93). Yamamoto and colleagues identified coexisting cirrhosis as one of the significant predictor for recurrence (85). Patients with cirrhosis have a shorter recurrence-free survival, high morbidity and mortality rate over 10% after resection when compared with noncirrhotic patient.

Metastases are the cause of high mortality rates (about 90% of cancer deaths) associated with cancer. Surgical resection is considered the best “curative treatment”, but >80% of patients have widespread HCC at the time of diagnosis and are not candidates for this treatment (71). Molecular signatures that can predict well early recurrence did not allow accurate prediction for the late intrahepatic recurrence (94, 95). Patients undergoing cirrhosis with portal vein thrombosis (PVT) but non-neoplastic usually are more prone to develop an advanced liver disease which may lead to hepatocellular carcinoma. The risk increases more than fivefold in the presence of mutation 20210 within prothrombin gene than of PVT (96).

1.4 Risk Factors

1.4.1 Viral Hepatocarcinogenesis

The two hepatitis viruses that play prominent roles in the evolution to HCC are HBV, a non-cytopathic and partially double-stranded hepatotropic DNA virus from *hepadnaviridae* family and HCV, a non-cytopathic RNA virus from *flaviviridae* family (26). The mechanisms of HBV-mediated HCC may involve directly the virus itself, firstly by integration of the genome into the chromosome of hepatocytes which may interfere with the host cancer-induced genes including TERT, PDGF β and MAPK1 (26), and secondly the HBx protein which transactivates the Ras-Raf-MAPK pathway and inactivates the tumor suppressor p53 *in vitro* (97). The indirect effect by HBV includes induction of T-cell immune response and chronic active virus infection resulting from incomplete HBV clearance, which provoke the repeated cycles of necrosis-inflammation-regeneration (26, 97) that leads to fibrosis and ultimately cirrhosis. The interaction between virus and endoplasmic reticulum of host cells provokes ER stress and oxidative stress which produces free radical leading to mutations and activates stellate cells which ultimately results in fibrosis (98). Chronic hepatitis B virus (HBV) infection confers to approximately 55–60% of hepatocellular carcinoma (HCC) cases in the world. Approximately 70% of HBV-associated HCC (97) and almost all cases of hepatitis C virus (HCV)-associated HCC occur in the presence of cirrhosis. Several case-control studies have reported that increasing serum HBV DNA level is not only correlated significantly to disease progression in chronic hepatitis B (99-101), but also provide bad prognosis including higher risk of death

(102-105) after surgery. The nonviral risk factors may include aflatoxin exposure, alcohol drinking, and host susceptibility factors (106, 107).

1.4.2 Cirrhosis

Cirrhosis is one of the strong risk factor for HCC development as about 80 to 90% of HCCs are found in patients with preexisting cirrhosis (36, 46, 108). Cirrhosis in uninvolved liver is one of the most prominent predicting lesion for HCC. Adachi and colleagues found that the necroinflammatory activity and high cell proliferation in the remaining cirrhotic liver to be highly correlated with intrahepatic recurrence in individual with completely resected small HCC (109).

Only a minority of 10 to 20% of all HCCs occur in noncirrhotic livers, except for some groups of African blacks, where about 40% of the HCCs are found in otherwise normal livers (110). The etiology of HCC in noncirrhotic livers remains less well-understood. Several studies reported an etiologic relationship between estrogen and noncirrhotic HCC, but this has never been convincingly proven (110, 111). Other factors that were suggested as the possible roles in the pathogenesis of noncirrhotic HCC in Western countries are ethanol abuse, hepatitis B, exposure to pesticides and environmental toxins, and cigarette smoking (110, 112). Based on etiologic and demographic differences, it has been reported that HCCs arising from cirrhotic and noncirrhotic liver may have different pathogenic backgrounds (113).

The survival rate in symptomatic noncirrhotic HCC patients seems to be more favorable than in symptomatic cirrhotic HCC patients partly because the HCC in noncirrhotic livers is less malignant and the age of this patient group is younger, which may contribute to higher functional hepatic reserve, respectability rates and the lower life expectancy of cirrhosis (110, 112).

1.4.3 Vascular Invasion

Vascular invasion or intravasation involves the interaction of tumor cells with extracellular matrix (ECM) followed by migration of the cells into the circulation of the primary tumors. This serves as the key step in tumor metastasis, and the tumor cell migration continues despite removal of the primary tumor (114). In other words, vascular invasion is an indication of the inherent malignancy of the lesion and a channel for intrahepatic and extrahepatic metastasis (115). Portal vein thrombosis which is related to portal vein invasion has been proposed by several authors as an independent adverse prognostic factor in both Eastern and Western series (108, 116). Invasion and growth into hepatic vein branches and the portal vein are common in HCC. For tumor size above 3 cm, the presence of vascular invasion appears to be the major indicator of recurrence (117). The extent of vascular invasion along with the size and multicentricity of tumor nodules, and the presence of local lymph-node metastases, are used as a guiding approach to select suitable candidates for liver transplantation (118, 119).

During invasion and intravasation, the intrinsic polarity and cell-cell junctions that are both required for maintaining integrity of normal epithelium, are lost in tumor cells. The extensive loss of extracellular matrix (ECM) networks unlocks cell shape and movement from dense networks of ECM. Cells that partially retain their polarity display a retraction fibrous tail and stress fibres front, to form focal adhesions, resembling the slow moving or static fibroblasts cells due to a tight attachment with ECM. In contrast, when all polarity are lost, cells become ‘amoeboid’ reflecting that these cells are loosely attached to the ECM and are undergoing marked shape alteration to form focal complexes, hence they can migrate fast. These cells are not fibroblastic during migration. In the process of understanding the mechanisms of invasion, the genes particularly those involved cell adhesion identified, will provide more efficient strategies to manage secondary and tertiary metastases (114).

1.5 Gene Expression Profiling in Cancer

Gene expression profiling is a global functional approach for molecular analysis of thousands of genes concurrently in various normal or diseased samples. The differential and interaction of expressed genes are discovered and extracted from huge data sets by using computational approaches. This yields the opportunity to classify tumor of unknown origin, to identify more efficient targets for diagnostic purpose and therapy design, and to find molecular signatures that explain cancer progression. Over the past two decades, the most common technologies that allow global comparative analysis of gene expression between multiple healthy and diseased

cells, are spotted complementary DNA (cDNA) microarrays and Affymetrix oligonucleotide GeneChip arrays.

The cDNA arrays have evolved from membrane-based to the glass slide-based. Arraying hundreds to thousands of cDNAs, produced by Polymerase Chain Reaction (PCR) amplification on a membrane for hybridization with radioactively labeled samples was the first DNA array approach (120). The technology took another leap forward when researchers at Stanford University started using robotics to spot cDNA onto glass slides at much greater densities than with nylon membrane (121). Another advancement is the use of two fluorophores instead of radioactive probes, typically red and green (cyanine 5 and cyanine 3), to label test and control sample before being hybridized to the same array. Each spot is measured in terms of the expression ratio between probes.

Oligonucleotide array is built by *in situ* light-directed synthesis for high-density DNA probes via photolithography chemical synthesis or solid-phase synthesis and followed by immobilization on a glass support (122). Each probe set on the chip is consists of 11 different sequences with each sequence composed of 25 base pairs (bp) in length. The latest human genome 3' IVT expression chip, HGU133 plus 2.0 array contains over 47,000 probe sets representing 38,500 well-substantiated human genes. The gene sequences for all the probe sets were from GenBank, dbEST, and RefSeq databases. The sequence clusters were created using updated UniGene cluster database

and further analyzed and compared using Washington University EST trace repository and the NCBI human genome assembly.

Microarray technique has been largely used as a routine global gene expression assay in many cancers including leukaemia (123), lymphoma (124), lung cancer (125), breast (126) and prostate (127). It has been reported to have wider application than the conventional techniques to verify the clonal relationship among different nodules in liver cancer (128). This may be due to its robust automated handling and standardization during manufacture such as array and probe synthesis, hybridization, and signal acquiring. The genomic characterization may allow a molecular diagnosis of HCC, but the gene-set proposed needs further validation. Recently, the genome-wide study conducted by Hoshida et al. have identified a gene signature that has direct implications for predicting patients at high risk of death due to late recurrence (129). This study has opened the door to identifying the gene-expression in the pathogenesis of hepatocellular carcinoma and bringing the research to utility for the clinical assessment of risk of recurrence as the gene signature was derived from formalin-fixed, paraffin-embedded nontumorous liver.

1.6 Support vector machine

Similar to other supervised learning techniques, Support vector machines (SVMs) (36, 37, 108) involve the use of training to specify in advance the samples to be clustered. When used on gene expression analysis, a cluster of genes that share the same function would be formed: for instance, genes accounted for ribosomal proteins

function or proteasome function. Meanwhile, genes that do not belong to the same cluster are also highlighted. These two clusters are integrated to generate training samples. Genes belong to the indicated cluster are positively labeled while genes do belong to the indicated cluster are negatively labeled. After supervised learning of the expression patterns of the cluster, SVM could discriminate new incoming genes from an independent test set. In general, an SVM would employ the information learnt from the training data not only to cross classify the characteristic of the expression features of an indicated group, but also estimate the indicated gene belong to the indicated group.

SVMs have two advantages over unsupervised clustering analysis including selforganizing maps and hierarchical clustering. Firstly, SVMs s can utilize distance metrics that work in the greatest high-dimensional feature spaces although other three algorithms can use distance metrics to compare measurements of gene expression patterns. Hence, this enables SVMs to consider assoication between gene expression measurements. Second, using prior knowledge information gained in the training data it can distinguish one gene cluster from the other. Unlike the unsupervised algorithms, when two genes are separated distantly from each other according to the distance metrics, the method can hardly justify that the genes are related (130).

1.7 Identification of Rac GTPase Activating Protein 1 (RACGAP1) as a HCC recurrence biomarker

1.7.1 RACGAP1 plays a role in cell differentiation

Rac GTPase activating protein 1 (RACGAP1), also known alternatively as MgcRacGAP (131), CYK4 (132) or lately as HsCYK-4 (133) in *Homo sapiens*, CYK-4 (134) in *Caenorhabditis elegans*, RacGAP50C and Tumbleweed (135) in *Drosophila melanogaster*, was first characterized as a human chimaerin-like protein in male germ cells (131). It consists of 3 main domain including myosin-like domain (amino acids 1-124), Protein kinase C-like cystein-rich domain (amino acids 280-340) and three consensus blocks conserved in RhoGAPs (amino acids 345-620). The N-terminal myosin-like domain and the cysteine-rich domain, but not GAP activity are required to induce differentiation of hematopoietic cells, HL-60 (136). Hence, overexpression of RACGAP1 suppressed cell growth and proliferation of HL-60, M1, Ba/F3, Jurkat, and HeLa cells, but induced macrophage differentiation of hematopoietic cells, HL-60 (136, 137).

1.7.2 RACGAP1 plays a role in regulating cytokinesis

The GAP domain of RACGAP1 strongly stimulates Rac1 and Cdc42, but is almost inactive on RhoA in human hematopoietic cells *in vitro* (131, 137). The other studies employing *Caenorhabditis elegans* (138) and *Drosophila* (139) system conclude that GAP domain acts on Rac1. The furrow ingression is attenuated as a

result of ectopic activation of Rac1 at the furrow in the absence of GAP activity of RACGAP1 orthology CYK-4, and this cytokinesis defect can be rescued by depletion of Rac1 ortholog CED-10 (138, 139). However, contradict result is observed in *Xenopus laevis* system in which the GAP domain acts primarily on RhoA rather than Rac1 or Cdc42 to assemble the cytokinetic apparatus (140). The expression of GAP-deficient variant replacing endogenous CYK-4 renders unfocused and hyper-accumulation of active RhoA and also F-actin, a downstream Rho target at the furrow, leading to failed cytokinesis. This study concludes that GAP domain of CYK-4 continuously inactivates RhoA during the ingression process (140). In summary, the functional requirement of GAP activity of RACGAP1 in regulating cytokinesis varies in different system.

The major role of RACGAP1 in cytokinesis has been previously reported (141). A recent high-throughput genome-wide RNA interference screening study that phenotypically profiles human genome using time-lapse microscopy further corroborates the important role of RACGAP1 in cytokinesis, as its loss exhibits binucleated and sometimes multinucleated cells, indicating an irreversible cytokinesis defect (142). RACGAP1 associates with MKLP1 (lately named KIF23) to form centralspindlin complex to recruit cortical localization of RhoA, which is essential for cytokinesis (143-146). It works with ECT2 in regulating the activation and function of Cdc42 spatio-temporally in mitosis (147-149). Its interaction with Polo-like kinase 1 (PLK1) (150), Anillin (ANLN) (151), Aurora B (AURKB) (152), Cep55 (153) in

regulating cytokinesis, were implicated in recent years. Rac GTPase activating protein 1 (RACGAP1) was previously identified in a prediction gene set (154).

1.7.3 RACGAP1 in other cancer databases or other aspects

The alteration of RACGAP1 together with its neighbourhood targets particularly PLK1, ANLN, ECT2, KIF23, and PRC1, was detected as part of the gene member with elevated expression in several other microarray cancer databases including epithelial ovarian cancer (155), cervical cancer (156), high grade breast cancer with transition from preinvasive to invasive stage (157) as well as ER positive breast cancer (158-160). Viral genome-wide RNAi/genetic screens using A549 cell line identify RACGAP1 as one of the host cellular factors required for early-stage influenza A virus replication (161), presumably by regulating IP3-PKC signaling pathway and cytoskeletal organization implicated respectively in viral entry and intracellular viral transport (162). Human cellular cofactors which are associated with cytoskeleton organization, is the only complex commonly implicated in influenza, HIV, HCV, dengue and West Nile virus (WNV) replication (162), based on the results of the previous screens.

1.8 Proteomics

The term 'proteome', initially coined by Wilkins et al. (163) describes the entire set protein encoded by the genome and the study of which is proteomics. Proteomics

refer to the study of almost everything post-genomics including all the proteins in any given cells, tissue or organism, in terms of the interaction between them, the isoforms, modifications, and structural description of proteins, in high-throughput setting. Proteomics is considered to complement other functional genomics approaches, including microarray-based expression profiles, systematic phenotypic profiles and systematic genetics to yield comprehensive insights into cellular function. In other words, the previous achievements of genomics has rendered the recent development of proteomics possible (164).

1.8.1 Principle of mass spectrometry-based proteomics

The traditional proteomics largely use two-dimensional gel electrophoresis (2-DE) to separate protein followed by mass spectrometer identification of protein spots. The protein mixture sample is first separated on an immobilized pH gradient strip, and later further separated into individual spots (proteins) by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) slab gel before staining with silver or Coomassie Brilliant Blue. After comparing the staining patterns of the proteins from two samples, the differentially expressed proteins of interest are defined, excised and cleaved into peptides by using proteolytic enzyme such as trypsin. The peptide digests are then identified by mass spectrometry (MS), particularly, the matrix-assisted laser desorption/ionization (MALDI), where a dry crystalline matrix is used to minimize the sample damage by direct laser beam and also serves to facilitate vaporization and ionization process by absorbing the laser radiation. The inherent

limitation of this gel-based separation approach is due to its inability to identify low abundant proteins and low resolution membrane proteins.

To meet the challenge by resolving higher degree of protein complexity and detecting low abundant proteins, the alternative approaches involve the use of liquid-based techniques consisting of high performance liquid chromatography (HPLC) and capillary electrophoresis that gradually substitutes the gel-based separation. In LC-based approach, the cleaved peptides mixture are bonded and then separated sequentially based on the hydrophobicity and charge of the solvents and analytes, and finally identified by MS. Electrospray ionization (ESI) is another sophisticated ionization method having advantage that can create ions directly out of a solution. The recent development of LC-based separation coupled with ESI-MS has greatly improved the technology to allow systematic analysis of cellular proteome with higher complexity particularly proteins involved in post-translational modification (PTM) and protein–protein interactions (165). In addition to the two types of ion sources described above, the other essential components of MS, consist mainly of a mass analyzer that measures gas phase ions according to the mass-to-charge ratio (m/z), and a detector which generates a current signal correlated with the abundances of respective ion fragments. Sensitivity, resolution, mass accuracy and the ability to obtain *de novo* protein sequence from peptide fragments via MS/MS spectra or tandem mass are always the key parameters that determine the performance of an analyzer (165, 166).

To date, lack of reliable quantitative methods remains one of the major challenge for system biology which aims to model cellular behaviour at global-system level (164). In both MALDI- and ESI-MS, the amount of analyte within a sample is not well reflected by the signal intensity generated if quantitative measures are not taken into account. Hence, the detection of protein abundance in a sample, especially changes in proteins level altered by perturbation of a biological system, is not possible if no reliable quantitative measures are conducted. This may lead to the false negative results as the protein-specific peptides fall beyond the mass range analyzed or they are too strongly attached to the chromatographic column, exceeding the detection threshold limit or the sensitivity of MS. The well proven technique of stable-isotope dilution has been employed as a quantitative measure in LC-MS/MS experiments. Several quantification methods developed and classified based on the labeling of stable isotope, are classified into 4 groups: (i) Spiking of the internal standards, the isotopically labeled peptide generated with indicated quantities (167), (ii) enzymatic incorporation by transferring ^{18}O from water to peptides (168, 169), (iii) isotope-tagging incorporation onto peptides and (iv) metabolically labelling the cells with stable isotope, by supplementing the stable isotope labeled amino acids in the growth media of a cell culture (SILAC). As a result, one cell state is metabolically labeled by, for instance, ^{13}C -labelled arginine (170). Of all the stable isotope-based quantification methods, iTRAQ has widely used recently as four samples can be analyzed simultaneously, hence increasing analytical output (171).

1.8.2 Protein profiling studies with proteomics approach

Large-scale analysis of MS-based proteomics has been commonly used in investigating protein expression profiles as a tool to infer cellular function. The systematic analysis of the proteins expressed by a cell or tissue has been reported recently by the use of high-throughput proteome mapping study of yeast (172) and the bacterium *Deinococcus radiodurans* (173). Intriguingly, stable-isotope dilution and LC-MS/MS are used to identify potentially clinically significant molecular patterns in cancer (174), such as HCC (175-178) and lung cancer (179), heart disease (180), as well as other common ailments. In addition, this technology is also employed to gain insight of specific cellular signaling pathways contributing to disease progression, by comparing healthy and diseased tissue samples (181-184) so as to identify protein profiles or drug-specific markers. Goufman et al. documented the use of mass spectrometry system to identify the biomarkers of serum samples from various cancers such as ovarian, uterus, and breast cancers, as well as samples from benign ovarian tumor (185). The results uncovered a few protein markers, including the down-regulation of R-1 acid glycoprotein and clusterin in breast cancer, and transthyretin, a specific down-regulated marker of ovarian cancer (185). ATP synthase was shown to be up-regulated in tumor tissues and localized at the plasma membrane of breast cancer cells. Treatment with ATP synthase inhibitor, aurovertin B, exhibited a significant decrease of cell density in breast cancer cells, indicating that aurovertin B can be exploited as an antitumorigenic agent (186).

1.9 Objective and approaches for identifying molecular markers for HCC recurrence

The objective of this study was to integrate both clinical pathological information and oligonucleotide microarray profiling to identify differentially expressed molecular signatures for accurate prediction of intrahepatic recurrence. The differential gene signature was also beneficial in understanding the progression of HCC to become recurrent even after complete hepatic resection. Amongst the differentially expressed genes, we focused on RACGAP1 as this gene showed the most consistent upregulation not only in HCC recurrence, but its expression was also elevated from normal normal liver and surrounding tumor to non-recurrent samples or HCC samples as a whole. We attempted to investigate the functional role of this gene in HCC recurrence using cell-based assays. The loss-of-function assay employing siRNA targeting RACGAP1 in metastatic HCCLM3 and MHCC97-H cell lines expressing high endogenous RACGAP1 and gain-of-function assay by transient transfection the exogenous RACGAP1 DNA in a non-metastatic PLC/PRF 5 cell line expressing low endogenous of this gene, were performed.

CHAPTER 2 Materials and Methods

2.1 Methods

(Some of the materials, reagents, and methods used in the following description are established standard protocols or materials recommended by the manufacturer. Therefore, it might be by no means commonly used by people in a similar approach. However, all the techniques and materials used in this study are designed to specifically suit the experimental requirement in this study)

2.1.1 Patient samples procurement

Cancerous and some of the corresponding surrounding non-cancerous liver tissues were obtained from consecutive patients who underwent partial hepatectomy as a treatment for HCC. All tumor tissues were divided to two parts and half portion of all the samples was stored immediately in liquid nitrogen until use. The other portion were stained with hematoxylin and eosin and evaluated by qualified pathologists to ascertain that the tissues studied were either cancerous (more than 80% cancerous tissues) or non-tumorous. All tissue samples employed in this study were approved and provided by the Tissue Repository of the National Cancer Centre, Singapore and conducted in accordance with the policies of its Ethics Committee. Informed consent was obtained from all participating patients prior to surgery. To protect patient confidentiality, linked clinical and histopathological data were collected from medical records for all patients who contributed tumor specimens and were rendered anonymous.

To assess recurrence, all treated HCC patients were monitored by routine clinical follow up once every 3 months. The level of alpha-fetoprotein level (AFP), liver function tests (LFT) were determined every three months and ultrasound scans were performed every six months. CT (computerised tomography) scans were performed when the level of AFP was greater than 400ng/ml or when the ultrasound results indicated the presence of recurrent disease. Biopsies of histologically non-tumorous tissues adjacent to the tumor of ten colorectal cancer patients with liver metastases, were used as reference normal liver tissues.

2.1.2 Total RNA isolation and quantitation

Frozen HCC biopsies were immediately homogenized using Omni tissue homogenizer (Omni International, Warrenton, Virginia) in 1 ml TRIzol reagent (Invitrogen, Carlsbad, CA) until no tissue aggregates was visible and total RNA was isolated as described in the manufacturer's protocol. Samples were mixed with chloroform by vortexing. After centrifugation at 12,000 rpm for 15 min at 4°C, the upper aqueous phase was removed to perform an additional extraction step using Acid Phenol: Chloroform (pH4.3 – 4.7, Ambion) followed by centrifugation, and chloroform mixture to eliminate DNA contamination. RNA was precipitated with equal volume of isopropanol followed room temperature incubation for 10 min and centrifugation at 12,000 rpm for 10 min at 4°C. After one time 70% ethanol wash, the RNA pellet was air-dried briefly and resuspended in 1x RNAsecureTM reagent (Ambion) and heated at 60°C for 20 min before storing at -80°C freezer. The

concentration of total RNA was determined by NanoDrop® ND-1000 spectrophotometric measurement (one absorbance unit at 260nm is equivalent to 40 µg/ml RNA). To ascertain the quality of the RNA, gel electrophoresis which allows the visualization of the ratio intensity of 28s and 18s RNA or Agilent 2100 Bioanalyzer analysis which provides RNA integrity number (RIN). This number range from 1 to 10 and 1 represents the most degraded profile while 10 indicates the most intact RNA profile.

2.1.3 Affymetrix GeneChip® hybridization

All first strand and second strand cDNA synthesis, cRNA synthesis, labeling, hybridizations, wash and scan procedures were performed according to protocols of Affymetrix GeneChip experiment. In brief, 5 µg total RNA was reversed transcribed to first strand cDNA using HPLC-Purified primer incorporated with the T7 RNA polymerase promoter binding site linked to oligo-dT₂₄ (Sigma-Proligo). Second-strand cDNA synthesis was carried out by using superscript II RNase H-reverse transcriptase (Invitrogen, Carlsbad, CA). The cDNAs were purified by phase-lock gel (Eppendorf, Hamburg, Germany) before being used for downstream *in vitro* transcription with RNA transcript labeling kit (Enzo Diagnostics, Farmingdale, NY) to produce amplified biotin-labeled complementary RNA (cRNA), which was then purified with Qiagen RNeasy kit (Qiagen, Hilden, Germany). The purified cRNA was fragmented and 15 µg was used for hybridization to human HG-U133A followed by the second

hybridization onto HG-U133B oligonucleotide probe arrays (Affymetrix, Santa Clara, CA), and the arrays were washed and scanned the next day as described in the previous publication (154).

2.1.4 Microarray data pre-processing and analysis

The raw image data generated from scanning were immediately analyzed by generating reports for all samples. Quality of each array was examined by assessing the ratio of 3'/5' of Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and beta actin, and the percentage of genes with “present” or “absent” calls. All the arrays were normalized against all probe sets and the signal intensity of the entire probe sets in both arrays of all samples were scaled to 500 prior to performing further statistical analysis. The data were filtered with absent/present calls using Affymetrix[®] Data Mining Tool 3.0 software. Probe sets that have at least 70% present calls in all samples grouped under a particular clinicopathological factor were kept for further analysis using Genedata Expressionist Analyst (version 1.0) (Genedata AG, Basel, Switzerland) and Partek[®] Genomics Suite[™] (Partek Inc., Missouri, USA). All data generated by Affymetrix[®] Microarray Suite version 5.0 in cel file format were exported to and refined by Genedata Expressionist Refiner using protocol of GD Affx diagnosis with reference. Refining of all HG-U133A arrays consist of both tumor and normal tissues were done separately from HG-U133B arrays. The cel files of all HG-U133A and HG-U133B arrays for all samples were imported separately into Partek[®] Genomics Suite[™] and the signal

intensity were all transformed into log base 2 before GC-RMA algorithm was used for background correction.

2.1.5 Statistical and computational analysis

The correlation between the clinicopathological factors of 80 HCC samples and the recurrence status were evaluated with the SPSS (version 15.0). Univariate analysis was used to investigate the association of recurrence and various clinical parameters including sex, age, presence of HBV or HCV, tumor size, capsulation, alpha-fetoprotein (AFP) level, single or multiple lesions, histological grading, presence or absence of cirrhosis and vascular tumor invasion. Multivariate analysis was performed to determine the independent risk factors for recurrence. Recurrence was entered as dependent factor and the other variables were entered as covariate factors into a forward stepwise logistic regression model. Kaplan Meier survival plots using log rank test was performed to assess the prognostic significance of the expression of a gene in relation to recurrence.

All HCC samples were clustered into groups according to their clinicopathological characteristics that have significant correlation with recurrence. Subsequently, prediction model generated from Genedata Expressionist Analyst was applied to classify HCC samples with recurrence potential. Each group was segregated randomly into training and test set. The combined parametric (Student T-test) and non-parametric (Wilcoxon test) analysis was applied to find genes that discriminate significantly ($P < 0.05$; median fold change ≥ 1.25) between predefined clinical groups

of the training sets. To keep the minimum misclassification error rate during leave-one-out cross-validation of the training sets, the optimal number of genes was obtained by selecting several top ranked gene sets containing fold change successively increase from 1.25 to 5. The training set cross validation and test set prediction were performed as described in the previous publication (154). Positive correlation was used as the distance metric and the same algorithms, Support Vector Machines (SVM), Sparse Linear Discriminant Analyst (SPLDA), and K Nearest Neighbour (KNN) were used for both cross-validation of the training sets and prediction of the independent test sets. Sensitivity is defined as rate of positive cases (recurrence) correctly predicted and specificity is rate of negative cases (non-recurrence) correctly predicted.

To select molecular signatures that can predict recurrence, first we employed both the parametric (Student T-test) and non-parametric (Wilcoxon test) statistical tests followed by median fold change (F) filter to select genes that could discriminate between two groups of each parameter including patients with vascular invasion versus without invasion, cirrhotic versus non-cirrhotic cases and recurrence versus non-recurrence cases respectively. Each gene was ranked based on the P value calculated by both the tests. The highest ranked gene should have the smallest P values from both the tests. After each step of filtering starting from the top 200 ranked genes, followed by successively increasing of fold change filter from greater or smaller than 1.25 to 5.0, 6 gene sets were generated. The same fold change filtering approaches were repeated for the top 190 ranked genes until the top 50 ranked genes. All the gene sets generated were used to estimate the training sets by “leave-one-out”

cross validation and those that yielded the minimum misclassification error rate were kept for further validation of the independent test set.

To display the expression pattern of the selected genes across different groups of HCC samples together with normal liver tissues, the signal intensity of each sample was first normalized against the average of all samples of that gene. Unsupervised hierarchical clustering algorithm was then carried out by CLUSTER and TREEVIEW software available from <http://rana.lbl.gov/eisen/> using mean centered correlation as the measure of similarity and average linkage as described in Eisen *et al.* (187). The target genes were then annotated using the Affymetrix NetAffx™ GeneOntology analysis system available from <http://www.affymetrix.com/site/login/login.affx> and FATIGO from <http://fatigo.bioinfo.cnio.es/>. The Proteome Biochemical Function, molecular function, cellular role, and biological process attributes stored in GeneOntology were used to functionally categorize the selected gene sets. The original microarray data files have been deposited in the EBI microarray public database available online from <http://www.ebi.ac.uk/arrayexpress/>.

2.1.6 Reverse Transcription

DNA contamination of all total RNA samples was treated with DNase I (Ambion). 5µg of DNase I-treated total RNA was mixed with 1ul of 3µg/ul random hexamer primer and 1 ul of 500ng/ul Oligo (dT)₁₂₋₁₈ primer (both are from Invitrogen,

Carlsbad, CA), and mixed with Diethylpyrocarbonate (DEPC) -treated water. The reaction mixture was then incubated at 70°C for 10 min and immediately chilled on ice before adding with the following components: 1 µl of 10mM dNTP (Promega), 2 µl of 0.1M DTT, 4 µl of 5x First Strand Buffer, 1 µl of 40U/ul RNasin® RNase Inhibitor (Promega) and 1 µl of 200U/ul SuperScript™ II RNase H⁻ Reverse Transcriptase (Invitrogen, Carlsbad, CA). The reverse transcription reaction was then carried out at 25°C for 10 min, 42°C for 1.5 hr and terminated at 70°C for 15 min. The first strand cDNA was then diluted 3 times with DEPC-treated water and 1 to 2 µl was used as the template for conventional PCR or quantitative real-time PCR reaction.

2.1.7 Polymerase chain reaction (PCR)

Each gene was first amplified by conventional PCR reaction prior to real-time PCR amplification. PCR reactions were performed in 20µl reaction mixture containing 1xPCR reaction buffer, 1.5mM MgCl₂, 50µM dNTPs, 0.25µM primer mix (forward and reverse primers) and 1 to 5U of 5U/µl *Taq* polymerase (Promega). The PCR reaction was performed in the condition as stated in the previous publication (154). The amplicon were then purified by Qiagen PCR purification kit (Qiagen, Hilden, Germany) for later use as template in constructing the standard curves with serial dilution for real-time PCR reaction.

2.1.8 Real-time PCR quantitation

The mRNA expression of the selected genes from 2 gene sets based on their top fold changes and significant differences in microarray experiments were verified using real-time PCR analysis. The forward and reverse primers of each gene flanking a region of approximately 200 base pairs were designed using Primer 3 software available online. All primers had a melting temperature ranging from 58°C to 60 °C. The real-time PCR amplification was performed in a total reaction volume of 20 µl containing 1x QuantiTect™ SYBR® Green PCR Master Mix (Qiagen, Hilden, Germany), 1 µM mixture of each forward and reverse primers, 1 µl of diluted cDNA and Nuclease-free water. All reactions were carried out with 45 cycles (94 °C, 15 sec; 55 °C, 30 sec; 72 °C, 30 sec) using Rotor-Gene RG 3000 (Cobertt Research, Australia).

All PCR reactions were performed in duplicate and the melting curve for each gene of all samples was monitored to ensure the specificity of the amplicons. Standard curves for each gene were generated independently by preparing 10x serial dilutions of the purified PCR products. The 18S ribosomal RNA was used as a house-keeping gene and loading control normalization to eliminate amount variation while adding sample to the reaction. For each gene, the averaged copy concentration of every samples calculated from the standard curves was normalized against its corresponding averaged concentration of 18S ribosomal RNA to obtain relative expression for comparison among HCC subgroups and normal liver tissues. Statistical analysis was performed using parametric (unpaired T test) and non-parametric test (Mann-Whitney

test) from GraphPad Prism 5.0 software. Median relative signal of each subgroup was used to calculate fold difference between groups. A p-value less than 0.05 was considered significant in all cases.

2.1.9 Immunostaining of RACGAP1

Frozen tissues embedded in Jung tissue freezing medium (Leica, Germany) on the specimen disc was immediately chilled in liquid nitrogen for a few seconds before transferring to Leica CM1950 cryostat (Leica, Nussloch, Germany) which was pre-cooled to -20°C. 6µm thick sections cut from each pair of frozen tumor and the corresponding matched surrounding non-tumorous tissues were mounted onto the same Superfrost Plus microscope glass slide (Thermo Fisher scientific, Waltham, MA), and stored in slide box at -80°C until use. All sections were fixed with prechilled acetone:methanol (70%:30%) at -20°C for 10 min and antigen-retrieval was performed by boiling the slides in 10mM citrate buffer for 20 min in microwave. Endogenous peroxidase was blocked for 15 min at room temperature by using Dako REAL Peroxidase-blocking solution (Dako, Denmark) followed by two washes of PBST (0.05% Tween-20 in 1x PBS) before the non-specific background block using 5% normal goat serum (Zymed, San Francisco, CA) for 30 min at room temperature. The slides were incubated with mouse monoclonal antibody against RACGAP1 (Abnova, Taiwan) at 4°C overnight. After 3 washes with PBST, the sections were incubated with secondary antibody against mouse IgG conjugated with peroxidase-coupled dextran. The signal was detected by adding freshly prepared mixture of 20 µl of

DAB+ chromogen and 1ml of substrate buffer. The secondary antibody incubation and color development using DAB+ system were recommended by the protocol in Dako REAL EnVision detection system (Dako, Denmark). The negative control was performed in the same procedures except that primary antibody was substituted with isotype-matched mouse IgG2b (Dako, Denmark). The slides were all finally counterstained with Hematoxylin (Merck, German). Immunostaining was quantitated using Image-Pro Plus 7.0 software (Media Cybernetics, Bethesda) as follows: the percentage of positive immunoreactive lesion for each image was obtained by counting number of cells presenting brown nuclei staining out of the total number of cells that either had or did not have the same staining. For each sample, the expression of RACGAP1 was represented by averaging the percentage of positive immunostaining for three independent images generated from three non-overlapping fields photographed in 20x magnification using Nikon eclipse 90i microscope. Statistical analysis was performed using parametric (unpaired T test) from GraphPad Prism 5.0 software.

2.1.10 Generation of pIRES2-RACGAP1-AcGFP1 plasmid construct

(The cloning method and reagents used in the following description are from an established standard protocol recommended by the manufacturer, BD ClonTech. Therefore, it might be by no means commonly used by people in a similar approach. However, the techniques used are quite different from the others and specifically designed to suit the experimental requirement in this study)

Cloning of the full length RACGAP1 into pIRES2-AcGFP1 vector (Clontech, Mountain View, CA) was performed by using the In FusionTM 2.0 Dry-Down PCR

Cloning Kit (Clontech, Mountain View, CA). The protocol recommended by the manufacturer involved plasmid linearization by amplification, and insert amplification by using Phusion High-Fidelity DNA polymerase (Finnzymes Oy, Finland) (188). Plasmid was linearized via PCR by using forward primer Linear EcoR1 (F) and reverse primer Linear Sal1 (R) under the following conditions: initial denaturation at 98°C for 30 sec, followed by 30 cycles of 98°C for 10 sec, 66 °C for 30 sec and 72 °C for 80 sec, and a final extension step at 72 °C for 10 min.

The cDNA template for PCR cloning was reverse-transcribed from 1 ug of total RNA extracted from a HCC recurrence tumor by using SuperScriptTM II RNase H⁻ Reverse Transcriptase (Invitrogen, Carlsbad, CA). The full length RACGAP1 insert was PCR-amplified from the cDNA template by using RACGAP1 specific primers, forward and reverse primers at which each 5' end includes 15 bases of sequence homology with the linearization site of the vector to ensure the success of the In-Fusion cloning reaction. As such, the RACGAP1 forward primer Rg_EcoR1 5' (F) includes sequence homology with the EcoRI restriction site and reverse primer Rg_Sal1 3' (R) carries sequence homology with the Sal1 restriction site of pIRES2-AcGFP1 vector at the 5' overhang (Chapter 2.2). The PCR amplification of the insert was carried out in 2-step condition: initial denaturation at 98°C for 30 sec, followed by 30 cycles of 98°C for 10 sec and 70 °C for 30 sec, and the final elongation step at 72 °C for 10 min.

The insert and linearized vector PCR products were examined by 1% agarose gel electrophoresis to assure that a single, specific product of the expected size had formed prior to incubation with 20U *DpnI* (New England Biolabs, Ipswich, MA) at 37°C for 2 h in the presence of 1x NEBuffer 4 to degrade the template DNA. The treated PCR products were then spin column purified using QIAquick PCR purification kit (Qiagen, Hilden, Germany) and eluted in 50µl of Elution Buffer. Subsequently, 5µl of the pure PCR product was treated with 2µl of Cloning Enhancer (Clontech, Mountain View, CA) at 37 °C for 15 min followed by 80 °C for 15 min in a PCR thermocycler. To achieve maximum cloning efficiency, 100 to 150 ng of linearized plasmid and 1 to 2µl of insert were mixed in a 2:1 molar ratio of insert to vector, volume adjusted to 10µl, and incubated with In-Fusion Dry-Down pellet for 15 min at 37°C followed by 15 min at 50°C, and then transferred to ice. The In-Fusion mixture was diluted with 40µl TE buffer (pH 8) and can be stored at -20 °C until transformation reaction is ready.

2.1.11 Bacterial transformation and colony screening

Two and a half microliters of the diluted In-Fusion reaction mixture was mixed gently to 50µl of DH5 α competent cells (Invitrogen, Carlsbad, CA) by tapping the tube and incubated on ice for 30 min. After heat shock at 42°C for 45 sec, and snap-cooling on ice for 2 min, 950 µl of pre-warmed SOC medium was added to the mixture and was incubated at 37°C for 1 hour with shaking at 250 rpm. Approximately 1/10 of the transformed cells (100 to 150 µl) were spread on LB agar plate containing

100 µg/ml ampicillin and incubated upside down overnight (~16 hr) at 37°C. The individual bacterial colonies (10 or more) were picked from each plate using sterile toothpick or pipette tips and part of each colony was transferred to a new ampicillin-containing LB agar plate for incubation overnight at 37°C, while the leftover colony in the tips was resuspended in 20 µl dH₂O. Half of the resuspension volume was inoculated in 3ml ampicillin-containing LB broth and incubated for 8 hr at 37°C with vigorous shaking at 300 rpm as a starter culture, and the other half volume was boiled for 5 min and analyzed by PCR screening to determine the presence and the orientation of insert. The starter cultures with positive insert were frozen as glycerol stocks by mixing 0.15 ml sterile glycerol (100%) with 0.85 ml culture, and stored at -80°C. LB agar plate of bacterial colonies with positive insert was sealed and stored upside-down at 4 °C for several weeks. Alternatively, the colony with the right insert was inoculated into 5ml 100 µg/ml ampicillin-containing LB medium and incubated overnight at 37 °C with shaking at 250 rpm.

2.1.12 Plasmid DNA isolation

Plasmid DNA isolated from the 5ml overnight *E.Coli* culture in selective LB broth using NucleoSpin Plasmid kit (Machery-Nagel, Germany) was sent for sequencing to confirm the correct orientation of the desired insert. Once the colony with the correct insert was identified, large scale plasmid isolation was carried out by inoculating a single colony in a 3ml starter culture of selective LB broth and incubated at 37°C for 8 hr with vigorous shaking at 250 rpm. Within the same day, the starter

culture was then diluted in 1:500 into a larger volume of selective LB medium and grown in a flask of at least 5 times the volume of culture at 37°C for 12 to 16 hr with vigorous shaking (~250 rpm). Plasmid DNA was then isolated using QIAGEN EndoFree Plasmid Maxi kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

2.1.13 Cell culture

HepG2, Hep3B and PLC / PRF 5 cell lines were purchased from American Type Culture Collection (Manassas, VA) and grown in Dulbecco's Eagle's medium (DMEM, Sigma) supplemented with 10% heat-inactivated fetal bovine serum (Hyclone, Ogden, UT), 2mM L-glutamine, 1mM sodium pyruvate, 0.1mM non-essential amino acids and 1x penicillin/streptomycin (GIBCO-Invitrogen, Carlsbad, CA). HCCLM3 and MHCC97-H were kindly provided by Prof. Tang Zhao-You and Liu Bin Bin (Liver Cancer Institute and Zhongshan Hospital, Fudan University, Shanghai, China), and were maintained in the same growing media, but without non-essential amino acids and penicillin/streptomycin. All cells were cultured in a humidified 37°C incubator with 5% CO₂. Cell lines were routinely checked for mycoplasma infection by staining with DAPI to detect for cytoplasmic and intercellular foci fluorescence. Suspected mycoplasma infections were cleared using Ciprofloxacin Hydrochloride (Cellgro, Manassas, VA).

2.1.14 DNA transfection

All cells were seeded in 6-well culture plate or glass coverslip within the 6-well plates one day ahead so that they reached 50 to 80% confluency at the day of transfection. 3 µg of plasmid DNA constructs diluted in 500 µl OptiMEM (Invitrogen, Carlsbad, CA) were combined with 3 µl of PLUS reagent and incubated at room temperature for 15 min before further incubation with 7 µl of Lipofectamine LTX (both from Invitrogen, Carlsbad, CA) for 25 min. All the DNA-Lipofectamine complex incubation was performed in polypropylene tubes before transferring to each well in a final volume of 1 ml OptiMEM. Transfection mixture was removed from each well after 4 to 6 hr incubation, and replaced with 2 ml fresh growing media.

2.1.15 siRNA transfection

Hepatocellular carcinoma cell lines: HCCLM3, MHCC97-H and PLC/PRF 5 were plated in 6-well plates one day ahead such that they reached 50 to 60% confluency at the time of transfection. 5 µl of Lipofectamine 2000 (Invitrogen, Carlsbad, CA) was incubated with 250 µl OptiMEM at room temperature for 5 min. 100 nM of the combined three siRNA duplexes targeting gene of interest and Stealth RNAi universal negative control medium GC duplex (both from Invitrogen, Carlsbad, CA) diluted in 250 µl OptiMEM was independently added with the Lipofectamine 2000/OptiMEM mixture for 20 min at room temperature. The siRNA-Lipofectamine complex was added to each well to a final volume of 1ml OptiMEM. After 4 hr

incubation, transfection medium was removed and replenished with fresh DMEM media with 10% FBS.

2.1.16 Western blot analysis

Cells were lysed with lysis buffer for 10 to 15 min on ice. Lysates were centrifuged for 15 min at 13,000 rpm, 4°C and supernatant was saved for protein quantitation using Bio-Rad Protein Assay (BioRad Laboratories, Hercules, CA) by reading absorbance at 595 nm via spectrophotometry. Thirty to 50 µg of total protein was resolved on 7% or 12% SDS-PAGE mini-gels, depending on the size of the protein to be detected, and was later transferred to HybondTM-C extra supported nitrocellulose membrane (Amersham, UK) by using wet transfer system at constant voltage 20V, 4°C overnight or at constant 110V, 4°C for 2 hr. The blots were either blocked with 5% nonfat dry milk or 3% BSA at room temperature for 1 hour (depend on the primary antibody) before incubating with primary antibody at 4°C overnight. Incubation with secondary antibody was performed at room temperature for 1 hr. The blots were finally developed using ImmobilonTM Western Chemiluminescent HRP substrate (Millipore) or SuperSignal® West Pico Chemiluminescent substrate (Pierce, Rockford, IL) and CL-Xposure film (Pierce, Rockford, IL). Blots were stripped and re-probed several times using Restore stripping buffer (Pierce, Rockford, IL) for 20 min at 42 °C or stripping buffer containing β-mercaptoethanol and SDS.

2.1.17 Cell migration and invasion assays

Migration assays were performed in Boyden chambers with polyethylene terephthalate filter membrane in 24 well plates containing 0.8 μm pores (BD Falcon, NJ, USA). Invasion assays were done in a similar way except that the filter inserts were readily precoated with growth factor reduced (GFR) Matrigel. To examine the RACGAP1 silencing and the overexpressing effect on migration and invasion, HCCLM3 and MHCC97H were transfected with siRNA targeting RACGAP1 for 2 days whereas PLC / PRF 5 cells were transfected with pIR and pIR-Rg plasmid DNA for 24 hours. 1.5×10^5 of the transfected cells were plated in 500 μl Opti-MEM medium into the inserts. The companion well was filled with 750 μl 10% fetal bovine serum (FBS)-DMEM as the chemoattractant. After culture for 48 hours, the non-invaded cells in the innerside of the inserts were removed with cotton swabs. The migrated or invaded cells on the underside were fixed with 1% formaldehyde and permeabilized with 0.2% Triton X-100 before staining the nuclei with Vectashield mounting containing DAPI (Vector Laboratories, Burlingame, CA). The migrated and invaded cells were photographed at five different fields using Nikon eclipse 90i fluorescent microscope and the average was taken. Each treatment was repeated in three independent experiments. Statistical analysis was performed using parametric Student's T-test from GraphPad Prism 5.0 software.

2.1.18 Cdc42, Rac1 and Rho A activation assay

Active form of Rho family small GTPases was detected using EZ-Detect Cdc42, Rac1 activation kit (both from Pierce, Rockford, IL) and Rho activation kit (Upstate, Temecula, CA). Cells were harvested into the ice-cold lysis buffer provided in the kit supplemented with protease inhibitors (Roche, Mannheim, Germany). The lysate was cleared by centrifugation at 4 °C and the supernatant was stored in aliquots at -80 °C if not used immediately. For Cdc42 and Rac1 pull down assay, 1 mg of total protein of each sample was incubated overnight at 4°C with GST-fusion protein containing p21-binding domain (PBD) of human Pak1 (GST-human Pak1-PBD) that specifically pulls down active or GTP-bound Cdc42 and Rac1.

For Rho pull down assay, the same amount of protein was incubated with GST-fusion protein containing Rho binding domain (GST-Rhotekin-RBD). Unbound inactive small GTPases (GDP-bound Cdc42, Rac1 and Rho) were washed a few times and supernatant was discarded. The active small GTPases were then resolved on 12% SDS-PAGE and transferred to nitrocellulose membranes (Amersham, Arlington Heights, UK), and finally detected by Cdc42, Rac1 and Rho antibody. To ensure the pull down procedures are working properly, positive and negative controls employing 0.1mM GTP-gamma-S and 100mM GDP were added separately into the same amount of lysates. The control mixtures were incubated at 30°C for 15 min and the reaction was terminated by adding 1M MgCl₂ to make up 60mM in final reaction. The downstream procedures were same as the other pull down assays, as per recommended

by the manufacturer's protocol. Quantitation of protein bands was done using Quantity One software (Bio-Rad Laboratories, Hercules, CA).

2.1.19 DNA fragmentation (TUNEL) assay by FACS and microscopy

MHCC97-H cells were grown triplicates in a 6 well plates one day prior to transfection with siRNA duplexes. Cells were harvested after 2, 4 and 6 day treatment and washed gently with 1x PBS and fixed with prechilled 1% formaldehyde for 20 minutes at 4°C. The fixed cells were then incubated with terminal deoxynucleotidyl transferase (TdT) and fluorescein labeled nucleotide mix for 1 hour at 37°C, and counterstained with propidium iodide (PI) in the presence of 0.5 µg/ml DNase-free RNase A (Sigma, Missouri, USA) as recommended in the protocol of ApoAlert® DNA Fragmentation assay kit (Clontech, Mountain View, CA). For experimental controls, a DNase-treated positive control was prepared by treating fixed cells with 100µl of DNase I Buffer containing 0.5 to 1µg/ml DNase (New England Biolabs, Ipswich, MA). For TdT-minus negative control, TdT enzyme was replaced with deionized water. The apoptotic cells exhibiting strong nuclear green fluorescence at 520±20nm were detected by BD FACSCanto™ II flow cytometer (BD Biosciences, San Jose, CA) or by fluorescence microscopy after counterstaining nuclei with DAPI, using Confocal Laser Scanning Microscope LSM 510 Meta (Carl Zeiss, Germany).

2.1.20 FACS analysis of Ki-67 proliferation assay

PLC/PRF 5 cells were harvested after 24 hours transfected with pIRES vector and RACGAP1 constructs and washed in 1x FACS staining buffer. Cells were then fixed and permeabilized with BD Fixation/Permeabilization solution and BD Perm/Wash buffer as recommended in BD Cytofix/Cytoperm Fixation/Permeabilization kit manual (BD Biosciences, San Jose, CA). Fixed cells were first incubated with mouse anti-human RACGAP1 primary antibody (Abnova, Taiwan) for 30 minutes at 4°C, followed by co-incubation of FITC-conjugated mouse anti-human Ki-67 antibody (BD Biosciences, San Jose, CA) and Alexa Fluor 546-conjugated goat anti-mouse secondary antibody (Molecular Probes) for another 30 minutes at 4°C. The transfected cells that co-express RACGAP1 and Ki-67 were detected by BD FACSCanto™ II flow cytometer (BD Biosciences, San Jose, CA).

2.1.21 Network Analysis of gene expression in cells depleting RACGAP1

The differential genes that had false discovery rate less than 10%, p value less than 0.05 and fold change greater than 1.25 between RACGAP1 siRNA-treated cells and RNAi negative control-treated cells were selected. The dataset was uploaded into a web-based software, Ingenuity Pathway Analysis 8.5 (Ingenuity Systems, Redwood City, California) and core-analyzed in the context of biological processes, pathway and networks. The analysis was performed using general setting of 70 molecules per network and 25 networks per analysis. The top 10 to 20 canonical pathways that were

deregulated after silencing RACGAP1, as ranked by the log p-value score generated by Fisher's Exact Test were kept for further analysis. The networks of RACGAP1 including other closely associated molecules that are deregulated was overlaid with a few top canonical pathways selected earlier as well as the RACGAP1 neighborhood genes which are literally curated in this software.

2.1.22 Protein profiling via LC-MS/MS analysis of cells depleting RACGAP1

Four days after treatment with RACGAP1 siRNA duplexes and RNAi universal negative control duplexes, medium supernatant of the cells was centrifuged to pellet the detached dead cells. Adherent live cells were harvested by trypsinization using 1 X Trypsin/ EDTA and pelleted. Both the live and dead cells pellets were combined and washed with 1 X PBS (pH 7.2) three times and pelleted by centrifugation before added with 200-600 μ L of lysis buffer containing 8 M urea, 4% (w/v) CHAPS, and 0.05% SDS (w/v). The lysate was vortexed vigorously for 1 min, and kept on ice for 20 min. After clearing by centrifugation at 15,300 rpm, 4 °C for 1 h, the supernatant containing the total protein was aliquoted and stored at -80 °C if not used immediately. The protein concentration was quantitated using the 2-D Quant Kit (GE Healthcare). To eliminate substances in protein sample that interfere with downstream applications, 100 μ g total protein from each sample was precipitated by adding four sample volume of cold (-20°C) acetone. Following vortexing, the samples were incubated for 2 hours at -20 °C and centrifugated at 13,000-15,000 x g for 10

min to obtain the protein pellet. The protein pellet in the microcentrifuge tube was left uncapped for about 30 min at room temperature to evaporate the residual acetone.

The precipitated protein pellet was kept at -80 °C or subjected to labeling with iTRAQ reagent by adding with 20 µl of dissolution buffer to dissolve the proteins. Following adding of one microliter of the denaturant to each tube to disrupt the hydrogen, hydrophobic, and electrostatic bonds of the proteins, 2 µl of reducing reagent was also added to reduce the disulfide bonds of the proteins. After incubation for 1 h at 60 °C, the cysteine group of proteins was blocked reversibly by adding with 1 µl cysteine-blocking reagent, and left at room temperature for 10 min. The proteins were then digested with 20 µl of 0.25 µg/ µl modified trypsin (Promega) at 37 °C overnight. Seventy microliter of ethanol was added to each vial of iTRAQ reagent and the mixture was transferred to one sample tube for reaction. After 1 hour reaction, each iTRAQ reagent-labeled sample tube were combined into a fresh tube, kept at -80 °C for subsequent experiments. The labeling with the iTRAQ tags are as follows: RNAi universal negative control-treated cells = iTRAQ 114; RACGAP1 siRNA duplexes-treated cells = iTRAQ 115. To achieve statistical confidence for the differential expressed proteins, two independent experiments for each treatment were performed. All the chemicals and reagent used for iTRAQ labeling in this study are from Applied Biosystems, unless stated otherwise. To eliminate the interference of sample preparation procedures with the digestion and labeling procedures, BSA standard solution (Pierce) was used as a technical negative control undergoing the

same experimental procedures as described above. These differentially labeled digests analyzed by LC-MS/MS.

The raw data analysis including protein identification was conducted using an Agilent 1200 nanoflow LC system (Agilent Technologies) interfaced with a QSTAR XL mass spectrometer (Applied Biosystems/MDS Sciex). Peptides were identified by using ProID software packages (Applied Biosystems). Each MS/MS spectrum was traced against the *Homo sapiens* protein database. To eliminate false positive results, all protein identification interpreted from single unique peptides were ignored and proteins were accepted only when their ProtScore values were greater than 2.0 with confidence of 99%. The database allowed for iTRAQ labelling at N-terminal residues, internal K and Y residues, and the methylmethanethiosulfate-labeled cysteine as fixed modification, together with one missed cleavage. The quantitative and statistical analysis were performed using ProQUANT 1.0. The tolerance settings for peptide identification in ProQUANT searches were 0.15 Da for MS and 0.1 Da for MS/MS. A peptide in each sample was quantitated by obtaining the relative amount of the peak areas at 115.1 m/z against the peak at 114.1 m/z . The relative peak area were then corrected for overlapping isotopic contributions, and were used to estimate the relative abundances of a particular peptide. To minimize the variation during protein loading, these ratios were again normalized against the overall ratios of all proteins in the sample, as recommended by the manufacturer (Applied Biosystems). In this study, the differentially expressed proteins with ratio greater than 1.1 and less than 1.0 were selected for further analysis.

2.2 Primers for PCR and quantitative realtime PCR

(The orientation symbols of the primers listed below, such as (F) and (R) might have been by no means commonly used before by other people . However, the sequences of all the primers are designed specifically for this study)

Primers for plasmid linearization

Linear EcoR1 (F) : 5'-GAATTCGAAGCTTGAGCTCGAGA-3'

Linear Sal1 (R) : 5'-GTCGACGGTACCGCGGGCCCGGGAT-3'

Primers for cloning of full length RACGAP1

Rg_EcoR1(F): 5'-CTCAAGCTTCGAATTCCCACCATGGATACTATGATGCTG-3'

Rg_Sal1(R) : 5'-CCGCGGTACCGTCGACCTTGAGCATTGGAGAAGCAAA-3'

Primers for screening insert orientation in pIRES2-RACGAP1-AcGFP1 constructs

Forward/correct orientation screening:

p520-41 (F) : 5'- GCAAATGGGCGGTAGGCGTGTA-3'

Rg_seq1-640(R) : 5'-TCATTCCCCTGGTCTACTGC-3'

Reverse/wrong orientation screening:

p709-29(F) : 5'- GCACACCGGCCTTATTCCAAG-3'

Rg_seq1-640(R) : 5'-TCATTCCCCTGGTCTACTGC-3'

Primers for realtime PCR

RACGAP1 (F) : 5'-CAGTGACTCCACCCTGAACA-3'

RACGAP1 (R) : 5'-CTGGATGAGAGACCACACG-3'

KCNK (F) : 5'-CTGCAAACCATTGAGCGTAG-3'

KCNK (R) : 5'-GGTCACAGCTTCTTTGTTCCA-3'

SMURF2 (F) : 5'-TTGGCTCTGCAGAAAGGATT-3'

SMURF2 (R) : 5'-AATCTTGCTCGTCGCTCTTC-3'

USH1C (F) : 5'-GACTGGATCGACCTTGTGGT-3'

USH1C (R) : 5'-GATGCCATCTGGTGTGTGTC-3'

GSTM3 (F) : 5'-ATCTGCCCTACCTCCTGGAT-3'

GSTM3 (R) : 5'-GGCTTCAGTTTTTCGTGGTC-3'

CNGA1 (F) : 5'-CCCAGATGCCAAAACATATGC-3'

CNGA1 (R) : 5'-GTTTCTGCTGCATGGACTCA-3'

INSIG1 (F) : 5'-TACGCTGATCACGCAGTTTC-3'

INSIG1 (R) : TTGCTCTCAGAATCGGTGGT-3'

2.3 siRNA sequences

(The orientation symbols of the primers listed below, such as Forward and Reverse might have been by no means commonly used before by other people . However, the sequences of all the primers are designed specifically for this study)

RACGAP1 (pool of 3 sequences, R1) (Stealth RNAi, Invitrogen)

Sequence 1 – Forward: 5'-CACACUGUCUGUCUCAGUUCUUGGC-3'

Reverse: 5'-GCCAAGAACUGAGACAGACAGUGUG-3'

Sequence 2 – Forward: 5'-UUUACUGUGCGGUCACAGCCAGAGA-3'

Reverse: 5'-UCUCUGGCUGUGACCGCACAGUAAA-3'

Sequence 3 – Forward: 5'- UUGCCUUGUCGUCCUAGGUUAGUGG-3'

Reverse: 5'- CCACUAACCUAGGACGACAAGGCAA-3'

Three sequences are combined to a final concentration of 100nM in each transfection

RACGAP1 (pool of 3 sequences, R2) (Stealth RNAi, Invitrogen)

Sequence 1 – Forward: 5'- UAUACAGGCCUGUCUCAGUCAGACC-3'

Reverse: 5'- GGUCUGACUGAGACAGGCCUGUAUA -3'

Sequence 2 – Forward: 5'- UUCUGCUGCUUCCAUAAGGCUCUG -3'

Reverse: 5'- CAGAGCCUUUAUGGAAGCAGCAGAA-3'

Sequence 3 – Forward: 5'- UUGAGAAGCUGAUGUUCAGGAGUGG-3'

Reverse: 5'- CCACUCCUGAACAUCAGCUUCUCAA-3'

Three sequences are combined to a final concentration of 100nM in each transfection

2.4 Reagents and antibodies

(The chemicals, reagents, formulation and antibodies listed below might have been by no means commonly used before by other people for a similar approach of different experimental study. However, all the reagents and antibodies were prepared and purchased to suit for a specific experimental requirement in this study)

2.4.1 General buffers

DEPC-treated water – 0.1% Diethylpyrocarbonate (Sigma) in double distilled water, stir overnight and inactivated by autoclaving

10x TBE – 89mM Tris-Base; 89mM boric acid; 2mM EDTA

10x sample loading Buffer – 0.25% (w/v) Bromophenol Blue; 25% (w/v) Xylene Cyanol; 50% Glycerol (v/v)

Phosphate buffered saline – 137mM NaCl; 2.7mM KCl; 4.3mM Na₂HPO₄; 1.4mM KH₂PO₄; pH 7.4

2.4.2 Reagents and materials for Affymetrix GeneChip experiment

T7-(dT)₂₄ Promoter Primer (HPLC purified) (Sigma-Proligo)
5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(dT)₂₄-3'

GeneChip® Human Genome U133 Set :

HG-U133A, HG-U133B arrays.

GeneChip® IVT Labelling kit components (30 reactions): 10x IVT labeling buffer, IVT labeling enzyme mix, IVT labeling NTP mix, 3'-labeling control (0.5µg/ul), RNase-free water.

GeneChip® Eukaryotic Hybridization control kit components (30 reactions): 20x hybridization control solutions composed of pre-mixed biotin-labelled bioB, bioC, bioD and cre for monitoring the hybridization process to facilitate troubleshooting, and Control Oligo B2 facilitating alignment signals for image analysis.

5x cRNA fragmentation buffer (200mM Tris-acetate, pH8.1; 500mM KOAc; 150mM MgOAc)

4.0 ml 1M Tris acetate pH 8.1 (Trizma Base, pH adjusted with glacial acetic acid)

0.64 g MgOAc

0.98 g KOAc

DEPC-treated H₂O adjust to 20 ml

12x MES stock buffer (1.22M MES; 0.89M [Na⁺])

70.4 g	MES free acid monohydrate
193.3 g	MES Sodium salt
800 ml	ddH ₂ O

Mix and adjust volume to 1000 ml. The pH should be between 6.5 and 6.7. Filter through 0.2 µm filter.

2x Hybridization buffer (200mM MES; 2M[Na⁺]; 40mM EDTA; 0.02% Tween 20)

8.3 ml	12x MES stock buffer
17.7 ml	5M NaCl
4 ml	0.5 M EDTA
0.1ml	10% Tween 20
19.9 ml	ddH ₂ O

Filter through 0.2 µm filter. Store at 4°C, and shield from light.

2x Stain buffer (200mM MES; 2M [Na⁺]; 0.1% Tween 20)

41.7 ml	12x MES stock buffer
92.5 ml	5 M NaCl
2.5 ml	10% Tween 20
113.3 ml	ddH ₂ O

Filter through 0.2 µm filter. Store at 4°C, shield from light.

Streptavidin phycoerythrin (SAPE) solution (1xMES; 2mg/ml acetylated BSA; 10ug/ml SAPE)

600 µl	2x Stain buffer
48 µl	50 mg/ml acetylated BSA
12 µl	1 mg/ml Phycoerythrin streptavidin (Molecular Probes)

540 μ l ddH₂O

Biotinylated anti-streptavidin IgG solution (1x MES; 2mg/ml acetylated BSA; 100 μ g/ml normal goat IgG; 3 μ g/ml biotinylated anti-streptavidin IgG)

300 μ l 2x Stain buffer

24 μ l 50 mg/ml acetylated BSA

6 μ l 10 mg/ml normal goat IgG (Sigma)

3.6 μ l 0.5 mg/ml biotinylated anti-streptavidin IgG (Vector Laboratories)

266.4 μ l ddH₂O

20x SSPE – 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA; adjust pH to 7.4

Wash A (Non-stringent wash buffer) (6x SSPE; 0.01% Tween 20)

300 ml 20x SSPE

1.0 ml 10% Tween-20

699 ml ddH₂O

Filter through 0.2 μ m filter

Wash B (Stringent wash buffer) (100mM MES; 0.1 M [Na⁺]; 0.01% Tween 20)

83.3 ml 12x MES stock buffer

5.2 ml 5 M NaCl

1 ml 10% Tween 20

910.5 ml ddH₂O

Filter through 0.2 μ m filter. Store at 4°C, shield from light.

2.4.3 Buffers for Immunohistochemistry

Wash buffer – 0.05% Tween 20; 1x PBS

Antigen retrieval buffer – 10mM citrate acid (pH 6.0); ddH₂O

Endogenous peroxidase blocking buffer – 3% H₂O₂; 0.015 M NaN₃; 1x PBS

Blocking solution – 5% Normal goat serum; 0.05% Tween 20; 1x PBS

Diluent buffer for 1^oC antibody - 0.05 M Tris-HCl (pH 7.2); 0.1% Tween; 0.015 M NaN₃

Dako REALTM EnVision Detection system, Peroxidase/DAB+, Rabbit/Mouse (500 test)

Dako REALTM EnVisionTM/HRP, Rabbit Mouse (ENV) – goat secondary antibody against mouse and rabbit IgGs conjugated with peroxidase-coupled dextran

Dako REALTM Substrate buffer – buffered solution containing H₂O₂

Dako REALTM DAB+ Chromogen - 3,3'-diaminobenzidine tetrahydrochloride

2.4.4 Reagents and materials for cloning and bacteria transformation

LB broth – 1% (w/v) Bacto-tryptone; 0.5% (w/v) Bacto-yeast extract; 1% (w/v) NaCl

LB Agar - 1% (w/v) Bacto-tryptone; 0.5% (w/v) Bacto-yeast extract; 1% (w/v) NaCl; 2% (w/v) bacto-agar

SOC media – 2% Bacto-tryptone; 0.5% yeast extract; 10mM NaCl; 2.5mM KCl; 10mM MgCl₂; 10mM MgSO₄; 20mM glucose

Figure 2-1. Map features and multiple cloning site of pIRES2-AcGFP1 vector.
MCS, multiple cloning site.

2.4.5 Buffers for protein analysis

Lysis Buffer – 10mM Tris-HCl (pH 7.4); 150mM NaCl; 1mM EDTA; 1% Triton X-100; 0.5% NP-40; 1x Complete protease inhibitor cocktail (Roche, Germany); 1x PhosSTOP (Phosphatase inhibitor cocktail) (Roche, Germany)

5x sample loading buffer – 0.3M Tris, pH 6.8; 10% (w/v) SDS; 30% (v/v) glycerol; 10% (v/v) β -mercaptoethanol; 0.002% (w/v) bromophenol blue

10x running buffer – 0.25M Tris, pH7; 1.92M glycine; 1% SDS (Bio-Rad)

1x transfer buffer – 25mM Tris-Base; 192mM glycine; 17.5% (v/v) methanol

Wash buffer - 0.05% Tween 20; 1x PBS

Blocking buffer – 5% (w/v) nonfat dry milk or 3%(w/v) BSA (depending on 1^o antibody); 1x PBS; 0.05% Tween 20

Diluent buffer for 1^o antibody - 5% (w/v) nonfat dry milk or 3%(w/v) BSA (depending on 1^o antibody); 1x PBS; 0.05% Tween 20; 0.1% NaN₃

Diluent buffer for 2^o antibody - 5% (w/v) nonfat dry milk; 1x PBS; 0.05% Tween 20

Stripping buffer – 100mM β-mercaptoethanol; 62.5mM Tris-HCl, pH 7; 2% SDS

Table 2.1. Protocol for preparing SDS-PAGE gels

Components	Resolving gel			Stacking gel
	7%	10%	12%	
ddH ₂ O	7.44 ml (3.72)	5.94 ml (2.97)	4.94 ml (2.47)	5.92 ml (2.96)
30% (w/v) Acrylamide/Bis solution (37.5:1) (Bio-Rad)	3.5 ml (1.75)	5 ml (2.5)	6 ml (3.0)	1.34 ml (0.67)
1.5M Tris, pH 8.8 (Bio-Rad)	3.75 ml (1.875)	3.75 ml (1.875)	3.75 ml (1.875)	-
0.5M Tris, pH 6.8 (Bio-Rad)	-	-	-	2.52 ml (1.26)
10% (w/v) SDS (Bio-Rad)	0.15 ml (0.075)	0.15 ml (0.075)	0.15 ml (0.075)	0.10 ml (0.05)
10% (w/v) Ammonium persulfate (Bio-Rad)	0.15 ml (0.075)	0.15 ml (0.075)	0.15 ml (0.075)	0.10 ml (0.05)
TEMED (Bio-Rad)	0.02 ml (0.01)	0.02 ml (0.01)	0.02 ml (0.01)	0.02 ml (0.01)
Total	15.01 ml	15.01 ml	15.01 ml	10.0 ml

Note: numbers in brackets represent volumes of reagents for 1 mini-gel

2.4.6 Buffers for Flow Cytometry (FACS)

PI staining buffer – 1x PBS; 0.1% (v/v) Triton-X 100; 50ug/ml PI; 100ug/ml RNase A

1x FACS staining buffer - 5% (v/v) fetal bovine serum; 5%(v/v) human serum; 0.01% NaN₃; 1X PBS

2.4.7 Antibodies

RACGAP1 antibody – mAb, mouse IgG2b, 0.5mg/ml (Abnova); working dilution – 1:500

Cdc42 antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Rac1 antibody – mAb, mouse IgG2b (Pierce); working dilution – 1: 1000

Rho (A, B, C) clone 55 antibody – mAb, mouse IgG1, 0.727mg/ml (Upstate); working concentration – 3ug/ml

Actin, pan Ab-5 (Clone ACTN05) – mAb, mouse IgG1,κ, 0.2mg/ml (Termo Scientific); working dilution – 1:5000

Caspase 7 antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Cleaved Caspase 7 (Asp198) antibody - pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Caspase 9 antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Cleaved Caspase 9 (Asp330) antibody - pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

PARP antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Cleaved PARP (Asp214) antibody - pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

ECT2 antibody – pAb, goat IgG, 0.5mg/ml (Imgenex); working dilution – 1:200

KIF23 (MKLP-1 (N-19):sc-867) antibody – pAb, rabbit IgG, 0.2mg/ml (Santa Cruz); working dilution – 1:200

AURKB (ARK-2 (H-75):sc-25426) antibody - pAb, rabbit IgG, 0.2mg/ml (Santa Cruz); working dilution – 1:200

Phospho-Aurora A(Thr288)/Aurora B(Thr232)/Aurora C(Thr198) antibody – mAb, rabbit IgG (Cell Signaling); working dilution – 1:2000

PRC1 antibody – mAb, rabbit IgG (abcam); working dilution – 1:5000

Phospho-PRC1 (Thr481) antibody – mAb, rabbit IgG (abcam); working dilution – 1:500

BRCA1 antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Phospho-BRCA1 (Ser1524) antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

PLK1 antibody – mAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

PLK4 antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

p44/42 MAP Kinase (Erk1/Erk2) antibody – pAb, rabbit IgG (Cell Signaling);
working dilution – 1:1000

Phospho-p44/42 MAP Kinase (phospho-Erk1/Erk2) antibody – mAb, mouse IgG1
(Cell Signaling); working dilution – 1:2000

Akt antibody – pAb, rabbit IgG (Cell Signaling); working dilution – 1:1000

Phospho-Akt (Thr308) antibody – pAb, rabbit IgG (Cell Signaling); working dilution
– 1:1000

Secondary Horseradish Peroxidase-conjugated anti-mouse antibody - pAb, goat IgG
(H+L), 0.8mg/ml (Pierce); working dilution – 1:10,000

Secondary Horseradish Peroxidase-conjugated anti-rabbit antibody – pAb, goat IgG
(Dako); working dilution – 1:2000

Secondary Horseradish Peroxidase-conjugated anti-goat antibody – pAb, rabbit IgG
(H+L) (Pierce); working dilution – 1: 10,000

FITC-conjugated anti-human Ki-67 – mAb, mouse IgG1, κ (BD Pharmingen); working
dilution – 1: 50

R-PE-conjugated anti-mouse – pAb, goat F(ab')₂ (Biosource); working dilution –
1:200

CHAPTER 3

Identification of molecular markers associated with the recurrence of human hepatocellular carcinoma by microarray analysis

3.1 Results

3.1.1 Clinical characteristics of HCC samples

To ensure that the results obtained following the analyses were representative for the overall HCC patients in this study, all consecutively and clinically obtained well-characterized HCC patients were included, without focusing on pre-selected subgroups of patients. The median age of 80 HCC patients in this study was 62 years old (range: 30 – 87 years old). Male patients were predominant with its ratio to female, 4.3:1 (81.2%; n=65 compared to 18.8%; n=15). Of all the patients in this study, 77.5% (n=62) were HBV carriers, 3.7% (n=3) were HCV carriers and 18.8% (n=15) were negative for both viral infection. There were no cases of hepatitis B and C coinfection in the present study. Almost half of the patients (48.8%; n=39) with resected tumor were partially and completely encapsulated while 51.2% (n=41) were not capsulated, and 53.8% (n=43) of patients had cirrhosis compared to 46.2% (n=37) without cirrhosis.

The majority of patients had single nodule (86.2%; n=69) compared to 13.8% (n=11) patients with multiple nodules. The median size of all resected tumor was 5.0

cm (range: 1.2 – 17 cm) and the α -fetoprotein (AFP) level was 25.6 ng/ml (range: 1.2 – 70700 ng/ml). The categorization of α -fetoprotein (AFP) revealed that 37.5% (n=30) cases scored low AFP level (< 10 ng/ml), 35% cases (n=28) had medium level (10 – 300 ng/ml) and 26.2% (n=21) patients had elevated AFP level (> 300 ng/ml). More than half of the cases (55%; n=44) were moderately differentiated with grade 2 whereas 12.5% (n=10) were well-differentiated with grade 1 and 25% (n=20) were grade 3. Very low number of cases (6.2%; n=5) had the poorest differentiation grade 4. Vascular invasion occurred in 36.2% (n=29) of the cases while 63.8% (n=51) of HCC patients had no invasion, and the median time of recurrence was 5.5 months (range: 0.6 – 35.2 months) in this study.

3.1.2 Vascular invasion and cirrhosis are associated with HCC recurrence

In an attempt to identify clinicopathological parameters that could be of significant diagnostic value to predict disease recurrence, we perform univariate analysis to assess the correlation of several common clinicopathological features that are available at diagnosis with the recurrence status of 80 HCC patients. The results in Table 3-1 showed that recurrence was associated significantly with vascular invasion ($P=0.006$), and weakly with cirrhosis ($P=0.07$) (154). No other clinicopathological parameter was found to be associated with the recurrence of HCC although patients with recurrence apparently had larger tumor size (median = 6cm) and elevated serum concentration of AFP (median = 41.35 ng/ml) compared to patients without recurrence (median AFP = 15.1 ng/ml).

Table 3-1. Univariate analyses to evaluate the correlation between clinicopathological factors and disease recurrence of 80 HCC patients (154)

Variable	Recurrence (n=39)	Non- recurrence (n=41)	<i>P</i> value ^a
Gender (men/women)	34/5	31/10	0.185
Age (median) (years)	62	61	0.607
Viral infection (HBV / HCV / non-B non-C)	29/2/8	33/1/7	0.738
Capsulation (partial or complete / no)	17/22	22/19	0.656
Tumor size (median) (cm)	6	4.2	0.132
AFP level (median) (ng/ml)	41.35	15.1	0.223
[low / medium / high]	[12/14/12] [¶]	[18/14/9]	
Lesion (single / multiple nodules)	33/6	36/5	0.679
Histological grading (G1/G2/G3/G2 to G4)	3/20/13/2 [¶]	7/24/7/3	0.191
Cirrhosis (yes / no)	25/14	18/23	0.070
Invasion (yes / no)	20/19	9/32	0.006

HBV, hepatitis B virus; HCV, hepatitis C virus; non-B non-C, negative for both HBV and HCV antigen; low, less than 10 ng/ml; medium, 10 to 300 ng/ml; high, more than 300 ng/ml; G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated; G4, extremely poor differentiation; [¶], clinical factor is not available at diagnosis in one recurrence case; ^a, *P* values were calculated using univariate analysis.

Multivariate analysis was employed to determine if clinicopathological features are significantly associated with the recurrence status of 80 HCC patients divided into the duration of 6 months, 12 months, 18 months and 35 months after hepatectomy resection. The findings ascertained that vascular invasion and cirrhosis as two significant independent risk factors for recurrence of HCC irrespective of various earlier durations of recurrence examined (Table 3-2). The risk of developing disease recurrence from as early as 6 months till up to 35 months after curative hepatectomy is

significantly higher in the presence of either microscopic vascular invasion or cirrhosis at diagnosis.

Table 3-2. Multivariate analyses to assess the association of clinicopathological parameters with HCC recurrence occurred at different duration after resection (154)

Variable	<i>P</i> value ^a			
	R≤6 months	R≤12 months	R≤18 months	R≤35 months
Gender (men/women)	0.196	0.096	0.073	0.040
Age (median) (years)	0.179	0.919	0.914	0.782
Viral infection (HBV / HCV / non-B non-C)	0.922	0.872	0.968	0.760
Capsulation (yes / no)	0.548	0.698	0.656	0.866
Tumor size (median) (cm)	0.609	0.755	0.796	0.379
AFP level (median) (ng/ml) [low / medium / high]	0.092	0.177	0.215	0.215
Lesion (single / multiple nodules)	0.930	0.572	0.487	0.659
Histological grading (G1/G2/G3/G2 to G4)	0.983	0.913	0.837	0.567
Cirrhosis (yes / no)	0.038	0.018	0.014	0.011
Invasion (yes / no)	0.001	0.0001	0.0001	0.0001

HBV, hepatitis B virus; HCV, hepatitis C virus; non-B non-C, negative for both HBV and HCV antigen; low, less than 10 ng/ml; medium, 10 to 300 ng/ml; high, more than 300 ng/ml; G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated; G4, extremely poor differentiation; R, duration of recurrence. ^a*P* value were calculated using multivariate cox regression analysis.

3.1.3 Stratification of HCC patients to recurrence and non-recurrence groups using either vascular invasion or cirrhosis at a time

We grouped 80 HCC patients into 2 subgroups, vascular invasion (n=29) and non-invasion (n=51), with each subgroup further divided into recurrence and non-recurrence. The same subdivision was applied to cirrhosis (n=43) and non-cirrhosis (n=37). The samples used in grouping invasion and cirrhosis respectively came from the same pool of patients. The result in Figure 3-1 showed that majority of the vascular invasion cases (69%; n=20), the yellow bars, had recurrence and slightly more than half of cirrhosis (58%; n=25), the blue bars had recurrence. The adverse patterns were observed in non-invasion cases in which majority, 63% (n=32), the light yellow bars had no recurrence and also 62% (n=23) of patients without cirrhosis, the light blue bars had no reported recurrence.

Figure 3-1. Stratification of recurrence and recurrence-free cases according to one clinical factor, invasion or cirrhosis at a time. Higher recurrence rate was observed in HCC patients either diagnosed with vascular invasion or cirrhosis at a time, while adverse patterns occurred in patients either had non-invasion or non-cirrhosis respectively. R, recurrence; NR, non-recurrence.

3.1.4 Stratification of HCC patients using the combined vascular invasion and cirrhosis reveals various recurrence risk groups

To determine if the recurrence risk would paramount by considering these 2 clinical factors simultaneously, we arbitrarily divided the 80 HCC patients into four subgroups in accordance to cirrhosis and vascular invasion jointly. Consequently, HCC patients who had vascular invasion and cirrhosis at diagnosis are subgroup 1, while HCC patients who had no vascular invasion and were non-cirrhotic at diagnosis are subgroup 4. Patients belonging to subgroups 2 and 3 were individuals, at diagnosis, with either vascular invasion but no cirrhosis, or with cirrhosis but no

vascular invasion, respectively. For HCC with recurrence occurred within six months (N=63) after surgical resection, those diagnosed with vascular invasion and cirrhosis (subgroup 1) were most likely to recur (78%), while patients diagnosed with non-invasion and non-cirrhotic (subgroup 4), were all recurrence-free (Figure 3-2A). Similar patterns were observed when the duration of recurrence increased to 12 months (N=74), 18 months (N=76) and 35 months (N=80) (Figure 3-2B, C and D).

Overall, patients who were positive for both vascular invasion and cirrhosis at diagnosis (subgroup 1) had the highest rate of developing recurrent disease, in contrast to patients that were negative for both lesions (subgroup 4). Hence, the combined clinicopathological features of both the vascular invasion and cirrhosis were adequately informative for the development of disease recurrence in HCC patients from subgroup 1 (78% to 83% of patients developed recurrence) and subgroup 4 (89% to 100% of patients remained disease free) (Figure 3-2). This suggested that patients adopted features as in Group 1 might have the most advanced and malignant stage in this study while the adverse condition occurred in Group 4 patients. However, in both subgroups 2 and 3 which covered more than half (> 60%) of the samples in all cutoff of recurrence duration (Figure 3-2), the combined features of vascular invasion and cirrhosis was apparently insufficient to estimate the risk of recurrence for these patients. In all 80 patients, slightly larger number of subgroup 2 (59%), and almost half of the subgroup 3 patients (48%) developed recurrence (Figure 3-2D). This suggested clinicopathological features of vascular invasion and the status of cirrhosis were not clinically useful, at diagnosis, to estimate the risk of recurrence for HCC

patients in subgroups 2 and 3, who have one lesion at diagnosis, either vascular invasion or cirrhosis.

Figure 3-2. Stratification of HCC patients to recurrence and non-recurrence guided by combined clinical factors. The combined use of cirrhosis and vascular invasion that segregates HCC patients with recurrent disease less than (A) 6 months, (B) 12 months, (C) 18 months, and (D) 35 months after surgical resection, results in 4 various recurrence risk groups, group 1 with the highest risk, group 4 with the lowest risk, and the remaining group 2 and 3 with risk ranging from 36% to 59%. R, recurrence; NR, non-recurrence; ■, Group 1 (HCC patients with vascular invasion and cirrhosis); ▤, Group 2 (HCC patients with vascular invasion but without cirrhosis); ▨, Group 3 (HCC patients without vascular invasion but with cirrhosis); □, Group 4 (HCC patients neither have vascular invasion nor cirrhosis); R, recurrence; NR, non-recurrence (154).

3.1.5 Prediction of recurrence is insufficient using gene signature derived from a high risk clinical factor analyzed at a time

Since vascular invasion and cirrhosis are strongly associated with the recurrence of HCC (Table 3-2), we postulated that differential gene list generated from each of this clinical parameter might contribute predictive value to disease recurrence. To explore predictive molecular signatures for recurrence, we randomly divided all 80 HCC samples into training and test sets according to their status of recurrence, vascular invasion, or cirrhosis, and analyzed independently. As a result, three top ranked independent gene sets (59, 48 and 40 probe sets) were derived separately by obtaining statistical significant comparison and satisfactory cross-validation of the training sets samples. The power of prediction for each of these gene set in their corresponding test set, was judged by its sensitivity and specificity. The former represents rate of recurrence cases correctly predicted and the latter is rate of non-recurrence cases correctly predicted.

The direct comparison of the parameter of interest, 20 recurrence cases versus 22 non-recurrence cases gave rise to 59 probe sets (57 genes). This gene set could estimate correctly 91% of the training sample (sensitivity=90%, specificity=91%). However, the accuracy of prediction of the test set (12 recurrence cases and 15 recurrence-free cases) was merely 74% (sensitivity=70%; specificity=77%). Secondly, 48 probe sets generated by comparing 18 invasion cases versus 33 non-invasion cases, could cross validate the training sets up to 90% overall accuracy (sensitivity=83%;

specificity=94%). Classification of an independent test set containing 16 recurrence and 13 non-recurrence cases using the same probe sets yielded 67% overall prediction accuracy (sensitivity=33%; specificity=83%). Finally, 39 probe sets generated from comparison between 25 cirrhotic and 20 non-cirrhotic patients, could estimate training set with 93% overall accuracy (sensitivity=96%; specificity=90%), but the prediction accuracy of test set (19 cases with recurrence and 16 cases without recurrence) decreased to 74% (sensitivity=86%; specificity=75%). In general, the leave-one-out cross-validation of each training sample set using gene set generated accordingly, all achieved 90% accuracy. However, the prediction power for each gene list reduced to approximate 70%, with either low sensitivity or specificity when independent test sets were analyzed (Table 3-3).

Table 3-3. Performance of various gene sets, derived by considering different clinicopathological factors in predicting recurrence of HCC (154)

Training set	*Filter criteria	Probe set	Accuracy of estimating training set (sensitivity, specificity)	Test set	Accuracy of predicting test set (sensitivity, specificity)
20 R, 22 NR	Top 60, F>1.25	59	SVM-90% (91%, 90%)	19R, 19NR	SVM-70% (67%, 70%)
18 VI, 33 NI	Top 200, F>2	48	SVM-90% (83%, 94%)	16R, 13NR	SVM-67% (33%, 83%)
25 Ci, 20 NCi	Top 50, F>1.25	40	SVM-93% (96%, 90%)	19R, 16NR	SVM-74% (89%, 59%)

VI, vascular invasion; NI, non-invasion; Ci, cirrhosis; Nci, non-cirrhosis; R, recurrence; NR, non-recurrence; F, median fold change; SVM, Support Vector machine.

3.1.6 Prediction of recurrence in group 2 and 3 using gene signature derived from the same group by combining high risk clinical factors

Patients grouping according to integrated vascular invasion and cirrhosis, was insufficient to accurately estimate recurrence especially in group 2 and 3, though extremely contrast outcomes were observed in group 1 and 4. Moreover, the prediction for recurrence for HCC samples using 3 gene sets generated by comparing one clinical factor at a time, including direct comparison between recurrence and non-recurrence cases, hardly reached 80% accuracy, suggesting the heterogeneity of the samples. As such, we attempted to employ molecular profiling of subgrouped patients stratified by these two factors (Figure 3-2) to determine if this could yield better or improved predictive results than by just considering on one factor at a time.

More than half of the patients (n=48; 60%) in this study were clustered in subgroup 2 and 3, yet the ratio of recurrence to recurrence-free cases were almost 50:50. In an attempt to uncover a molecular signature to predict recurrent disease for these HCC patients, these group of patients were randomly divided into training (n=23) and test set (n=25). Parametric (Student T-test) and non-parametric (Wilcoxon test) statistical tests of the Genedata Expressionist Analyst (version 1.0) were employed to select genes that could discriminate between recurrent disease (n=11) and non-recurrent disease (n=12) in the training set. The genes obtained were ranked by p values and the top 500 probe sets with the most significant p values were selected. To generate additional genes sets to cross-validate the training set, probe sets were further

filtered from this gene list by progressively selecting those with increased median fold changes from 1.25 to 5. Gene set that gave the lowest error rate for the cross-validation in the training group was eventually employed to estimate the recurrence outcome for an independent test group (n=14 for recurrent disease; and n=11 for non-recurrent disease) of HCC patients.

The 61 probe sets (57 genes), selected from the 130 top-ranked probe sets with median fold change greater than 1.5, could cross-validate the training set consisting of recurrence up to 35 months, with the best estimation score ranging from 96% to 100% using Support Vector Machine (SVM), Sparse Linear Discriminant Analyst (SLD) and K Nearest Neighbor (KNN) (K=3) (Table 3-4). Intriguingly, when this 57 gene signature was employed to predict a totally independent test set of 14 HCC patients with recurrence (followed up till 35 months) and 11 HCC patients without recurrence, the accuracy was between 76% (sensitivity = 78.6%; specificity = 73%) and 84% (sensitivity = 86%; specificity = 82%) using SLD, SVM and KNN classifier algorithms (Table 3-4). Similar patterns were observed irrespective disease recurrence occurred earlier within duration of 6, 12, and 18 months. This gene cluster of 61 probe sets (57 genes) identified is tentatively designated as the HCC recurrence prediction molecular gene signature.

Table 3-4. Performance of the 57-member gene signature in predicting disease recurrence for HCC patients in combined group 2 and 3 (154).

Duration of recurrence	No. of HCC patients in training set	Accuracy of estimating training set	No. of HCC patients in test set	Accuracy of predicting test set (sensitivity, specificity)
6 months	7R, 12NR	SVM-100% SLD-100% KNN-100%	8R, 11NR	SVM-79% (75%, 82%) SLD-74% (75%, 73%) KNN-84% (87.5%, 82%)
12 months	8R, 12NR	SVM-100% SLD-100% KNN-100%	13R, 11NR	SVM-83% (85%, 82%) SLD-75% (77%, 73%) KNN-83% (85%, 82%)
18 months	10R, 12NR	SVM-100% SLD-100% KNN-100%	13R, 11NR	SVM-83% (85%, 82%) SLD-75% (77%, 73%) KNN-83% (85%, 82%)
35 months	11R, 12 NR	SVM-100% SLD-100% KNN-96%	14R, 11NR	SVM-84% (86%, 82%) SLD-76% (78.6%, 73%) KNN-84% (86%, 82%)

VI, vascular invasion; NI, non-invasion; Ci, cirrhosis; Nci, non-cirrhosis; R, recurrence; NR, non-recurrence; F, median fold change; SVM, Support Vector machine; SLD, Sparse Linear Discriminant Analyst; KNN, K Nearest Neighbor; Subgroup 2, patients with invasion; but without cirrhosis; Subgroup 3, patients with cirrhosis but without invasion. The 57 genes are derived from the top 130 ranked genes resulting from analyses with fold greater than 1.5.

3.1.7 Kaplan Meier analysis of 57-gene set prediction reveals poorer prognosis in predicted recurrent than non-recurrent patients from both group 2 and 3

When Kaplan-Meier plot employing log-rank tests was performed to assess the predicted outcomes of the test set patients from these two groups by considering the duration of disease free after resection as time factor, the predicted recurrence (n=14)

and non-recurrence (n=11), generated with the 57 gene-set displayed statistical significant difference ($p=0.002$) (Figure 3-3). The predicted recurrent patients would always have significant poorer prognosis or lower tumor free rate than the predicted non-recurrent patients.

Figure 3-3. Kaplan-Meier analysis of the prediction outcomes of HCC recurrence. A log-rank test of the prediction outcome of the 57 gene-set in test set patients from both group 2 and 3, by considering the duration of disease free after resection as time factor, revealed statistical significant difference between patients predicted to have recurrence and no recurrence (154).

3.1.8 The 57-gene set is specifically applicable in predicting HCC recurrence of patients from both group 2 and 3

To investigate if the prediction performance of the 57-gene signature is confounded by any clinicopathological factors, multivariate analysis was performed

and the result showed that it was solely confounded by the outcome of the predicted recurrence but not by any clinical factors (Table 3-5).

Table 3-5. Multivariate analyses to determine if prediction performance of the 57-gene signature is confounded by any clinical factors.

Variable	<i>P</i> value ^a
Predicted recurrence	0.001
Sex (men/women)	0.439
Age (median) (years)	0.344
Viral infection (HBV / HCV / non-B non-C)	0.368
Capsulation (yes / no)	0.210
Tumor size (median) (cm)	0.429
AFP level (median) (ng/ml) [low / medium / high]	0.493
Lesion (single / multiple nodules)	0.267
Histological grading (G1/G2/G3/G4)	0.982
Cirrhosis (yes / no)	0.341
Invasion (yes / no)	0.341
HBV, hepatitis B virus; HCV, hepatitis C virus; non-B non-C, negative for both HBV and HCV antigen; low, less than 10 ng/ml; medium, 10 to 300 ng/ml; high, more than 300 ng/ml; G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated; G4, extremely poor differentiation; R, duration of recurrence.	
^a <i>P</i> value were calculated using multivariate cox regression analysis	

Furthermore, when the predictability of this 57 gene signature was investigated in groups 1 and 4, the prediction scores were mostly below 80% (Table 3-6). The result was not as good as the extreme distribution of HCC recurrence using combined vascular invasion and cirrhosis in each group 1 and group 4 which were readily predicted to be most likely to have recurrence and vice versa (Figure 3-2). This further supported that the 57-member gene set could effectively distinguish patients with and without recurrence, particularly at diagnosis, with vascular invasion but no cirrhosis (group 2) and those have cirrhosis but no invasion (group 3).

Table 3-6. Performance of the 57-member gene signature in predicting recurrence for HCC patients in group 1 and group 4 respectively.

No. of HCC patients	Accuracy of prediction (sensitivity, specificity)
10R, 2NR (group 1)	SVM-67% (60%, 100%) SLD-33% (30%, 50%) KNN-67% (60%, 100%)
4R, 16NR (group 4)	SVM-75% (50%, 81%) SLD-70% (50%, 75%) KNN-80% (50%, 87.5%)

R, recurrence; NR, non-recurrence; SVM, Support Vector machine; SLD, Sparse Linear Discriminant Analyst; KNN, K Nearest Neighbor; group 1, patients diagnosed with cirrhosis and vascular invasion; group 4, patients without cirrhosis or vascular invasion at diagnosis.

3.1.9 Performance comparison of 65-gene set predictor of group 1 and 4 versus the 57-gene set predictor of group 2 and 3

To investigate whether the prediction using molecular expression profile could reproduce the clinicopathological findings for disease recurrence in group 1 and 4 (Figure 3-2), we applied the same approach as in Figure 3-4A and identified an initial of 1181 genes that discriminated ($P < 0.05$; $F > 1.25$) between recurrence of group 1 ($n=10$) and non-recurrence of group 4 ($n=12$). Subsequently, a 65-gene set with cutoff value $P < 0.001$ and median fold change > 1.25 (Figure 3-4B) after the multi-steps filtering, could best classify the training set with consistent accuracy up to 86% (sensitivity=80% and specificity=92%; Table 3-7).

Figure 3-4. Multi-step filtering approach generates 2 optimized gene sets to predict 2 combination groups of HCC patients. The use of parametric (Student T-test) and non-parametric (Wilcoxon test) tests, p value ranking and median fold change filters generate (A) the 57-gene set that specifically predict recurrence in group 2 (Gp2) and 3 (Gp3), and (B) the 65-gene set that predict recurrence outcome of group 1 (Gp1) and 4 (Gp4). R, recurrence; NR, non-recurrence; Gp1, patients with invasion and cirrhosis; Gp2, patients with invasion, but without cirrhosis; Gp3, patients with cirrhosis but without invasion; Gp4, patients negative for both invasion and cirrhosis; F, median fold change.

The same gene set could predict an independent test set with 70% to 80% overall accuracy (sensitivity=33% to 67%; specificity=86%) (Table 3-7). The low sensitivity of the predictive value could be due to the small sample size of the recurrence cases (n=3). In general, the result indicated that the use of expression profile for estimation of recurrence in group 1 and 4 was not as good as the stratification findings based on the combined clinical factors approach (Figure 3-2), as well as the prediction result of the 57-gene set in group 2 and 3 (Table 3-7). Nonetheless, the gene list would help to elucidate the molecular features reflecting mechanism involved in the progression of HCC to advanced stage especially in group 1 that was likely to recur after surgical resection.

Table 3-7. Performance comparison of 2 gene set predictors in their corresponding training and test samples using different classifiers.

Training set	Filter criteria	Probe set (gene)	Accuracy of estimating training set	Test set	Accuracy of predicting test set
11R, 12 NR (Gp2 and Gp3)	Top 130, F>1.5	61 (57)	SVM- 100% SLD- 100% KNN-96%(100%,92%)	14R, 11NR	SVM-84% (86%, 82%) SLD-76% (79%, 73%) KNN-84% (86%, 82%)
10 R(Gp1), 12 NR (Gp4)	Top 70, F>1.25	70 (65)	SVM-86%(80%, 92%) SLD-86%(80%, 92%) KNN-86%(80%, 92%)	3R,7NR	SVM-80%(67%, 86%) SLD-80%(67%, 86%) KNN-70%(33%,86%)

R, recurrence; NR, non-recurrence; F, median fold change; Gp1, patients with invasion and cirrhosis; Gp2, patients with invasion; but without cirrhosis; Gp3, patients with cirrhosis but without invasion; Gp4, patients negative for both invasion and cirrhosis; SVM, Support Vector machine; SLD, Sparse Linear Discriminant Analyst; KNN, K Nearest Neighbour.

3.1.10 Comparison analysis of two HCC recurrence-related gene sets

Having identified two differential gene set predictors that could sufficiently discriminate recurrence in two combined groups of patients, we next attempted to gain additional insight of the biological differences of these two combined groups by investigating the initial 2 larger gene sets from which the 2 predictor sets evolved. Gene list A contained 518 genes, from which 57 gene signature was derived, that differentiate between recurrence and recurrence-free cases from group 2 and 3. Gene list B represented 1181 genes that eventually gave rise to 65 gene member which differentially expressed between recurrent cases in group 1 and non-recurrent samples in group 4 (Figures 3-4). There was small number of genes overlapped between 2 gene sets. Of the 35 genes common in gene set A and gene set B (Figure 3-5), 27 genes shared the same expression patterns.

Figure 3-5. Venn diagram showing comparison of the two gene sets. Gene list A represents 284 up-regulated genes and 234 down-regulated genes in recurrence cases compared to non-recurrence cases, derived from group 2 and group 3. Gene list B represents 587 up-regulated genes and 594 down-regulated genes in recurrent HCC of group 1 compared to non-recurrent cases of group 4.

Fourteen out of 27 genes had overexpression in both gene list A and gene list B (Table 3-8) while 13 genes were underexpressed concurrently in the two gene lists (Table 3-9). The remaining 8 genes had inversed expression pattern between the 2 gene lists (Table 3-10).

Table 3-8. Fourteen commonly overexpressed genes in both gene list A and B

Gene Title	Symbol	Function
apoptosis antagonizing transcription factor	AATF	Apoptosis/anti-apoptosis
H2A histone family, member X	H2AFX	c'some organization
SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1	SMARCE1	c'some organization
chromobox homolog 1 (HP1 beta homolog Drosophila)	CBX1	c'some organization
interleukin 8	IL8	G-protein signaling
adenylosuccinate lyase	ADSL	metabolism(nucleotide)
238829_at	---	metabolism(nucleotide)
testis-specific kinase 1	TESK1	protein processing
DnaJ (Hsp40) homolog, subfamily C, member 10	DNAJC10	Transport
ubiquitin-conjugating enzyme E2E 1 (UBC4/5 homolog, yeast)	UBE2E1	Ubiquitination
acidic (leucine-rich) nuclear phosphoprotein 32 family, member B	ANP32B	NA
KIAA0102 gene product	KIAA0102	NA
flotillin 1	FLOT1	NA
OTU domain, ubiquitin aldehyde binding 1	OTUB1	NA

c'some, chromosome; NA, not available

Table 3-9. Thirteen commonly underexpressed genes in both gene list A and B

Gene Title	Symbol	Function
acid phosphatase, testicular	ACPT	acid phosphatase activity
insulin induced gene 1	INSIG1	Cell proliferation
hypothetical protein FLJ10986	FLJ10986	CHO Metabolism
claudin 14	CLDN14	cytoskeleton
solute carrier organic anion transporter family, member 2B1	SLCO2B1	ion transport
solute carrier family 27 (fatty acid transporter), member 2	SLC27A2	Metabolism/biogenesis
chromosome 5 open reading frame 4	C5orf4	Metabolism/biogenesis
hypothetical protein MGC15875	MGC15875	pyridoxal phosphate binding /// transaminase activity
solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 25	SLC25A25	Transport
F-box and leucine-rich repeat protein 17	FBXL17	Ubiquitination
tigger transposable element derived 2	TIGD2	NA
LOC440283	---	NA
adult retina protein	LOC153222	NA

NA, not available

Table 3-10. Eight genes with inversed expression between gene list A and B

Gene Title	Symbol	Function	Gene list A	Gene list B
cyclin-dependent kinase 7 (MO15 homolog, <i>Xenopus laevis</i> , cdk-activating kinase)	CDK7	DNA repair	↓	↑
UDP-glucose ceramide glucosyltransferase	UGCG	Development	↓	↑
occludin	OCLN	protein processing	↓	↑
TGF beta-inducible nuclear protein 1	TINP1	Metabolism(a acid/protein)	↓	↑
ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)	ENPP4	Metabolism(nucleotide)	↓	↑
CDW92 antigen	CDW92	Transport	↓	↑
similar to RIKEN cDNA 2310030G06 gene	MGC14839	NA	↑	↓
228030_at	---	NA	↑	↓

↑, overexpression; ↓, underexpression

3.1.11 Biological and functional comparison analysis of two HCC recurrence-related gene sets

Gene list A had exclusively 483 genes (268 genes with overexpression and 215 with underexpression), independent from 1146 genes in B comprising exclusively of 567 overexpressed and 579 underexpressed genes (Figure 3-5). To gain insight of the biological functions of these genes, we clustered them based on the functional canonical pathway annotation derived from Ingenuity Pathway analysis (IPA) as shown in Figure 3-6. The $-\log(p\text{-value}) > 1.3$, which is equivalent to $p < 0.05$, was used as a cutoff threshold point depicting the significant association of our datasets with the given canonical pathways curated in IPA. Figure 3-6A represented 30 most significantly disrupted canonical pathways ($-\log(p\text{ value})$ between 1.39 and 2.61) of 483 altered genes (purple red bar) exclusively in recurrent cases of the co-existing group 2 and 3. Figure 3-6B demonstrated the top 30 most aberrant canonical pathways ($-\log(p\text{ value})$ between 2.76 and 17.1) conferred by 1146 genes exclusively disrupted

in recurrent cases of group 1 versus non-recurrent cases in group 4. Although the majority of the disrupted pathways were mutually exclusive in gene set A and B, pathways in charge of ATM signaling and xenobiotic metabolism were both significantly affected in the recurrence cases of these two comparison settings, as indicated by red arrows in Figure 3-6. Intriguingly, majority of the members in gene set B, were underexpressed in recurrent groups and were significantly associated with pathways governing various metabolism process for instance fatty acid, bile acid and amino acid metabolisms while these were not observed in gene set A. This might serve as the unique underlying mechanisms that confer to the enhanced aggressiveness of group 1 patients, in which majority are cirrhotic patients with invasion. Meanwhile, NRF2-mediated oxidative stress response, LPS-stimulated MAPK signaling, HGF and Cdc42 signaling were exclusively activated in recurrence patients from combined group 2 and 3.

Figure 3-6. Functional classification of 2 sets of HCC recurrence-related gene signatures. Pathway annotation using Ingenuity Pathway Analysis results in (A) gene list A comprising of 483 differential expressed genes between recurrence and non-recurrence cases exclusively derived from group 2 and group 3, and (B) gene list B comprising of 1146 genes exclusively and differentially expressed between recurrent samples of group 1 and non-recurrent samples of group 4. Arrows depict common pathways in both gene list A and B.

3.1.12 Hierarchical clustering and functional annotation of two HCC recurrence gene set predictors

The hierarchical clustering analysis of 57 genes revealed four distinct groups according to their expression when compared to surrounding non-tumorous tissues of the same HCC patient (ST) and normal liver tissue of liver metastases from colon cancer patients (NN) (Figure 3-7). Twenty two genes showed elevated expression in recurrence groups compared to recurrence-free cases of subgroups 2 and 3 as well as surrounding non-tumorous tissue and normal normal liver (Figure 3-7A). In contrast, 16 genes had decreased expression in recurrence as compared to the non-recurrent and normal controls groups (Figure 3-7C). Figure 3-7B and 3-7D represented 7 and 12 genes that were specifically underexpressed and overexpressed in non-recurrence patients diagnosed with vascular invasion and cirrhosis, from group 2 and 3. The several top significantly biological functions associated significantly with the 57 gene signature including organ development, reproductive system function, auditory and vestibular system, injury and abnormalities, lipid metabolism, cell cycle, connective tissue development, gene expression and cancer were summarized in Table 3-11.

Figure 3-7. Gene expression pattern of 57 gene set. Hierarchical clustering of 57 gene predictor with the best recurrence predictive value between recurrence and non-recurrence cases of group 2 and group 3 results in (A) 22 genes with higher expression in recurrence as compared to non-recurrence, surrounding non-tumorous tissues and normal normal liver, (B) 7 genes with distinct lower expression in non-recurrence as compared to the other 3 groups, (C) 16 down-regulated genes in recurrence groups as compared to non-recurrence, surrounding non-tumorous tissue and normal normal liver, and (D) 12 genes with distinct higher expression in non-recurrence groups than the other 3 groups. Gp1, patients with invasion and cirrhosis; Gp2, patients with invasion; ST, surrounding tissue; NN, normal normal liver (from colon metastases). red color represents up-regulation; black color represents no change; green represents down-regulation.

Table 3-11. The top 30 significant functional annotation of 57 genes.

		-log (p value)		Expression	Molecules
		max	min		
Organ Development	4.66	4.66	1.32	Up Down	USH1C ID2,INSIG1,ETS2,MAFB,SKP2
Reproductive System Development	4.66	4.66	1.32	Up Down	- ID2,ETS2,HEXB,SKP2
Auditory and Vestibular System	3.30	3.30	2.46	Up Down	USH1C INSIG1,HEXB,MAFB
Organ Morphology	3.30	3.30	1.79	Up Down	PTPN11,USH1C ID2,INSIG1,MAFB,SKP2
Organismal Injury and Abnormalities	3.01	3.01	1.79	Up Down	SMURF2,DNAJC10 MAFB,SKP2
Renal and Urological Disease	3.01	3.01	1.39	Up Down	SMURF2 MAFB,SKP2
Lipid Metabolism	2.83	2.83	1.39	Up Down	PTPN11,PARD3 CYP17A1,SLC27A2,INSIG1,HEXB
Cell Cycle	2.75	2.75	1.32	Up Down	CETN2,SKA2(FAM33A),PTPN11,RACGAP1,MPHOSPH9,PARD3, USH1C ID2,PGGT1B,SKP2
Connective Tissue Development	2.75	2.75	1.57	Up Down	PTPN11 ID2,CDC42SE1,SKP2
Gene Expression	2.61	2.61	1.32	Up Down	SMURF2,PARD3 ID2,ETS2,MAFB
Cancer	2.56	2.56	1.33	Up Down	PTPN11,RACGAP1 CYP17A1,GCH1,ID2,SLC27A2,INSIG1,PGGT1B,MAFB,SKP2
Cell Morphology	2.56	2.56	1.57	Up Down	PTPN11,RACGAP1,PARD3 HEXB,ETS2,CDC42SE1,SKP2
Cellular Development	2.56	2.56	1.30	Up Down	PTPN11,RACGAP1 ID2,ETS2,CDC42SE1,MAFB,SKP2
Cellular Growth and Proliferation	2.56	2.56	1.34	Up Down	PTPN11,SH3GLB2 (includes EG:56904) ID2,ETS2,MAFB,SKP2
Cellular Movement	2.56	2.56	1.79	Up Down	RACGAP1,SH3GLB2 (includes EG:56904) ID2,PTP4A1,ETS2,MAFB,SKP2
DNA Replication, Recombination, and Repair	2.56	2.56	1.79	Up Down	CETN2 NUDT9
Digestive System Development and Function	2.56	2.56	1.39	Up Down	CUL4B,PTPN11 ID2,INSIG1,ETS2,MAFB,SKP2
Embryonic Development	2.56	2.56	2.56	Up Down	- ID2
Endocrine System Development and Function	2.56	2.56	1.96	Up Down	PTPN11 ETS2,MAFB
Genetic Disorder	2.56	2.56	1.57	Up Down	- CYP17A1
Hematological System Development	2.56	2.56	1.38	Up Down	PTPN11,SMURF2,USH1C CYP17A1,GCH1,INSIG1,PGGT1B,HEXB,MAFB,CNGA1, SAR1B(SARA2)
Hematopoiesis	2.56	2.56	1.34	Up Down	PTPN11,RACGAP1 ID2,ETS2,MAFB
Immunological Disease	2.56	2.56	1.57	Up Down	PTPN11,RACGAP1 ID2,MAFB
Nervous System Development and Function	2.56	2.56	1.72	Up Down	PTPN11 SKP2
Neurological Disease	2.56	2.56	1.48	Up Down	PTPN11,GSTM3 (includes EG:2947),PARD3,USH1C MAFB
Nucleic Acid Metabolism	2.56	2.56	1.46	Up Down	PARD3,USH1C GCH1,MARVELD2,HEXB
Nutritional Disease	2.56	2.56	1.96	Up Down	- NUDT9
Small Molecule Biochemistry	2.56	2.56	1.66	Up Down	PTPN11 GCH1
Tissue Development	2.56	2.56	1.39	Up Down	PTPN11,PARD3 CYP17A1,GCH1,SLC27A2,INSIG1,NUDT9,PGGT1B,HEXB
			1.34	Up Down	PTPN11 ID2,ETS2

The $-\log(p \text{ value})$ were calculated using Fisher's exact test in all functional categories depicted from Ingenuity Pathway analysis and ranked.

The hierarchical clustering analysis of 65 differential expressed genes with the best predictive value for recurrence between group 1 and group 4 reflected the asymmetric expression pattern (Figure 3-8). The genes with underexpression in recurrence were greater in number than the overexpressed genes. Twenty two genes had the most elevated expression in group 1 followed by group 4, and their expression were homogenously lower in both surrounding tissue and normal normal liver (Figure 3-8A). Forty three genes showed decreased expression in group 1, the poorest prognostic group in comparison to group 4, surrounding non-tumorous tissue and normal normal liver from colon metastases (Figure 3-8B). A large number of these deregulated genes were involved in cellular metabolism activities for instance lipid, vitamin and mineral, amino acid, and drug metabolism (Table 3-12). Several typical liver specific genes including cytochrome P450, alcohol dehydrogenase and prealbumin were also found in this gene list.

Figure 3-8. Gene expression pattern of 65 gene set. Hierarchical clustering of 65 gene predictor with the best recurrence predictive value between group 1 and group 4 results in (A) 43 down-regulated genes in group 1 as compared to group 4, normal surrounding tissue and normal normal liver, and (B) 22 up-regulated genes in group 1 as compared to group 4, normal surrounding tissue and normal normal liver. Gp1=patients with invasion and cirrhosis; Gp2=patients with invasion; ST=surrounding tissue; NN=normal normal liver (from colon metastases); red color represents up-regulation; black color represents no change; green represents down-regulation.

Table 3-12. The top 30 significant functional annotation of 65 genes

	Functional annotation	-log (p value)		Expression	Molecules
		max	min		
Lipid Metabolism	Lipid Metabolism	4.58	1.38	Up Down	APEX1, CD24, P4KA (LOC220686 // PK4CA) CYP3A43, AR, CYP2C9, RORA (LOC283666), NR12, GLYAT, DHRS4, ADH4, CYP8B1
Small Molecule Biochemistry	Small Molecule Biochemistry	4.58	1.38	Up Down	RRM1 (includes EG:6240), ADSL, CD24, APEX1, P4KA (LOC220686 // PK4CA) DPYS, CYP2C9, GLYAT, MTHFD1, DCXR, ADH1 (MTCBP-1), CYP8B1, GSTZ1 (includes EG:2954), CYP3A43, AR, NR12, RORA, ALAS1, ALDH6A1, ADH4, DHRS4
Vitamin and Mineral Metabolism	Vitamin and Mineral Metabolism	4.58	1.61	Up Down	CD24 CYP2C9, NR12, MTHFD1, ADH4, DHRS4
Endocrine System Development	Endocrine System Development	3.90	1.53	Up Down	- CYP3A43, AR, CYP2C9, NR12
Genetic Disorder	Genetic Disorder	3.66	1.38	Up Down	MCM6, ADSL, KIF2A (KIF2), AP3B1, CD24, P4KA (LOC220686 // PK4CA) CYP2C9, MTHFD1, CYP8B1, AR, SMARCA2, NR12, D2HGDH (MGC25181)
Metabolic Disease	Metabolic Disease	3.66	1.71	Up Down	MCM6, ADSL AR, CYP2C9, NR12, D2HGDH (MGC25181), CYP8B1
Neurological Disease	Neurological Disease	3.66	1.35	Up Down	KIF2A, CD24, P4KA AR, RORA, SMARCA2, NR12, MTHFD1, CYP8B1
Amino Acid Metabolism	Amino Acid Metabolism	3.37	1.40	Up Down	CD24 GSTZ1 (includes EG:2954), NR12, MTHFD1, ALDH6A1, ADH1
Reproductive System Disease	Reproductive System Disease	3.33	1.38	Up Down	GINS1 (KIAA0186), CD24, P4KA, RRM1 (includes EG:6240) AR, NR12, TGFBR3
Cell Cycle	Cell Cycle	3.03	1.33	Up Down	SFRS2, KIF2A (KIF2), RAN, SOCS5 AR, SMARCA2
Cell Death	Cell Death	3.03	1.35	Up Down	APEX1, CD24 AR, NR12
Cellular Assembly and Organization	Cellular Assembly and Organization	3.03	1.38	Up Down	KIF2A, RAN, AP3B1, TRAPPC4, CD24 AR, SMARCA2
DNA Replication, Recombination, and Repair	DNA Replication, Recombination, and Repair	3.03	1.38	Up Down	KIF2A, RAN, APEX1, RRM1 (includes EG:6240) AR
Drug Metabolism	Drug Metabolism	2.67	1.38	Up Down	- CYP3A43, AR, CYP2C9, NR12, MTHFD1, ADH4
Cancer	Cancer	2.48	1.30	Up Down	MCM6, RAN, RRM1 (includes EG:6240), ADSL, RPS16, APEX1, SOCS5, CD24, P4KA MASP2, LECT2, TGFBR3, GINS1 (KIAA0186), AR, NR12, SMARCA2, RORA (LOC283666), MUTHYH (CYP2C19 // CYP2C9), ALAS1
Carbohydrate Metabolism	Carbohydrate Metabolism	2.48	1.38	Up Down	P4KA NR12, DCXR
Cardiovascular System Development	Cardiovascular System Development	2.48	1.35	Up Down	SFRS2 AR, CYP2C9, TGFBR3
Cell-To-Cell Signaling and Interaction	Cell-To-Cell Signaling and Interaction	2.48	1.41	Up Down	SFRS2, CD58, CD24 AR, NR12
Cellular Development	Cellular Development	2.48	1.36	Up Down	SFRS2, AP3B1, SOCS5, CD24 COX17, AR, RORA (LOC283666), TGFBR3
Cellular Function and Maintenance	Cellular Function and Maintenance	2.48	1.38	Up Down	KIF2A, AP3B1, CD24, SOCS5 AR, NR12, SMARCA2, RORA (LOC283666)
Cellular Movement	Cellular Movement	2.48	1.35	Up Down	CD58, CD24 AR, TGFBR3, ADH1
Developmental Disorder	Developmental Disorder	2.48	1.71	Up Down	- AR
Gene Expression	Gene Expression	2.48	1.32	Up Down	APEX1 HRSP12, AR, NUDT6, NR12, RORA (LOC283666), SMARCA2, TGFBR3
Hair and Skin Development and Function	Hair and Skin Development and Function	2.48	1.64	Up Down	- AR, RORA (LOC283666), TGFBR3
Hematological System Development	Hematological System Development	2.48	1.45	Up Down	AP3B1, CD58, SOCS5, CD24 RORA (LOC283666)
Hepatic System Disease	Hepatic System Disease	2.48	1.34	Up Down	- GSTZ1 (includes EG:2954), NR12
Infection Mechanism	Infection Mechanism	2.48	1.35	Up Down	SFRS2 AR, SMARCA2
Inflammatory Response	Inflammatory Response	2.48	1.49	Up Down	CD58 NR12
Molecular Transport	Molecular Transport	2.48	1.39	Up Down	RAN, APEX1, P4KA, RRM1 (includes EG:6240) COX17, SFXN5, AR, RORA (LOC283666), NR12, DCXR, ADH4, CYP8B1

The $-\log(p \text{ value})$ were calculated using Fisher's exact test in all functional categories depicted from Ingenuity Pathway analysis and ranked.

3.1.13 Verification of genes selected from two gene set predictors by quantitative real-time PCR

To validate the differential expression pattern of the selected genes with the highest median fold change in 63-gene set found by microarray analysis, quantitative real-time PCR was conducted. Due to the constraint of the independent new samples, half of the samples used for verification in real-time PCR were the same as those used in microarray analysis. All 7 genes selected from 57-gene set, comprising of 5 up-regulated (RACGAP1, KCNK1, SMURF2, USH1C and GSTM3) and 2 down-regulated genes (CNGA1 and INSIG1) in recurrence cases of group 2 and 3, showed consistent result between the two techniques used (Figure 3-9). The significant altered expression of these genes were observed right from the comparison between non-malignant tissue (normal normal liver or surrounding tissue) and non-recurrence HCC as well as between non-recurrence and recurrence HCC.

Figure 3-9. Validation of 7 genes selected from 57 gene set predictor. Comparison of the expression of 5 representative up-regulated genes and 2 representative down-regulated genes selected from the 57-member HCC recurrence-associated gene set studied by Affymetrix microarray (upper panel) and quantitative real-time PCR (lower panel). NN, normal liver tissues from patients with colon cancer metastases to liver; ST, matched surrounding non-tumorous tissue; NR, non-recurrence; R, HCC recurrence; columns, mean; bars, standard error; *, $P < 0.05$.

Figure 3-10 showed the validation of 6 genes selected from 65-gene set by quantitative real-time PCR. The results corresponded well with the microarray analysis for all 6 genes including 2 up-regulated (GINS1 or previously known as KIAA0186 and MCM6) and 4 down-regulated genes (LECT2, DXCR, DPYS and PROZ) as the differences between Group 1 and Group 4 were significant. However, the significant differences occurred at earlier stage between surrounding nonmalignant tissue or normal normal liver and Group 4 of HCC as indicated by microarray analysis were not observed in real-time PCR analysis for GINS1, MCM6 and DXCR genes.

3.2 Discussion

3.2.1 Cross assessment of prediction ability of public HCC recurrence-related gene expression data using clinical samples in this study

Gene expression profiling in hepatocellular carcinoma and pre-malignant liver is an expanding area of interest. Various powerful molecular technologies such as oligonucleotide arrays, cDNA microarrays, PCR-based arrays and recently proteomic approaches have been employed aiming to translate the molecular information into useful clinical practice in cancer research (189, 190). Over the past 20 years, compelling studies of expression profiling in HCC have identified lists of genes that can differentiate HCC tumor from their adjacent normal tissue (191, 192), discriminate patients according to etiological factors (193, 194) and disease state (195, 196). Lately molecular signatures associated with pre-neoplastic lesions including cirrhosis are also uncovered (192, 194).

Although great advancement has been made in the surgical treatment of HCC, recurrence after surgery as a result of intrahepatic metastasis and de novo primary tumor remains a key challenge and complicates 70% of the cases at 5 years (33, 71, 197). Up-to-date several studies have identified gene signatures to predict survival, intrahepatic metastasis and early recurrence (198-201) of HCC with overall accuracy range from 73% to 93%. To evaluate these genes in our dataset on recurrence prediction, we transform the data into our platform. There were 621, 231, 29 and 37

probe sets representing 406, 153, 12 and 20 genes found in our Affymetrix gene expression data from studies of Lee *et al.* (198), Ye *et al.* (199), Iizuka *et al.* (200) and Kurokawa *et al.* (201). We employed them to predict our clinical by selecting samples that match the criteria including early recurrence less than one year employed by Iizuka *et al.* and recurrence less than 2 years specified by Kurokawa *et al.* The overall accuracy scores can merely achieve 65% for Lee *et al.*'s gene set, 57% for Ye *et al.*'s gene set, 61% for Iizuka *et al.*'s gene set and 48% for Kurokawa *et al.*'s gene set. This discrepancy might arise from the differences in algorithms or computational approaches, microarray technologies used (oligonucleotide versus cDNA and PCR-based arrays), patients' backgrounds (HBV versus HCV associated HCC and racial or geographic origins).

3.2.2 Comparison of public HCC recurrence-related gene expression data with the prediction signatures in this study

Insulin-induced gene 1 (INSIG) was the only common gene detected between our 57-members gene signature and that reported by Lee *et al.*(198). The top most up-regulated gene in recurrent HCC samples, USH1C encodes the gene product harmonin, a PDZ-containing protein. PDZ domains are modular protein interaction domains that play a role in protein targeting and the assembly of large protein complexes involved in signaling or subcellular transport and was the first mutation in a PDZ-encoding gene linked to a human disease (202, 203) Another gene, which shows remarkable up-regulation in the recurrent HCC samples, is Rac GTPase

activating protein 1 (RACGAP1). Rho GTPases, which include Rho, Rac and CDC42, play pivotal roles in regulating the actin cytoskeleton necessary for cell motility, cell-cell contact and malignant transformation in many cell types (204), are responsible for the regulation of many downstream kinases such as CDC42. Interestingly, CDC42 small effector 1, which inhibits GTPase activity, was also found to be down-regulated in the recurrent HCC samples studied. Moreover, casein kinase that regulates the cytoskeletal organization through small GTPases via the Wnt signaling pathway (205) is also down-regulated in the recurrent HCC samples. Expression of the detoxification enzyme cytochrome P450 and enzymes involved in metabolism (AS3MT and HEXB) that are predominantly expressed in differentiated hepatocytes were down-regulated in recurrent HCC samples. The suppression of these genes could reflect tumor de-differentiation following the progression of malignancy. On the other hand, the gene that is noticeably up-regulated in recurrent HCC samples is La ribonucleoprotein domain family, member 6 (LARP6), a RNA-binding protein..

When the 2 cluster of genes (518 and 1181 gene sets) of our dataset were compared with 406 survival-related genes of Lee *et al.*, 7 genes (3 up-regulated and 4 down-regulated) and 124 genes (46 up-regulated and 78 down-regulated) were overlapped with their list. Among the 124 genes, most of them are annotated as metabolism-related genes with large number involved in the downregulation of lipid and glucose metabolism. We also compared our two gene sets with 153 intrahepatic metastasis-related genes of Ye *et al.* There were 3 up-regulated and 3 down-regulated

genes were overlapped with 518 gene list while 5 down-regulated genes represented in our 1181-gene list. To compare with intrahepatic recurrence of HCV-associated HCC databases from Iizuka *et al.*, one up-regulated gene from cluster A list (1181-gene set) and two down-regulated genes from cluster B list (518 gene sets) were overlapped. The first two intersects with Lee *et al.* and Ye *et al.* represent the consistent gene signature for recurrence of HBV-associated HCC. The third intersect with Iizuka *et al.* represents common biomarkers for progression of HCC into recurrent stage regardless of HBV- or HCV- associated factors.

3.2.3 Differential gene expression to predict recurrence of various risk groups

The clinical usefulness of the signatures derived merely from recurrence-based profiling are limited and what is urgently needed is the ability to accurately identify, at diagnosis, patients at different risk for recurrence after curative hepatectomy for primary HCC. The ability to predict the clinical outcome after surgery would enable the implementation of optimal clinical treatments and follow-up strategies prior to clinical manifestation of recurrent disease. Therefore, a prediction system for recurrence customized according to our dataset is necessary. Univariate analysis revealed that vascular invasion and cirrhosis associated most closely with recurrence. This correlates with the previous findings showing that vascular invasion (117) and the remaining cirrhotic liver after complete resection of HCC tumor (109) respectively serve as risk factor for intrahepatic recurrence. HCC patients diagnosed with these two

lesions usually have poorer outcome as our result indicated that the majority of them (83%) have recurrence. This is true even in non-neoplastic condition as cirrhotic patients with portal vein thrombosis were prone to develop advanced liver disease and HCC as a result of the mutation in prothrombin gene (96).

Using supervised learning analysis as summarized in Figure 3-11, the present study identified 2 clusters of genes. The first cluster (gene set A) contains 518 genes that differentiate recurrence and recurrence-free patients of group 2 and group 3 diagnosed with either cirrhosis or invasion. The second cluster (gene set B) contains 1181 genes that discriminate significantly cirrhotic patients with invasion (group 1) from patients negative for both of these lesions (group 4). Gene set A from which a gene set predictor, the 57 gene member is evolved, can effectively predict recurrence with 84% overall accuracy rate after curative resection in HCC patients diagnosed either with cirrhosis or vascular invasion. Gene set B from which the 65 gene member is derived, serves as specific molecular signature in patients with cirrhosis and vascular invasion, the two clinical factors proven to be markedly associated with disease recurrence. The use of 65 gene set to predict recurrence of an independent test set from group 1 and group 4 patients achieves overall 70% to 80% accuracy. Collectively, these two non-overlapping exclusive gene set predictors yield higher prediction accuracy than differential genes derived from direct comparison between all recurrence and non-recurrence samples which can only estimate up to 74% accuracy.

Figure 3-11. Overview of the findings of two differential gene set predictors. Two gene set members, 57 and 65 get set predictors were identified employing supervised learning analysis and multistep filtering approaches. R, recurrence; NR, non-recurrence; VI+, patients diagnosed with vascular invasion; VI-, negative in invasion; Cirr+, positive in cirrhosis; Cirr-, no cirrhosis.

3.2.4 Comparative functional annotation of different gene sets

The unique existence of a large number of aberrantly expressed genes involved in metabolism is observed in gene cluster B. We found that huge number of genes involved in various metabolism such as fatty acid, amino acid, bile acid, xenobiotics and glucose metabolism were largely down-regulated within the differential 1146-

gene set of cluster A in recurrence of group 1. This suggests the complete deregulated function or incapability of hepatocytes to regulate its normal metabolic activity, bile acid biosynthesis and detoxication process. Furthermore, the up-regulated genes involved in carbohydrate metabolism are mostly associated with gluconeogenesis while only a few down-regulated genes are involved in glycolysis. From this observation, it is possible to infer that aggressive tumor microenvironment favors the glucose synthesis. This is discordant with a report suggesting the upregulation of glycolysis is consistently correlated with poor prognosis and increased tumor aggressiveness from pre-malignant lesions to invasive cancer (206).

The presence of large number of down-regulated genes in fatty acid metabolism may result in accumulation of lipid in the liver, which in turn induces cytokine production such as IL-6 and TNF (59), generation of oxygen radical due to fatty acid oxidation, hepatocytes damage and apoptosis, and ultimately fibrosis, a molecular mechanism similar to obese or diabetic (60) patients. High incidence of HCC has been associated significantly with obesity and diabetes (58). This pattern of functional group distribution does not occur in the 483-gene set that differentiate recurrence from non-recurrence samples of Group 2 and Group 3, but rather NRF 2-mediated oxidative stress response, LPS-stimulated MAPK signaling, HGF and Cdc42 signaling are predominant. There seems to be a unique mechanism for instance the predominant of NRF2-oxidative stress response exploited by the recurrence samples of Group 2 and Group 3 to maintain their survival. This pathway has been recently

implicated in playing a potential role in HCV pathogenesis by linking between genes involved in HCV-induced perturbation of lipid metabolism and oxidative stress (207).

Another distinct patterns including genes involved in coagulation, proteolysis and peptidolysis were exclusively observed in gene set B but not in gene set A. The majority of the genes in almost all these groups are down-regulated. These probably serve as additional oncogenic strategies, unique to the most aggressive group, the cirrhotic patients with invasion in order to favor tumor progression towards the development of intrahepatic recurrence after surgery. SERPINC1, PROS1, HGFAC and CPB2 are involved in the coagulation and proteolysis pathway. The downregulation of these genes has been linked to thromboembolism and cancer development (208). PLG, a proteolysis regulator, whose downregulation in HCC has been reported to play a critical role in cancer invasion and metastasis (209).

As the pre-malignant liver tissues, primarily the inflamed or cirrhotic liver from patients without HCC are unavailable, we assumed that all the differential genes selected are related to the alteration in a more advanced stage of HCC based on the sample labels, recurrence and non-recurrence cases used in this study. To evaluate the relationship of the gene sets selected in this study with transition from pre-malignant to early stage of HCC, we compare again the two gene sets with 273 HCC-associated signature of Kim *et al* (192). The 273 genes were selected by Kim and colleagues by comparison between the combined HCC samples with high risk cirrhotic patients and low risk cirrhotic patients. Cirrhotic patients with end-stage chronic liver disease

(CLD) but without HCC and had been associated with HBV infection, HCV infection, Wilson's disease and hemochromatosis are clustered as high risk group as they are more likely to develop HCC. While the low risk group consisted of cirrhotic patients who had CLD but not HCC and had been associated with alcoholic liver disease, primary biliary cirrhosis and autoimmune hepatitis. The 6 up-regulated genes and 45 down-regulated genes overlapped with cluster A gene set, and 3 up-regulated and 9 down-regulated genes overlapped with cluster B gene set, may serve as the useful biomarkers for early detection of HCC as well as those that are likely to have recurrence.

CHAPTER 4

Rac GTPase activating protein 1 (RACGAP1) up-regulation as a molecular marker with prognostic indication for recurrence of human hepatocellular carcinoma

4.1 Results

4.1.1 Significant overexpression of RACGAP1 in HCC recurrence within 2 years

Using microarray approach, RACGAP1 was not only significantly upregulated in primary HCC with recurrence within 2 years (R) compared to non-recurrence (NR) (Fold change = 2.31, $p < 0.0001$, FDR (false discovery rate) = 0.043%), but also statistically higher in non-recurrent biopsy (n=41) than either in the matched surrounding non-tumorous tissues (ST) (n=27) (Fold change = 2.5, $p < 0.0001$; Figure 4-1A). Consistent with microarray results, quantitative real-time PCR, demonstrated that HCC recurrence samples had pronounce higher fold expression than non-recurrence (Fold change = 2.03, $p < 0.02$), as well as significantly higher expression in non-recurrent HCC compared to the matched surrounding non-tumorous tissues (ST) (Fold change = 2.12, $p < 0.025$; Figure 4-1B). As such, these suggest the role of RACGAP1 in HCC progression from early cancer development to aggressive stage.

Figure 4-1. Significant up-regulation of RACGAP1 in HCC recurrence within 2 years. Differential overexpression of RACGAP1 transcripts was significantly detected in primary HCC tumors that develop recurrence within 2 years after curative resection (R) compared to non-recurrence tumor (NR), and also in non-recurrence tumor compared to surrounding non-tumorous (ST) tissues by (A) oligonucleotide microarray analysis and (B) real time PCR analysis. NN, normal normal liver from colon metastases; FDR, false discovery rate; bars, median values; P values were calculated from Student's t-test.

4.1.2 Immunostaining validated RACGAP1 upregulation in HCC recurrent samples

We further determined whether the significant upregulation of RACGAP1 transcript in HCC recurrence detected by both microarray analysis and quantitative realtime PCR was recaptured in protein level. We randomly selected 9 pairs of the frozen matched primary HCC tumor with recurrence less than 24 months and their surrounding non-cancerous tissues as well as the other 9 pairs of matched primary non-recurrence tumor and their surrounding tissues to perform immunostaining with RACGAP1 antibody. RACGAP1 staining was mainly confined to the nucleus of tumor cells in which this was mostly detected in primary HCC recurrent and minority in the non-recurrent samples (Fig 4-2A). The frequency of RACGAP1 positively

Figure 4-2. Immunostaining validates clinically significant overexpression of RACGAP1 in HCC recurrence. (A) Representative images of RACGAP1 staining illustrated heterogenous dissemination of RACGAP1+ cells (brown stain). Its expression was mainly localized in the nuclei of tumorous tissues, but was almost undetected in the surrounding non-tumorous (ST) tissues from HCC samples with (Rec) and without (NR) disease recurrence. The images in the lower left corner of each image (20x objective magnification), were photographed in 40x objective magnification. Nuclei of the tissues were counterstained with hematoxylin. (B) RACGAP1 staining was quantitated for 4 subgroups of samples. Bars, median values; the statistical significance between Rec and NR, and between tumor and its matched surrounding tissues, were calculated from the Student's t test and paired sample t test respectively. * denotes significant differential expression of RACGAP1.

113

positive signal = 0.6%). Using 2% average positive staining as a cutoff value, all recurrent tumor samples(100%) and 3 out of 9 non-recurrent tumor (33%) used in IHC analyses in this study showed positive RACGAP1 expression. In contrast, almost all matched surrounding tumor tissues of both recurrent and non-recurrent samples, which were non-cancerous by histological definition did not show immunoreactivity of RACGAP1 (median positive signal = 0%).

Table 4-1. Immunohistochemistry staining of RACGAP1 in all individual HCC recurrence and non-recurrence samples

Non-recurrence			Recurrence			
Average staining			Average staining			
	Matched Normal (%)	Tumor (%)		Months	Matched Normal (%)	Tumor (%)
L137T	0.0	0.3	L82T	1	0.1	4.6
L72T	0.4	1.5	L92T	8	0.0	14.5
L154T	0.0	9.9	L134T	2.3	0.9	27.2
L80T	0.2	0.4	L157T	8	0.0	4.4
L99T	0.0	0.2	L77T	0.6	0.0	4.0
L98T	0.7	2.6	L158T	2	0.2	3.6
L66T	0.0	0.6	L37T	12.6	0.0	7.0
L85T	0.0	0.3	L104T	4.5	0.0	29.6
L152T	0.1	25.3	L71T	9.7	0.0	8.5
Median positive signal	0.0	0.6	Median positive signal		0.0	7.0
no. case positive	0/9	3/9	no. case positive		0/9	9/9

Cases showing more than 2% immunoreactive staining were considered positive

4.1.3 RACGAP1 overexpression is a significant independent risk factor for HCC recurrence

Since recurrence cases consistently demonstrated elevated RACGAP1 expression and several clinical conditions have been previously reported to correlate with HCC recurrence (154), we performed univariate and multivariate Cox regression analysis to further determine whether RACGAP1 expression was confounded by various clinical factors considered in this study. In Univariate analysis, RACGAP1 expression in addition to vascular invasion and cirrhosis were found to be significantly associated with recurrence status of HCC. RACGAP1 expression had equivalent effect (relative risk (RR)=3.42, $p=0.001$) with vascular invasion (RR=3.17, $p=0.001$), but far superior to cirrhosis (RR=1.91, $p=0.048$) in association with disease recurrence after resection. In other words, patients with elevated RACGAP1 expression had recurrence rate that is about 3.4 times higher than those with low RACGAP1. Patients with presence of vascular invasion and cirrhosis had recurrence rate that is about 3 times and double that of those without either one, respectively (Table 4-2).

In a multivariate Cox regression analysis, factors showing non-significance by univariate analysis were eliminated before significant factors were further assessed. Similar to the effects of univariate analysis, RACGAP1 expression, vascular invasion and cirrhosis appeared to be independent risk factors for recurrence. The multivariate analysis showed a 2.7 increased risk of developing recurrence after resection for patients with high RACGAP1 expression compared with those having low expression,

while a respective of 4.06 and 2.55 increased risk of recurrence were observed in patients with vascular invasion and cirrhosis compared with those that did not have either disease (Table 4-2).

Table 4-2. Univariate and multivariate analyses to examine the association of clinicopathological factors and RACGAP1 expression with HCC recurrence

Variable	Univariate analysis		Multivariate analysis	
	RR (95% CI)	p value	RR (95% CI)	p value
Gender	1.56 (0.61-4.02)	0.358		ns
male (n=61) vs female (n=15)				
Age	1.07 (0.55-2.08)	0.839		ns
>60 yr (n=41) vs ≤60 yr(n=35)				
Hepatitis		0.585		ns
HBV (n=59) vs non-B/C (n=14)	0.76 (0.33-1.76)	0.525		
HCV (n=3) vs non-B/C (n=14)	1.47 (0.31-7.09)	0.632		
Encapsulation		0.590		ns
partial (n=19) vs no (n=40)	0.81 (0.36-1.82)	0.604		
complete (n=17) vs no (n=40)	0.63 (0.26-1.57)	0.323		
Tumor size	1.62 (0.84-3.16)	0.153		ns
>5 cm (n=35) vs ≤5 cm (n=41)				
AFP level		0.328		ns
10-300 ng/ml (n=27) vs <10ng/ml (n=31)	1.66 (0.72-3.78)	0.232		
>300 ng/ml (n=21) vs <10ng/ml (n=31)	1.85 (0.78-4.35)	0.161		
Lesion	1.18 (0.46-3.04)	0.732		ns
multiple (n=10) vs single (n=66)				
Differentiation		0.626		ns
G2 (n=42) vs G1(n=10)	1.58 (0.47-5.38)	0.460		
G3 (n=18) vs G1(n=10)	2.23 (0.62-8.01)	0.218		
G4 (n=5) vs G1(n=10)	1.61 (0.27-9.65)	0.601		
Cirrhosis	1.91 (0.95-3.85)	0.048*	2.55 (1.20-5.40)	0.0148*
yes (n=41) vs no (n=35)				
Vascular Invasion	3.17 (1.62-6.21)	0.001**	4.06 (1.93-8.53)	0.0002***
yes (n=29) vs no (n=47)				
RACGAP1	3.42 (1.64-7.16)	0.001**	2.71 (1.27-5.74)	0.0096*
high (n=39) vs low (n=37)				

RR, relative risk; CI, confidence interval; *, P<0.05; **, P<0.005; ***, P<0.001 ; ns., not significant

4.1.4 RACGAP1 expression has prognostic indication for disease recurrence

To determine if RACGAP1 was related to patient prognosis, Kaplan-Meier recurrence analysis of primary HCC was performed based on the transcript signals of

RACGAP1. It appeared that high RACGAP1 expressor group had a significantly shorter mean recurrence-free times (14.6 months, 14.6 months and 16.2 months) when compared to the mean recurrence-free times of low expressor group (29.5 months, 29.5 months and 32 months; $p=0.0026$, $p=0.0008$, $p=0.0005$ for all three comparisons, respectively) regardless those whose disease recurred earlier within 6, 12 until 24 months (Figure 4-3). The 6, 12 and 24-month overall recurrence-free rate were 66.1%, 55.4% and 53.9% respectively. Taken together, these results evidenced that RACGAP1 could be a potentially postoperative strong prognosticator resembled the presence of vascular invasion which is a commonly reported indication of high risk for HCC recurrence.

Figure 4-3. Kaplan-Meier recurrence-free survival analysis of HCC patients according to RACGAP1 expression. The median level of RACGAP1 from all non-recurrent samples and recurrence samples with respective disease occur (A) within 6 months, (B) within 12 months and (C) within 24 months, was designated as the cut-off point for separating RACGAP1 low-expression tumors from RACGAP1 high-expression tumors. R, recurrence; NR, non-recurrence. Statistical significance between two groups was calculated using log rank test (Mantel-Cox).

4.1.5 Endogenous RACGAP1 expression in human hepatocellular carcinoma and other cancer cell lines

The microarray data of the clinical findings evidenced that RACGAP1 expression was elevated in primary HCC samples obtained from patients who developed disease recurrence within two years after curative surgery. In an initial attempt to understand the functional role of this gene in disease recurrence in cell-based assay, we examined its endogenous expression in a few metastatic and non-metastatic HCC cell lines. As previously reported by the providers, MHCC97-H and HCCLM3 were established from the same parental human metastatic HCC cell line, MHCC97 (210). They have identical genetic background and the stepwise increasing metastatic potentials are observed from MHCC97-H (211) to HCCLM-3 (212). Consistent to the clinical observation, our study showed that both metastatic HCCLM3 and MHCC97-H cell lines expressed far higher endogenous RACGAP1 protein (Figure 4-4A, left panel) and mRNA (Figure 4-4B) than the other 6 non-metastatic (HepG2, Hep3B, PLC/PRF 5 and Huh7) HCC cell lines. In addition, the expression pattern of RACGAP1 as revealed by the presence of two bands, was only observed in some cell lines, particularly HCCLM3 and MHCC97-H cells. This may imply different level of phosphorylation in response to their different origins of genetic background, except that HCCLM3 and MHCC97-H were derived from a same parental clone. RACGAP1 up-regulation was found not only specific to HCC, as its expression was also detected in other cancerous cell lines, for instance lung cancer (A549), breast cancer (MCF7), nasopharyngeal

carcinoma (CNE2), keloid fibroblasts (KF4 P/P), as well as normal fetal lung fibroblast (MRC-5) cell lines (Figure 4-4A, right panel).

Figure 4-4. RACGAP1 endogenous expression in HCC and other cancer cell lines. Detection of RACGAP1 endogenous (A) protein expression in various HCC cell lines (left panel) and other cancer cell lines (right panel) by using western blot and (B) its mRNA expression in HCC cell lines by using quantitative real-time PCR, demonstrated concordant results. HCC, hepatocellular carcinoma tissues; ST, surrounding non-malignant tissues.

4.1.6 siRNA-mediated silencing of RACGAP1 in metastatic HCC cell lines

Having found that the upregulation of RACGAP1 in HCC correlates with poor outcome and its strong prognostication effect for recurrence of HCC, we investigated the phenotypical changes disfavoring the survival of tumor cells *in vitro*, by silencing RACGAP1 in HCCLM3 and MHCC97-H metastatic cell lines expressing high endogenous RACGAP1 protein (Figure 4-5). The repression of endogenous RACGAP1 mediated by siRNA duplex was evaluated at 3 timepoints. The depletion of this gene was checked by performing western blot of the cell lysate harvested at day 2, day 4 and 6 after transient transfection using 2 pools of siRNA duplexes (siR1 and siR2) against RACGAP1. The suppression in both cell lines began at day 2 using either siR1 or siR2 with a more prominent effect observed in HCCLM3 cells. Depletion of endogenous RACGAP1 employing siR2 was superior to siR1 in MHCC97-H cells on both day 2 and day 4 (lower panel, Figure 4-5). The blockage effect was drastically observed in MHCC97-H cells and consistently remained in HCCLM3 at day 4 using either one pool of siRNA, but the endogenous protein expression of RACGAP1 recovered to almost the same level as the RNAi duplex universal negative control transfected cells on day 6. Hence, the subsequent RACGAP1 repression experiments were all performed using siR2 and its endpoint analysis were fixed on day 4 post silencing of RACGAP1.

Figure 4-5. Time course analysis of siRNA-mediated silencing effect on RACGAP1 protein level in HCCLM3 and MHCC97-H cells. The depletion of endogenous RACGAP1 protein in two HCC metastatic cell lines, HCCLM3 and MHCC97-H after treatment with two independent pools of siRNA duplexes (siR1 and siR2) at a final concentration of 100nM each, was determined and compared to the RNAi duplexes universal negative control (C) treated and non-treated (Null) cells for 2, 4 and 6 days by using western blot.

4.1.7 Silencing of RACGAP1 results in suppressed migration and invasion in HCC cells

Following transfection with siRNA targeting RACGAP1 for 2 days, the transfected cells were reseeded in the migration or invasion chamber for another 2 days before migration or invasion measurement of migration and invasion were taken. The migration and invasiveness of both HCCLM3 ($p=0.002$ and 0.018) and MHCC97-H ($p=0.027$ and 0.002) cell lines were significantly impaired after suppression of endogenous RACGAP1 for 2 days followed by seeding in the migration or invasion

chamber for another 2 days, as compared to the RNAi negative control-treated cells (Fig. 4-6), suggesting the role of RACGAP1 in HCC recurrence.

Figure 4-6. Suppressed cell migration and invasion after siRNA-mediated depletion of RACGAP1 *in vitro*. Silencing of RACGAP1 in (A) HCCLM3 and (B) MHCC97-H cells significantly impaired cell migration (left panel) and invasion (right panel). Representative fluorescent images of the migrated (upper panel) and invaded (lower panel) (C) HCCLM3 and (D) MHCC97-H cells were stained with DAPI 2 days after depletion of RACGAP1. Null, no treatment; C, RNAi duplex universal negative control; siR2, pool of RACGAP1 siRNA duplexes. Each of the treatment shown was representative of three independent experiments and average value was obtained from five non-overlapping fields photographed in each membrane insert.

4.1.8 Silencing of RACGAP1 promotes cell death in HCC cells

One of the hallmark of apoptosis, DNA fragmentation increased drastically in both HCCLM3 and MHCC97-H cell lines irreversibly beginning on day 4 after siRNA-mediated repression of RACGAP1 expression, as detected by flow cytometry analysis (Figure 4-7A and B). The positive signal, green fluorescent stained nuclei co-localized with DAPI nuclei staining, as detected by fluorescent microscopy, were pronounced in both cell lines starting from day 4, and peaked on day 6 irreversibly after suppression of RACGAP1 (Figure 4-7C). A number of molecules involved in a cascade of caspase activation events which ultimately lead to cell death, including cleaved caspase 9, cleaved caspase 7 and cleaved PARP were activated (Figure 4-7D), suggesting the cells undergoing apoptosis, in the absence of RACGAP1.

Figure 4-7. TUNEL assay after siRNA-mediated silencing of RACGAP1. DNA fragmentation was significantly augmented in (A) HCCLM3 and (B) MHCC97-H cells 4 days after depleting of RACGAP1 as compared to the respective RNAi negative control treated cells. Data were obtained from 3 independent experiments (C) Representative fluorescent images of TUNEL-positive MHCC97-H cells undergoing DNA fragmentation illustrated by green nuclei stain overlaid with DAPI nuclei counterstain following knockdown of RACGAP1 starting from day 4 and peaked at day 6. (D) Suppression of RACGAP1 initiated a series of caspase activation including cleaved Caspase-9, 7 and PARP, which lead to cell death. Null, non-treated cells; C, RNAi negative control transfected cells; siR2, RACGAP1 siRNA duplexes transfected cells.

4.1.9 Silencing of RACGAP1 represses Cdc42 and Rac1 rather than RhoA activation

Given the GAP domain activity of RACGAP1 strongly stimulates Cdc42 and Rac1 GTPase but is almost inactive on RhoA (131, 137), alteration of RACGAP1 expression might affect the GTP bound or active form of such small GTPases. To obtain semi-quantitative pull down assay, fold ratio of the pull-downed active GTPases (GTP-Cdc42, GTP-Rac1 and GTP-Rho A) were generated by normalizing expression signal of each sample to its corresponding total protein, and the ratios were again normalized to the untreated sample (Null) of the same cell line. The fold ratio for RACGAP1 in each treatment was generated in the similar way. The ratio signal of the respective lower and upper bands in relative to the corresponding actin expression were calculated, followed by normalizing the ratio signals respectively to the corresponding ratio of the bands in the untreated control (Figure 4-8).

It appeared that the siRNA-mediated silencing of RACGAP1 in HCCLM3 was more effective than in MHCC97-H cells. Both lower and upper bands signal (fold ratio of lower band = 0.03 and upper band = 0.55) in HCCLM3 cells were much reduced when compared to their corresponding RNAi negative controls (fold ratio of lower band = 0.98 and upper band = 0.79). On the other hands, only the lower band in MHCC97-H cells was repressed (fold ratio = 0.32) when compared to its corresponding lower band signal in RNAi duplex negative control treated cells (fold ratio = 0.67), but upper band signal (fold ratio = 0.74) remained close to their

corresponding RNAi duplex negative control treated cells (fold ratio of upper band = 0.69).

Among 3 active GTPases examined, GTP-Rac1 was reduced consistently and drastically in both HCCLM3 and MHCC97-H cells (Fold ratio = 0.22 and 0.4) upon RACGAP1 depletion (siR2), compared to the RNAi negative universal control (C) treated-cells (Fold ratio = 1.14 and 1.07). GTP-Cdc42 was mildly repressed in RACGAP1-depleting HCCLM3 cells (Fold ratio = 2.24) compared to its corresponding RNAi negative control treated-cells (Fold ratio = 3.19). Meanwhile, GTP-Cdc42 reduction was more prominent in RACGAP1- depleting MHCC97-H cells (Fold ratio = 0.43), in comparison to RNAi negative control cells (Fold ratio = 1.12). Nonetheless, active Rho was not affected in HCCLM3 depleting RACGAP1, although slight reduction was detected in RACGAP1-silencing MHCC97-H cells (Fold ratio = 0.79) (Figure 4-8). Taken together, siRNA mediated-silencing of RACGAP1 in HCCLM3 and MHCC97-H impaired primarily active Cdc42 and Rac1 rather than Rho.

Figure 4-8. Deactivation of small GTPase pull-down assay after silencing of RACGAP1. RACGAP1 silencing impairs small GTPases activation, specifically GTP-Cdc42 and Rac1, rather GTP-Rho A. Numbers below RACGAP1 blot represent fold ratio of each individual band relative to untreated control (Null) of the same blot after normalization against actin of the same treatment. For quantitation of GTP-small GTPases, fold ratio of each band was obtained in the same way, except after normalization against respective total GTPases. Null, untreated cells; C, RNAi duplex negative universal control treated-cells; siR2, RACGAP1 siRNA duplexes treated-cells.

4.1.10 Exogenous overexpression of RACGAP1 results in increased cell proliferation in PLC/PRF 5 cells

Having observed that the reduction of RACGAP1 resulted in a few phenotypic changes that attenuated the growth of HCC cells, we further investigated if the inverse phenotypes were true by etopically overexpressing full-length of RACGAP1 encoded in pIRES vector (Rg-pIR) in a low endogenous RACGAP1-expressing HCC cell line determined earlier (Figure 4-4), PLC/PRF 5 cells. As a result, a significant increase of cell proliferation after etopic RACGAP1 overexpression for 24 hours, determined by FACS sorting of a cell proliferating marker, Ki-67 together with RACGAP1 staining (Figure 4-9A). To examine the specificity of primary anti-RACGAP1 antibody, isotype control was included using isotype-matched IgG2b as a replacement for primary RACGAP1 antibody. The Ki-67 positive population in RACGAP1 transfected cells (median signal = 5.0%) was increased significantly ($p=0.004$), with about 3 fold higher than those in empty IRES vector (pIR) transfected cells (median = 1.5%), as demonstrated in Figure 4-9B, suggesting cell proliferation was induced under the forced expression of exogenous RACGAP1.

Figure 4-9. Overexpression of RACGAP1 promotes cell proliferation in PLC/PRF 5 cells. (A) FACS sorting analysis of Ki-67 and RACGAP1 positively stained population in cells transfected with pIRES encoding RACGAP1(Rg-pIR) for 24 hours, compared to cells transfected with pIRES vector alone (pIR) and untreated cells (Null). Isotype control was performed by replacing RACGAP1 antibody with isotype-matched IgG2b. (B) Significant increase of Ki-67 signal was observed in the RACGAP1-overexpressing (Rg-pIR) cells as compared to pIR-transfected cells. Data shown were obtained from 3 independent experiments.

4.1.11 Exogenous overexpression of RACGAP1 enhances cell migration and invasion

In contrast to silencing endogenous RACGAP1 in HCCLM3 and MHCC97-H cells, overexpression of this gene in PLC/PRF 5 cells enhanced cell migration and invasion. The number of migrated and invaded cells were both increased prominently 2 days after seeding the 24 hours RACGAP1 DNA post-transfected (Figure 4-10A) cells in the membrane inserts, as compared to the empty pIRES transfected cells (Rg-

pIR versus pIR in Figure 4-10B). The increase of cell migration and invasion was significant ($p = 0.036$ and 0.02 ; Figure 4-10B) after performing 3 sets of independent experiments.

Figure 4-10. Overexpression of RACGAP1 promotes cell migration and invasion in PLC/PRF 5 cells. (A) Transient overexpression of RACGAP1 for 24 hours in PLC / PRF 5 using pIRES vector results in ectopic expression of the third band (the uppermost band) as compared to the pIRES-transfected and

untreated cells. (B) The cells migration and invasion were augmented profoundly after reseeding the RACGAP1-overexpressing (Rg-pIR transfected) cells as compared to pIR-treated cells. Data shown were collected from 3 independent experiments. (C) Representative fluorescent images of the migrated (upper panel) and invaded (lower panel) cells 2 days after reseeding RACGAP1-overexpressing cells 24 hours post transfection. The cells remained in the innerside were removed and those migrated or invaded to the underside of the membrane were fixed and counterstained with DAPI.

4.1.12 Exogenous overexpression of RACGAP1 activates Cdc42, Rac1 and RhoA

Unlike the RACGAP1-silencing experiments, stimulation of GTP-RhoA (fold ratio = 6.64) was observed in the RACGAP1-overexpressing cells (Rg-pIR in Figure 4.11). Both Cdc42 and Rac1 were also activated (fold ratio = 1.6 and 1.75) as compared to their respective pIRES transfected-cells (fold ratio = 1.16 and 0.82), which were close to the untreated controls. In an attempt to quantitate the multiple bands representing various phosphorylation after transiently overexpression of RACGAP1, we normalized signal of each band to actin expression of the same sample to generate fold ratio as in the previous RACGAP1-silencing experiments. Subsequently, the fold ratio was normalized to RACGAP1-overexpressing (Rg-pIR) cells, instead of the untreated control (Null), due to the unique existence of ectopic hyper-phosphorylated band (Upper band in Figure 4-11). To correlate small GTPase activation with RACGAP1 overexpression, we found that all three GTPases were activated when the unphosphorylated (lower band) form and the ectopic hyper-phosphorylated (upper band) form of RACGAP1 were stimulated, but the phosphorylated RACGAP1 (middle band) remained unchanged under the forced

expression of exogenous RACGAP1, using pIRES plasmid DNA. This suggested the GTP-bound of Cdc42, Rac1 and Rho A were all activated when the unphosphorylated and ectopic hyper-phosphorylated RACGAP1 were stimulated in the low endogenous unphosphorylated RACGAP1-expressing PLC/PRF 5 cells.

Figure 4-11. Overexpression of RACGAP1 activates small GTPases in PLC/PRF 5 cells. Transient overexpression of RACGAP1 enhances small GTPases activation, including GTP-Cdc42, GTP-Rac1 and GTP-Rho. Numbers below RACGAP1 blot represent fold ratio of each individual band relative to Rg-pIR of the same blot after normalization against actin of the same treatment. For quantitation of GTP-small GTPases, fold ratio of each band was relative to untreated control after normalization against respective total GTPases. Null, untreated cells; pIR, pIRES-transfected cells; Rg-pIR, RACGAP1-pIRES transfected cells.

4.1.13 Genome-wide analysis of siRNA-mediated silencing of RACGAP1 in MHCC97-H cells

To gain foresight concepts of the potential deregulated molecules or pathways upon deactivation of RACGAP1, which might be useful in clinical setting of HCC recurrence, we performed microarray analysis to compare high endogenous RACGAP1 expressing metastatic MHCC97-H cells 4 days after siRNA-mediated silencing of this gene with RNAi negative control-treated and the untreated cells *in vitro*. Four samples each treatment were generated and the cel files were analysed using GC-RMA algorithm provided in Partek software for normalization. All data were preprocessd by removing scan date batch effect before performing two filtering approaches. Probe sets that have the least difference between RNAi negative control-treated and untreated samples were first generated with statistic filtering criteria which are inverse of the second criteria (p value less than 0.05, false discovery rate (FDR) less than 0.1 or 10% and fold change greater than 1.25) used later to differentiate RACGAP1 siRNA-treated from RNAi negative control-treated samples. Consequently, a common list of 2569 probe sets including 1145 up-regulated and 1424 down-regulated probe sets that passed the 2 filtering approaches was obtained (Figure 4.12A). Principle component analysis of this 2569 differential probe sets demonstrated a distinct segregation of the RACGAP1-silencing samples from the RNAi negative control-treated and the untreated control samples (Figure 4.12B). More importantly the target gene of knockdown, RACGAP1 is significantly repressed in RACGAP1 siRNA-treated samples in comparison to RNAi negative control-treated

samples($p=0.0016$, Fold change $=-3.5$, FDR=3.02%; Figure 4.12C).

Figure 4-12. Gene expression profiling of siRNA-mediated RACGAP1-depleting cells. Differential gene expression analysis using microarray approach to compare the RACGAP1-depleting MHCC97-H cells with the RNAi negative control cells following (A) a series of filter criteria to yield a final list of 1145 up-regulated and 1424 down-regulated probe sets. (B) Principle component analysis demonstrated the clustering of the RACGAP1-deficient samples were distinct from the RNAi negative control-treated and the untreated control samples using the 1145 up-regulated and 1424 down-regulated probe sets. (C) RACGAP1 is significantly repressed in RACGAP1 siRNA-treated samples in comparison to RNAi negative control-treated samples. Statistical significance was calculated using one way ANOVA test from Partek software. C, RNAi duplex negative universal control treated-cells; siR2, RACGAP1 siRNA duplexes treated-cells.

4.1.14 Functional analysis of differential genes after knockdown of RACGAP1 in MHCC97-H cells

To elucidate the specific biological functions that are significantly associated with the 2569 differential probe sets generated earlier, we employed Ingenuity

Pathway analysis and the top ten canonical pathways that were distinctly deregulated were shown in Table 4-3. The two most deregulated pathways are molecules associated with mitotic polo-like-kinase as well as BRCA1-related genes in DNA damage pathways. Most of the molecules related to these 2 pathways are suppressed in response to RACGAP1 downregulation.

Table 4-3. Top 10 deregulated canonical pathways in RACGAP1-deficient MHCC97-H cells.

Pathways	-log(p-value)	Ratio	Differential expression	Molecules
Mitotic Roles of Polo-Like	5.56	0.29	Up Down	PPP2R5B,CDC16 KIF23,CDC25C,CCNB1,ESPL1,CDC20,PPP2R5C,PTTG1,PRC1,CCNB2,ANAPC10, PLK1, CDC2,SLK,PLK4,PKMYT1, PPP2R5E
Role of BRCA1 in DNA	4.75	0.302	Up Down	GADD45A RBL2,FANCG,FANCF,SMARCD2,RPA1,PLK1,SMARCA4,RAD51,POU2F1,FANCD2,RFC4,RFC2,BRCA1,FANCA, RFC3
ATM Signaling	3.44	0.269	Up Down	CREB3,CREB5,JUN,GADD45A CDC25C,CCNB1,MDC1,CCNB2,CDC2,RAD51,FANCD2,CREB1,BRCA1,CDK2
Molecular Mechanisms of Cancer	3.36	0.142	Up Down	MRAS,GNA15,GNA13,SMAD2,ARHGEF4,CREBBP,RHOQ,MAX,GAB1,FYN, RELA,RALA,HIF1A,EP300,JUN,NLK,TGFB2,PIK3C2A,NF1,RBPJ,BCL2L11, BIRC2 APH1B,MAPK1,SMAD3,SUV39H1,MYC,PAK1,CAMK2D,HIPK2,BRCA1, SMAD1,PIK3C2B,RALB,CDK6,AURKA,RND3,CDK2,CTNNA1,BMPR2,PRKAG1,FANCD2,RHOV,CASP8,PRKCA,PRKDC,CDC25C,GNAQ,PIK3CB,PRKCB, CTNND1,FZD7.
ERK5 Signaling	2.94	0.217	Up Down	LIF,LOC729991-MEF2B,CREB3,CREB5,GAB1,MRAS,GNA13,MAP3K2 SH2D2A,RP86KB1,GNAQ,EGF,MYC,CREB1,WNK1
Apoptosis Signaling	2.84	0.194	Up Down	BIRC2,CAPN3,DFFB,TRA2,MAP4K4,MRAS,PLCG1,RELA,ROCK1 AIFM1,CASP8,CDC2,DFFA,LMNA,MAPK1,PARP1,PRKCA,TNFRSF1A
HGF Signaling	2.81	0.192	Up Down	DOCK1,ETS2,GAB1,HGF,IL6,JUN,MAP3K12,MAP3K2,MRAS,PIK3C2A,PLCG1 CDK2,MAP3K13,MAPK1,PAK1,PIK3C2B,PIK3CB,PRKCA,PRKCB,RAPGEF1
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	2.37	0.233	Up Down	EP300,GADD45A BRCA1,CCNB1,CCNB2,CDC2,CDC25C,PLK1,PRKDC,SKP2
p53 Signaling	2.27	0.191	Up Down	DRAM (includes EG:55332),EP300,GADD45A,JUN,PIK3C2A,SCO2 (includes EG:9997),SIRT1,STAG1,TP53INP1 BIRC5,BRCA1,CDK2,HIPK2,MED1,PIK3C2B,PIK3CB,PRKDC
TGF- β Signaling	2.27	0.186	Up Down	ACVR1B,ACVR1C,CREBBP,EP300,HOXC9,JUN,MRAS,RUNX2,SERPINE1, SMAD2,TGFB2,VDR BMPR2,MAPK1,SMAD1,SMAD3

Cutoff point: $-\log(p \text{ value}) > 1.3$ equals to $p \text{ value} < 0.05$. Up, up-regulation; Down, down-regulation

4.1.15 Comparative functional analysis of cell-based depletion and clinically overexpression of RACGAP1

Using Ingenuity Pathway network analysis which depicts RACGAP1 neighborhood molecules or the literaturally well curated molecules associated directly or indirectly with RACGAP1, the affected targets (PRC1, AURKB, CDC2, ECT2, KIF23, PAK1 and PPP2R5E) were all downregulated after overlaying with the significant 2569 differential signatures (Figure 4-13A) . In contrast, the overlay with significant differential list from comparison between HCC clinical samples with disease recurrence within 2 years and non-recurrent samples revealed overexpression in all affected RACGAP1 target molecules (PRC1, TGFB1, ECT2, SH3RF1, SFN, PRSS23) shown in Figure 4-13B. After comparing two deregulated RACGAP1 neighborhood gene lists (Figure 4-13A versus 4-13B), we found two RACGAP1 related genes, PRC1 and ECT2 which were elevated in HCC recurrence samples at diagnosis, but down-regulated upon suppression of RACGAP1 in metastatic MHCC97-H cells *in vitro*.

Figure 4-13. Comparison of aberrant signaling networks resulted from cell based-depletion of RACGAP1 and clinical up-regulation of RACGAP1 in HCC recurrence. Ingenuity network comparison analysis demonstrated that PRC1 and ECT2 were altered in inverse pattern, by comparing the overlays of RACGAP1 signaling network with differential genes in (A) RACGAP1-depleting MHCC97-H cell lines and in (B) HCC clinical samples with disease recurrence within 2 years. Green and red highlights represent decreased and increased expression; Solid red arrows depict direct activation, dashed red arrows depict indirect activation, solid pink arrows depict phosphorylation or dephosphorylation, Grey dashed arrows depict indirect expression, solid blue lines depict protein interaction or binding. R, HCC recurrence within 2 years; NR, non-recurrent samples; C, RNAi duplex negative universal control treated-cells; siR2, RACGAP1 siRNA duplexes treated-cells.

4.1.16 Interaction of the RACGAP1 signaling networks and pathways that are closely associated with cell motility

To further explore the intermediate regulators that relate these two distinct perturbed pathways to RACGAP1 following depletion of this genes, we overlay these pathways with the perturbed RACGAP1 network molecules as shown in Figure 4-14. The results indicated that RACGAP1 depletion linked closely to deregulation of

138

RACGAP1 signaling network. For visualization, network was mainly restricted to genes showing altered regulation after depleting of RACGAP1, and interaction between genes from other pathways are not all indicated.

4.1.17 Validation of deregulated molecules in response to RACGAP1 silencing

To ascertain deregulation mRNA expression of the molecules detected by genome-wide expression profiling, in response to RACGAP1 suppression, we investigated the protein expression of a few selected genes by performing western blot. The results revealed down-regulated expression of the selected molecules in protein level, of which mostly are in phosphorylated form including phospho-PRC1, ECT2, phospho-AURKB, PLK4, phospho-BRCA1, phospho-ERK1/2 and phospho-AKT^{Thr308} (Figure 4-15). These correlate with the observed decreased mRNA level detected in microarray expression profiling, except that PLK1 protein level remained unchanged, despite its down-regulated mRNA level.

Figure 4-15. Verification of the altered target molecules upon silencing of RACGAP1 by western blot. The alteration of a few RACGAP1 associated molecules selected from the altered RACGAP1 signaling networks and the top deregulated pathways, were validated by immunoblot. Null, untreated cells; C, RNAi duplex negative universal control treated-cells; siR2, RACGAP1 siRNA duplexes treated-cells.

4.1.18 LC-MS/MS analysis of RACGAP1-depleted MHCC97-H cells

In an attempt to correlate the perturbed mRNA transcripts level with protein expression for the list of 2569 differential signatures in a global-system level, we performed proteomic approach by employing LC-MS/MS analysis, the

same way as in microarray analysis, to compare RACGAP1 siRNA-treated with RNAi negative control-treated MHCC97-H cells. Two independent duplicates were performed for each treatment and an initial of 85 proteins consisting of 46 up-regulated and 36 down-regulated proteins were found to be differentially expressed between RACGAP1 deficient cells and RNAi negative control cells. When we overlapped the 85 differential proteins with the 2569 differential transcripts found earlier, we obtained 8 consistent differential signatures comprising of one up-regulated molecule (HSPA5) with ratio greater than 1 and 7 down-regulated molecules with ratio less than 1 (AKR1B10, ALDH1A1, ANXA1, PTMA, SET, TMPO and UGDH) as shown in Table 4-4. However, the primary deregulated expression of RACGAP1 was not captured in the list.

Table 4-4. Differential molecules consistently detected by microarray and LC-MS/MS approaches in RACGAP1-depleting MHCC97-H cells.

Probe Set ID	Gene Symbol	Gene Title	Cytoband	SwissProt	siR2 versus C	
					mRNA (Fold)	Protein (ratio)
230031_at	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	9q33-q34.1	P11021	2.558	1.43
206561_s_at	AKR1B10	aldo-keto reductase family 1, member B10 (aldose reductase)	7q33	O60218	-1.751	0.94
212224_at	ALDH1A1	aldehyde dehydrogenase 1 family, member A1	9q21.13	P00352	-2.297	0.93
233011_at	ANXA1	Annexin A1	9q12-q21.2 9q12-q21.2	P04083	-2.749	0.96
211921_x_at	PTMA	prothymosin, alpha	2q35-q36	P06454	-1.263	0.88
215780_s_at	hCG_1644608 /// SET	SET translocation (myeloid leukemia-associated) pseudogene /// SET nuclear oncogene	9q34 /// Xq21.1	Q01105	-2.166	0.84
200630_x_at	SET	SET nuclear oncogene	9q34	Q01105	-1.431	0.84
200631_s_at	SET	SET nuclear oncogene	9q34	Q01105	-1.586	0.84
213047_x_at	SET	SET nuclear oncogene	9q34	Q01105	-1.511	0.84
40189_at	SET	SET nuclear oncogene	9q34	Q01105	-1.600	0.84
203432_at	TMPO	thymopoietin	12q22	P42166	-1.895	0.82
203343_at	UGDH	UDP-glucose dehydrogenase	4p15.1	O60701	-1.254	0.86

siR2, RACGAP1 siRNA duplexes-treated cells; C, RNAi duplex universal negative control-treated cells.

4.2 Discussion

4.2.1 Clinical significance of RACGAP1 up-regulation in HCC recurrence

We here report that RACGAP1 (Rac GTPase activating protein 1), which we have previously found as one of the most consistent up-regulated gene from a 57 molecular predictor for HCC recurrence with either vascular invasion or cirrhosis (154), is significantly overexpressed in all primary HCC samples with early recurrence within 2 years after surgical resection compared to non-recurrent samples. The confidence of our finding in transcript level (fold change =2.31, $p < 0.0001$) employing microarray approach is statistically high as the false discovery rate is less than 1% (FDR=0.066%). RACGAP1 has not been previously implicated in HCC. Its upregulation or its role in favoring cancer cell proliferation has been described within a cluster of genes, in a more aggressive cancer stage of several other cancer microarray analyses including epithelial ovarian cancer (155), cervical cancer (213), high grade breast cancer with transition from preinvasive to invasive stage (214) as well as Estrogen Receptor (ER) positive breast cancer (158-160). The expression of this gene in either all the non-recurrent HCC or the tumorous HCC samples ($n=76$) is remarkably higher compared to the surrounding tumor tissues that are histologically non-tumorous. The observation is strongly concordant with the result obtained by using alternative method, quantitative realtime PCR which validates the elevated mRNA expression of RACGAP1, and immunohistochemistry staining findings which correlate the RACGAP1 mRNA up-regulation with protein overexpression in HCC

recurrence. This extrapolates the role of RACGAP1 in disease progression particularly from primary cancer development to disease recurrence which is the major obstacle for treatment of HCC via partial hepatectomy (84). However, the immunostaining of RACGAP1 reveals no significant difference between paired non-recurrent and the corresponding surrounding non-tumorous samples, despite the significant difference between paired recurrent and surrounding non-tumorous tissues exists.

4.2.2 Up-regulation of RACGAP1 has prognostic indication for HCC recurrence within 2 years.

Our results suggest that HCC patients with primary tumor overexpressing RACGAP1 are significantly at high risk of developing cancer recurrence. Univariate analysis demonstrates that RACGAP1 has equal significant effect as vascular invasion but more prominent than cirrhosis in association with HCC recurrence. The recurrence rate, about 3.4 times was observed in patients with elevated RACGAP1 expression relative to tumor with low RACGAP1, which is close to the recurrence rate in patients diagnosed with vascular invasion (3.17). In addition to vascular invasion and cirrhosis, multivariate analysis further proves that RACGAP1 expression, is an independent prognostic factor for disease recurrence within two years and its effect is far superior to cirrhosis. Hence, our attempt has validated that patients with high tumor RACGAP1 expression have far poorer prognosis than those with low expression, regardless disease recurrence occur within 6, 12 and 24 months, strengthening that RACGAP1 as

a promising independent prognosticator for early HCC recurrence after curative resection.

Recurrence of HCC is characterized to two categories, firstly, a true metastasis, which is resulted from the residual intrahepatic recurrence due to tumor factor or tumor cells dissemination from primary HCC before resection, and secondly, multicentric metastasis which reflects a de-novo primary tumor in the preneoplastic liver remnant influenced by the underlying liver disease such as continuous virus infection, inflammation and cirrhosis (83-85). The former mainly gives rise to early recurrence which is undetectable within the first 2 years after surgery (87, 88) and in general carries poorer prognosis than that with multicentric occurrence which usually appears as late recurrence (215, 216). A recent study employing viral genome-wide RNAi/genetic screens in A549 cell line identified a number of host cellular factors including RACGAP1 that were required for early-stage influenza A virus replication (161), presumably by regulating IP3-PKC signaling pathway and cytoskeletal organization during viral entry and intracellular viral transport (162). Human cellular cofactors, associated with cytoskeleton organization, is the only complex commonly implicated in influenza, HIV, HCV, dengue and West Nile virus (WNV) replication (162). As RACGAP1 was implicated in HCC recurrence in patients either having vascular invasion or cirrhosis in our earlier study (154), we postulate that RACGAP1, besides its role in early HCC recurrence due to tumor factor caused by vascular invasion, could also be involved in a similar host cellular mechanism facilitating HBV replication in the non-tumorous liver remnant, leading to cirrhosis and ultimately late

recurrence. However, the disease recurrence of our patient cohort in this study only capture up to 35 months after surgery, thus limiting the study of the late recurrence group.

4.2.3 Functional analysis of RACGAP1 via knockdown and overexpression experiments using cell-based approach.

Despite many studies reported RACGAP1 in regulating furrow ingression for completion of cytokinesis via interaction with Rac1, Cdc42 and RhoA, and mitosis in hematopoietic cells, the mechanism by which RACGAP1 contributes to HCC progression, particularly disease recurrence remains unclear. Our study to elucidate the functional role of this gene in HCC recurrence via cell-based assay using siRNA-mediated silencing approach in the high endogenous RACGAP1 and metastatic cell lines, HCCLM3 and MHCC97-H, results in reduced cell migration and invasion, suppressed active Cdc42 and Rac1, and a strong augmentation of DNA fragmentation leading to massive cell death *in vitro*. Intriguingly, the inverse phenotypes, including increased cell migration and invasion, activated Cdc42, Rac1 and Rho A, and enhanced cell proliferation are observed by exogenously overexpressing this gene in low endogenous RACGAP1-expressing PLC/PRF 5 cells.

RACGAP1 was recently uncovered by phosphoproteomic approach as a phosphoprotein to interact with Aurora B (217), PP2A and Cdk1 in a phosphoregulation manner during the progression of mitosis to cytokinesis (152). The

disparity between the gain and loss of RACGAP1 function assays in which deactivation of GTP-bound Cdc42 and Rac1, but not RhoA in RACGAP1 knockdown experiment, and stimulation of all three active GTPase molecules in the overexpression experiment might be due to the acquisition of an additional ectopic hyper-phosphorylated form of RACGAP1 in the latter experiment. Taken together, our results propose that RACGAP1 regulates primarily Rac1 and Cdc42 activation, which are known to influence actin cytoskeleton reorganization necessary for cell morphology, cell migration, chemotaxis and establishment of cell polarity (218, 219), that may lead to tumor metastases. The knockdown experiment is concordant with the previous report which characterized RACGAP1 as a human chimaerin-like protein in male germ cells showing GAP activity mainly toward Rac1 and Cdc42 (131) that are crucial for initiating cytokinesis (136, 144). Moreover, microarray profiling of the knockdown experiments showed that PAK1, a main downstream effector of Rac1 and Cdc42 is suppressed in response to RACGAP1 depletion. Its overexpression in human HCC and its role in cancer metastasis has been reported (220). Hence, the knockdown of RACGAP1 may abolish or affect the PAK1-mediated mechanism contributing to cancer progression and metastasis including cell motility, survival, mitosis and angiogenesis (221).

Suppression of RACGAP1 promotes apoptosis. The cleavage of one of the initiator caspase, caspase-9 propagates a cascade of caspase processing (222, 223) events by activating an effector, caspase-7 which in turn cleaves PARP (224, 225), thereby triggering apoptosis. This is in line with the DNA fragmentation that

augments in response to blockage of RACGAP1. In contrast, cell proliferation increases as the Ki-67 co-expression increases with the forced expression of RACGAP1. Hence, we postulate that RACGAP1 plays a pivotal role in sustaining survival and proliferation of HCC cells *in vitro*. The findings of cell-based assay that manipulates the expression of RACGAP1 by siRNA-mediated depletion was compared to the clinical finding, and both were summarized in Figure 4-16. This might be useful in clinical management of HCC recurrence, by inferring potential deregulated molecules or pathways upon deactivation of RACGAP1 *in vitro*, if therapy is designed to target RACGAP1.

Figure 4-16. Summary of the significant findings of RACGAP1 in both clinical and cell based assays. Significant clinical implications of RACGAP1 in (A) disease recurrence within 24 months versus non-recurrence of HCC. (B) Inverse phenotypes were found in cell-based assay by knocking down and overexpression of RACGAP1 in HCC cell lines.

4.2.4 Pathway and network analysis of RACGAP1 loss-of-function in vitro reveals mitotic polo-like-kinase as the most affected signaling pathway.

In vertebrate cells, RACGAP1 associates with KIF23 in a complex known as centralspindlin which plays an essential role in cytokinesis(143). RhoGAP RACGAP1 interacts with RhoGEF Ect2 in a cell cycle-dependent manner to recruit cortical localization of RhoA, which is involved in activating actin polymerization and myosin-II necessary for assembly and ingression of the contractile ring to complete the cytokinesis process (144-146). RACGAP1, together with ANLN, ECT2, AURKB, PRC1 and KIF23 (MKLP1) are found to be cytokinesis-related cluster genes. The knockdown of these genes individually employing high-throughput RNAi profiling in HeLA cells confers to cell division defect, particularly cytokinesis defect phenotypes (226). Here, we use Ingenuity knowledge base analyses in an attempt to identify specific molecular targets associated with RACGAP1, that govern the invasive and survival phenotypes that may lead to recurrence of HCC. Intriguingly, the overlay between RACGAP1 neighbourhood clusters and the significant differential signatures in HCC recurrence in relative to non-recurrence samples (Figure 4-13B) reveals that RACGAP1 is a downstream target of TGFB1 and it may be involved in TGFB pathway. TGFB signaling is known to play a dual role in tumorigenesis. It acts as a tumor suppressor at early tumorigenesis stage, but enhances tumor progression, invasion, metastasis, and induces epithelial to mesenchymal transition at later stages (227).

The most perturbed canonical pathway is mitotic roles of polo-like kinase and this is caused by RACGAP1, the cytokinesis regulator upon its deactivation is thought to deregulate the major mechanism involved in cell division. The majority of the molecules in this pathway are suppressed in response to RACGAP1 downregulation. The overlay with RACGAP1 network derived from the differential molecules between RACGAP1 siRNA treated cells relative to the control siRNA treated cells uncovers a number of intermediate modulators including KIF23, PRC1, PPP2RE, PPP2RC and PPP2RB that connect RACGAP1 and polo-like kinase mediated mitosis. Indeed, the Cdc42, Rac and actin cytoskeleton signaling are indirectly deregulated by RACGAP1 blockage via a common association with ERK pathway mediated by AURKB. ERK, one of the family member in MAPK, once downregulated in response to RACGAP1 silencing, exerts distinct cellular phenotypic changes such as enhanced apoptosis and decreased invasion seen in this study, that has been previously reported (228-230). In addition to GEF domain, ECT2 contains two BRCA1COOH terminus (BRCT) domains in its NH₂-terminal region that bind to RACGAP1 via phosphorylation (144). Here, we postulate that the blockage of RACGAP1 abolish interaction with ECT2, leading to deactivation of this binding domain which in turn switching off the BRCA1 mediated DNA damage response. Hence, we proposed that the interaction between RACGAP1 and ECT2 via the BRCT domains may be essential for maintaining BRCA1 function in DNA damage response pathway.

4.2.5 Validation of perturbed molecules in response to RACGAP1 suppression by western blot and LC-MS/MS analysis

The perturbed mRNA level of most of the molecules, upon siRNA-mediated suppression of RACGAP1 as captured in microarray-based profiling, correlated well with its protein expression when detected individually by western blot, except that PLK1 protein expression remained unaltered, though its overexpression has been implicated in HCC (231-233). In contrast, silencing of RACGAP1 decreased PLK4 expression, which was suggested to play a tumor suppressor role in HCC (232). This technique has always been laborious and time-consuming as only one protein can be detected per blot at a time, thus restricting quantitative analysis and reproducibility.

Hence, the preliminary proteomics analysis which provides a global insight of the protein expression within or across system, coupled with microarray profiling, was performed. The main discrepancy between LC-MS/MS and microarray experiments is that the primary change of RACGAP1 was not detected and of all the 8 consistent differential proteins dysregulated both in mRNA and protein level, their association with RACGAP1 have not been reported before, thus could not link to the most significantly perturbed pathway or networks generated from microarray analysis. A number of studies using the same proteomics approaches have implicated the overexpression of HSPA5 (heat shock 70kDa protein 5) or commonly known as GRP 78 (glucose-regulated protein, 78kDa) in HCC (178, 234-237), associated non-specifically either in HCV or HBV-related HCC (176). However, the knockdown of

RACGAP1 did not seem to reduce, but instead elevated HSPA5 expression, the mechanism of which remains unclear in this study. The overexpression of ANXA1 (annexin A1) has also been implicated in HCC (177, 238), and in transgenic mice during liver regeneration and transformation (239). Recent study employing MALDI-TOF-MS/MS has identified UGDH (UDP-glucose dehydrogenase) and aldehyde dehydrogenase 1 as part of the aberrantly α 1,6-fucosylated glycoproteins which are related to HCC metastasis (240).

Despite the inconsistency of the proteomic data in this study and its little impact on the overall findings, it clearly offers new approach that integrates phenotypic, genome wide and proteomic analysis in the hope to combine strengths from individual approach to provide robust cross validation so as to uncover new direction for future and further investigation. Further efforts are needed to fine-tune this technique to be equivalent mature as the microarray analysis. These include more protein coverage as the initial amount is small as compared to microarray experiments, therefore leading to small overlap consistent genes (8 genes) between the two platforms. More sophisticated algorithms and software for automated protein identification, quantitation and detection of post-translational modifications are also needed especially phosphoprotein population as RACGAP1 and some of its neighbourhood molecules are phosphoproteins.

CHAPTER 5 Conclusion and future perspectives

In conclusion, the present study describes an effective strategy, exploiting combined conventional clinicopathological judgement and microarray approach, to identify HCC recurrence with an overall accuracy above 80% for a cohort of HBV-predominant HCC patients (77.5% HBV⁺, 3.7% HCV⁺, and 18.8% non-HBV and non-HCV). The two recurrence-related prediction gene sets, especially the 57-member gene set would serve as lead targets for the development of novel strategy to greatly improve the clinical management of HCC patients either with invasion or cirrhosis, and carry uncertain risk of disease recurrence following curative partial hepatectomy. The comparative analysis of our gene sets and the earlier similar studies generate lists of robust, reliable and reproducible candidate genes despite differences in patient populations and technology platforms. Among the genes that were differentially expressed between subgroups of HCC in association with recurrence, genes with decreased expression especially large number of those observed between Group 1 and Group 4 can be of importance as they could be aggressive stage or recurrence suppressor genes. This asymmetric distribution of up-regulated and down-regulated genes in HCC in comparison to non-tumor liver tissue or less advanced stage is also reported in the recent studies (128, 192, 241). This indicates the importance of a large number of suppressed genes in HCC and their potential role as therapeutic treatment if the key regulator or pathway is identified and restored.

For up-regulated genes, the localization of those whose products are membrane associated or secreted are of particular interest for their potential as therapeutic targets or as serological markers for detection and diagnosis of recurrence. This will assist to speed up the structural genomic and proteomic researches. As such, the custom made antibody array against the secreted gene products can be developed for serological screening of HCC patients for recurrence. By silencing the master gene with elevated expression using a small interfering RNA approach or by raising antibodies against the products of these genes, their culprit roles for cancer progression can be characterized *in vitro* and perhaps new candidate targets for inhibiting cancer progression can be identified. Vascular invasion has been reported as an important tumor factor, contributing to true metastasis, which usually gives rise to early recurrence. To uncover the specific early recurrence prediction signature, the size of the patients with vascular invasion, from group 2 has to be increased, so that recurrence time could be shortened to 2 years, and specific differential genes can be obtained from comparing recurrence versus non-recurrent samples from group 2 alone. To identify late recurrence-specific signature, the same approach is anticipated in patients with cirrhosis from group 3, except that the recurrence time will have to be extended to beyond 3 years, as cirrhosis is usually associated with de novo primary tumor which contributes to late recurrence.

Collectively, this study presented a remarkable upregulation of RACGAP1 in HCC cases with aggressive tumor behaviour and its expression has distinct prognostic indication for disease recurrence within 2 years after resection. RACGAP1 expression

influences cell behaviour including cell migration, invasion and proliferation implicated in HCC tumor cell survival and metastases by modulation of Rho GTPase molecules such as Rac1 and Cdc42. Having observed the inverse phenotypes after loss of and gain of RACGAP1 function *in vitro*, it will be of interest to investigate if the knock-down and overexpression of RACGAP1, employing lentiviral approach can suppress or induce the metastases of HCC *in vivo*.

We also depict the molecular networks including mitotic polo-like kinase and BRCA1-associated DNA damage pathways that relate aggressive phenotypes contributing to HCC recurrence, as modulated by RACGAP1 expression *in vitro*. Most of the dysregulated pathways or networks after suppression of RACGAP1 are interpreted from the changes in mRNA expression of lists of aberrant genes that are depicted from microarray experiments, and mRNA population is often poor indication of protein level or cellular behavior. Almost half of the dysregulated genes selected for confirmation in protein expression by western blot had perturbation in phosphorylated form, not in the total protein.

Proteomics serves to complement functional genomics to comprehensively elucidate cellular behaviour at the global-system level. Therefore, we employ high-throughput LC-MS/MS experiments in this study to compare protein expression and cross-validate them to mRNA level because proteins are the active agents, which could serve as the most direct inference of cellular function as well as drug-activity targets. An affirmative conclusion from the proteomic analysis could not be drawn in

this study as the data is preliminary and has little impact on the overall findings, due to the discrepancy discussed earlier. Further needs will include more protein coverage as the initial amount is small as compared to microarray experiments, more sophisticated algorithms and software for automated protein identification, detection of post-translational modifications especially phosphoprotein population as well as reliable quantitative measures. Taken together, the combine use of physical, phenotypic and expression data sets between proteomic and genome-wide analysis will become necessary in future research, as this not only allow robust cross-validation, but can also generate new hypotheses and new information that would otherwise not arise from one individual approach.

REFERENCES

1. Ross MH, Romrell LJ, Kaye GI. Histology. A Text and Atlas. Third ed. Baltimore: Williams and Wilkins, 1995.
2. Mascarenhas R, Gurakar A. Liver Transplant Candidates. Comprehensive Transplant Centre. Johns Hopskin Medicine. http://www.hopkinsmedicine.org/transplant/Programs/liver/liver_candidates (accessed April 10, 2010).
3. Standring S. Gray's anatomy. The anatomical basis of clinical practice. Philadelphia: Elsevier Churchill Livingstone, 2005.
4. Bismuth H. Surgical anatomy and anatomical surgery of the liver. World J Surg 1982;6:3-9.
5. Smithuis R. Liver: Segmental Anatomy. <http://www.radiologyassistant.nl/en/> (accessed April 18,2010). the Radiology Assistant 2006.
6. Sawitza I, Kordes C, Reister S, Haussinger D. The niche of stellate cells within rat liver. Hepatology 2009;50:1617-1624.
7. Hendriks HF, Verhoofstad WA, Brouwer A, de Leeuw AM, Knook DL. Perisinusoidal fat-storing cells are the main vitamin A storage sites in rat liver. Exp Cell Res 1985;160:138-149.
8. Zender L, Hutker S, Liedtke C, Tillmann HL, Zender S, Mundt B, Waltemathe M, et al. Caspase 8 small interfering RNA prevents acute liver failure in mice. Proc Natl Acad Sci U S A 2003;100:7797-7802.
9. Mani H, Kleiner DE. Liver biopsy findings in chronic hepatitis B. Hepatology 2009;49:S61-71.
10. Schoemaker MH, Gommans WM, Conde de la Rosa L, Homan M, Klok P, Trautwein C, van Goor H, et al. Resistance of rat hepatocytes against bile acid-induced apoptosis in cholestatic liver injury is due to nuclear factor-kappa B activation. J Hepatol 2003;39:153-161.

11. Gujral JS, Farhood A, Bajt ML, Jaeschke H. Neutrophils aggravate acute liver injury during obstructive cholestasis in bile duct-ligated mice. *Hepatology* 2003;38:355-363.
12. Black D, Bird MA, Samson CM, Lyman S, Lange PA, Schrum LW, Qian T, et al. Primary cirrhotic hepatocytes resist TGFbeta-induced apoptosis through a ROS-dependent mechanism. *J Hepatol* 2004;40:942-951.
13. Friedman SL. Hepatic stellate cells: protean, multifunctional, and enigmatic cells of the liver. *Physiol Rev* 2008;88:125-172.
14. Peng D, Han Y, Ding H, Wei L. Hepatic steatosis in chronic hepatitis B patients is associated with metabolic factors more than viral factors. *J Gastroenterol Hepatol* 2008;23:1082-1088.
15. Walewska-Zielecka B, Madalinski K, Jablonska J, Godzik P, Cielecka-Kuszyk J, Litwinska B. Composition of inflammatory infiltrate and its correlation with HBV/HCV antigen expression. *World J Gastroenterol* 2008;14:4040-4046.
16. Maini MK, Boni C, Lee CK, Larrubia JR, Reignat S, Ogg GS, King AS, et al. The role of virus-specific CD8(+) cells in liver damage and viral control during persistent hepatitis B virus infection. *J Exp Med* 2000;191:1269-1280.
17. Das A, Hoare M, Davies N, Lopes AR, Dunn C, Kennedy PT, Alexander G, et al. Functional skewing of the global CD8 T cell population in chronic hepatitis B virus infection. *J Exp Med* 2008;205:2111-2124.
18. Chang JJ, Thompson AJ, Visvanathan K, Kent SJ, Cameron PU, Wightman F, Desmond P, et al. The phenotype of hepatitis B virus-specific T cells differ in the liver and blood in chronic hepatitis B virus infection. *Hepatology* 2007;46:1332-1340.
19. Romo EM, Munoz-Robles JA, Castillo-Rama M, Meneu JC, Moreno-Elola A, Perez-Saborido B, Mancebo E, et al. Peripheral blood lymphocyte populations in end-stage liver diseases. *J Clin Gastroenterol* 2007;41:713-721.
20. Ganem D, Prince AM. Hepatitis B virus infection--natural history and clinical consequences. *N Engl J Med* 2004;350:1118-1129.
21. Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology* 2004;127:S35-50.

22. Paradis V, Laurendeau I, Vidaud M, Bedossa P. Clonal analysis of macronodules in cirrhosis. *Hepatology* 1998;28:953-958.
23. Anthony PP, Ishak KG, Nayak NC, Poulsen HE, Scheuer PJ, Sobin LH. The morphology of cirrhosis. Recommendations on definition, nomenclature, and classification by a working group sponsored by the World Health Organization. *J Clin Pathol* 1978;31:395-414.
24. Kalai M, Van Loo G, Vanden Berghe T, Meeus A, Burm W, Saelens X, Vandenabeele P. Tipping the balance between necrosis and apoptosis in human and murine cells treated with interferon and dsRNA. *Cell Death Differ* 2002;9:981-994.
25. Vanden Berghe T, Kalai M, van Loo G, Declercq W, Vandenabeele P. Disruption of HSP90 function reverts tumor necrosis factor-induced necrosis to apoptosis. *J Biol Chem* 2003;278:5622-5629.
26. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. *Nat Rev Cancer* 2006;6:674-687.
27. Theise ND, Park YN, Kojiro M. Dysplastic nodules and hepatocarcinogenesis. *Clin Liver Dis* 2002;6:497-512.
28. Anthony PP, Vogel CL, Barker LF. Liver cell dysplasia: a premalignant condition. *J Clin Pathol* 1973;26:217-223.
29. Koo JS, Kim H, Park BK, Ahn SH, Han KH, Chon CY, Park C, et al. Predictive value of liver cell dysplasia for development of hepatocellular carcinoma in patients with chronic hepatitis B. *J Clin Gastroenterol* 2008;42:738-743.
30. Sugitani S, Sakamoto M, Ichida T, Genda T, Asakura H, Hirohashi S. Hyperplastic foci reflect the risk of multicentric development of human hepatocellular carcinoma. *J Hepatol* 1998;28:1045-1053.
31. Borzio M, Fargion S, Borzio F, Fracanzani AL, Croce AM, Stroppolini T, Oldani S, et al. Impact of large regenerative, low grade and high grade dysplastic nodules in hepatocellular carcinoma development. *J Hepatol* 2003;39:208-214.
32. Nakashima Y, Nakashima O, Tanaka M, Okuda K, Nakashima M, Kojiro M. Portal vein invasion and intrahepatic micrometastasis in small hepatocellular carcinoma by gross type. *Hepatol Res* 2003;26:142-147.

33. Okuda K. Hepatocellular carcinoma. *J Hepatol* 2000;32:225-237.
34. Bruix J, Boix L, Sala M, Llovet JM. Focus on hepatocellular carcinoma. *Cancer Cell* 2004;5:215-219.
35. Colombo M. The natural history of hepatocellular carcinoma in Western countries. *Hepatogastroenterology* 1998;45 Suppl 3:1221-1225.
36. Porte RJ CP. Epidemiology, etiology, and natural history of liver tumors. In: Clavien PA. editors. *Malignant liver tumors: current and emerging therapies*. Sudbury, Massachusetts: Jones and Bartlett Publishers, 2004: 27-38.
37. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. *Int J Cancer* 2001;94:153-156.
38. Lee HP, Chew L, Chow KY, Jin AZ, Sandra L, Chia KS, Seow A, et al. Trends in Cancer Incidence in Singapore 2003-2007. Singapore Cancer Registry Report. Interim Report 2009.
39. Parkin DM WS, Ferlay J, Teppo L, Thomas DB. Cancer incidence in five continents. Volume VIII. IARC Scientific Publications No. 155. : IARC Press, 2003.
40. El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. *N Engl J Med* 1999;340:745-750.
41. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin* 2009;59:225-249.
42. Tanaka Y, Hanada K, Mizokami M, Yeo AE, Shih JW, Gojobori T, Alter HJ. Inaugural Article: A comparison of the molecular clock of hepatitis C virus in the United States and Japan predicts that hepatocellular carcinoma incidence in the United States will increase over the next two decades. *Proc Natl Acad Sci U S A* 2002;99:15584-15589.
43. McCracken M, Olsen M, Chen MS, Jr., Jemal A, Thun M, Cokkinides V, Deapen D, et al. Cancer incidence, mortality, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities. *CA Cancer J Clin* 2007;57:190-205.
44. El-Serag HB. Hepatocellular carcinoma: an epidemiologic view. *J Clin Gastroenterol* 2002;35:S72-78.

45. Wang XW, Hussain SP, Huo TI, Wu CG, Forgues M, Hofseth LJ, Brechot C, et al. Molecular pathogenesis of human hepatocellular carcinoma. *Toxicology* 2002;181-182:43-47.
46. Monto A, Wright TL. The epidemiology and prevention of hepatocellular carcinoma. *Semin Oncol* 2001;28:441-449.
47. Chu CM. Natural history of chronic hepatitis B virus infection in adults with emphasis on the occurrence of cirrhosis and hepatocellular carcinoma. *J Gastroenterol Hepatol* 2000;15 Suppl:E25-30.
48. Takano S, Yokosuka O, Imazeki F, Tagawa M, Omata M. Incidence of hepatocellular carcinoma in chronic hepatitis B and C: a prospective study of 251 patients. *Hepatology* 1995;21:650-655.
49. Oikawa T, Ojima H, Yamasaki S, Takayama T, Hirohashi S, Sakamoto M. Multistep and multicentric development of hepatocellular carcinoma: histological analysis of 980 resected nodules. *J Hepatol* 2005;42:225-229.
50. Cantarini MC, Trevisani F, Morselli-Labate AM, Rapaccini G, Farinati F, Del Poggio P, Di Nolfo MA, et al. Effect of the etiology of viral cirrhosis on the survival of patients with hepatocellular carcinoma. *Am J Gastroenterol* 2006;101:91-98.
51. Cacoub P, Geffray L, Rosenthal E, Perronne C, Veyssier P, Raguin G. Mortality among human immunodeficiency virus-infected patients with cirrhosis or hepatocellular carcinoma due to hepatitis C virus in French Departments of Internal Medicine/Infectious Diseases, in 1995 and 1997. *Clin Infect Dis* 2001;32:1207-1214.
52. Garcia-Samaniego J, Rodriguez M, Berenguer J, Rodriguez-Rosado R, Carbo J, Asensi V, Soriano V. Hepatocellular carcinoma in HIV-infected patients with chronic hepatitis C. *Am J Gastroenterol* 2001;96:179-183.
53. Craig J. Tumors of the liver. In: Zakim D, Boyer TD, editors. *Hepatology: a Textbook of Liver Disease*. Philadelphia: Saunders, 2003.
54. Caballeria L, Pares A, Castells A, Gines A, Bru C, Rodes J. Hepatocellular carcinoma in primary biliary cirrhosis: similar incidence to that in hepatitis C virus-related cirrhosis. *Am J Gastroenterol* 2001;96:1160-1163.
55. Ikeda K, Saitoh S, Koida I, Arase Y, Tsubota A, Chayama K, Kumada H, et al. A multivariate analysis of risk factors for hepatocellular carcinogenesis: a

- prospective observation of 795 patients with viral and alcoholic cirrhosis. *Hepatology* 1993;18:47-53.
56. Park SZ, Nagorney DM, Czaja AJ. Hepatocellular carcinoma in autoimmune hepatitis. *Dig Dis Sci* 2000;45:1944-1948.
 57. Riordan SM, Williams R. The Wilson's disease gene and phenotypic diversity. *J Hepatol* 2001;34:165-171.
 58. Starley BQ, Calcagno CJ, Harrison SA. Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection. *Hepatology* 2010;51:1820-1832.
 59. Park EJ, Lee JH, Yu GY, He G, Ali SR, Holzer RG, Osterreicher CH, et al. Dietary and genetic obesity promote liver inflammation and tumorigenesis by enhancing IL-6 and TNF expression. *Cell* 2010;140:197-208.
 60. Farrell GC, Larter CZ. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. *Hepatology* 2006;43:S99-S112.
 61. Colli A, Fraquelli M, Casazza G, Massironi S, Colucci A, Conte D, Duca P. Accuracy of ultrasonography, spiral CT, magnetic resonance, and alpha-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. *Am J Gastroenterol* 2006;101:513-523.
 62. Burrel M, Llovet JM, Ayuso C, Iglesias C, Sala M, Miquel R, Caralt T, et al. MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation. *Hepatology* 2003;38:1034-1042.
 63. Junqueira LC, Cossermelli W, Brentani R. Differential staining of collagens type I, II and III by Sirius Red and polarization microscopy. *Arch Histol Jpn* 1978;41:267-274.
 64. Puchtler H, Waldrop FW. Silver impregnation methods for reticulum fibers and reticulin: a re-investigation of their origins and specificity. *Histochemistry* 1978;57:177-187.
 65. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, Feng ZZ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol* 2002;97:2614-2618.
 66. Brunetti E, Silini E, Pistorio A, Cavallero A, Marangio A, Bruno R, Filice C. Coarse vs. fine needle aspiration biopsy for the assessment of diffuse liver

- disease from hepatitis C virus-related chronic hepatitis. *J Hepatol* 2004;40:501-506.
67. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;38:1449-1457.
 68. Colloredo G, Guido M, Sonzogni A, Leandro G. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. *J Hepatol* 2003;39:239-244.
 69. Guido M, Rugge M. Liver biopsy sampling in chronic viral hepatitis. *Semin Liver Dis* 2004;24:89-97.
 70. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, Sherman M, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. *J Natl Cancer Inst* 2008;100:698-711.
 71. Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. *Lancet* 2003;362:1907-1917.
 72. Llovet JM, Schwartz M, Mazzaferro V. Resection and liver transplantation for hepatocellular carcinoma. *Semin Liver Dis* 2005;25:181-200.
 73. Llovet JM, Bruix J. Novel advancements in the management of hepatocellular carcinoma in 2008. *J Hepatol* 2008;48 Suppl 1:S20-37.
 74. Bismuth H, Majno PE. Hepatobiliary surgery. *J Hepatol* 2000;32:208-224.
 75. Lau WY, Leung TW, Ho SK, Chan M, Machin D, Lau J, Chan AT, et al. Adjuvant intra-arterial iodine-131-labelled lipiodol for resectable hepatocellular carcinoma: a prospective randomised trial. *Lancet* 1999;353:797-801.
 76. Takayama T, Sekine T, Makuuchi M, Yamasaki S, Kosuge T, Yamamoto J, Shimada K, et al. Adoptive immunotherapy to lower postsurgical recurrence rates of hepatocellular carcinoma: a randomised trial. *Lancet* 2000;356:802-807.
 77. Muto Y, Moriwaki H, Ninomiya M, Adachi S, Saito A, Takasaki KT, Tanaka T, et al. Prevention of second primary tumors by an acyclic retinoid, polyprenoic acid, in patients with hepatocellular carcinoma. Hepatoma Prevention Study Group. *N Engl J Med* 1996;334:1561-1567.

78. Mazzaferro V, Romito R, Schiavo M, Mariani L, Camerini T, Bhoori S, Capussotti L, et al. Prevention of hepatocellular carcinoma recurrence with alpha-interferon after liver resection in HCV cirrhosis. *Hepatology* 2006;44:1543-1554.
79. Poon RT, Fan ST, Ng IO, Lo CM, Liu CL, Wong J. Different risk factors and prognosis for early and late intrahepatic recurrence after resection of hepatocellular carcinoma. *Cancer* 2000;89:500-507.
80. Imamura H, Matsuyama Y, Tanaka E, Ohkubo T, Hasegawa K, Miyagawa S, Sugawara Y, et al. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *J Hepatol* 2003;38:200-207.
81. Llovet JM, Fuster J, Bruix J. Intention-to-treat analysis of surgical treatment for early hepatocellular carcinoma: resection versus transplantation. *Hepatology* 1999;30:1434-1440.
82. Wu JC, Huang YH, Chau GY, Su CW, Lai CR, Lee PC, Huo TI, et al. Risk factors for early and late recurrence in hepatitis B-related hepatocellular carcinoma. *J Hepatol* 2009;51:890-897.
83. Bismuth H MP. Hepatobiliary surgery. . *J Hepatology* 2000;32:208-224.
84. Llovet JM BA, Bruix J. . Hepatocellular carcinoma. *Lancet* 2003;362:1907-1917.
85. Yamamoto J KT, Takayama T, Shimada K, Yamasaki S, Ozaki H, Yamaguchi N, Makuuchi M. Recurrence of hepatocellular carcinoma after surgery. *Br J Surg* 1996;83:1219-1222.
86. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 2007;132:2557-2576.
87. Imamura H MY, Tanaka E, Ohkubo T, Hasegawa K, Miyagawa S, Sugawara Y, Minagawa M, Takayama T, Kawasaki S, Makuuchi M. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *J Hepatology* 2003;38:200-207.
88. Llovet JM SM, Mazzaferro V. Resection and liver transplantation for hepatocellular carcinoma. *Semin Liver Dis* 2005;25:181-200.

89. Hung IF, Poon RT, Lai CL, Fung J, Fan ST, Yuen MF. Recurrence of hepatitis B-related hepatocellular carcinoma is associated with high viral load at the time of resection. *Am J Gastroenterol* 2008;103:1663-1673.
90. Kim BK, Park JY, Kim do Y, Kim JK, Kim KS, Choi JS, Moon BS, et al. Persistent hepatitis B viral replication affects recurrence of hepatocellular carcinoma after curative resection. *Liver Int* 2008;28:393-401.
91. Kubo S, Hirohashi K, Tanaka H, Tsukamoto T, Shuto T, Yamamoto T, Ikebe T, et al. Effect of viral status on recurrence after liver resection for patients with hepatitis B virus-related hepatocellular carcinoma. *Cancer* 2000;88:1016-1024.
92. Huang Y, Wang Z, An S, Zhou B, Zhou Y, Chan HL, Hou J. Role of hepatitis B virus genotypes and quantitative HBV DNA in metastasis and recurrence of hepatocellular carcinoma. *J Med Virol* 2008;80:591-597.
93. Jang JW, Choi JY, Bae SH, Yoon SK, Woo HY, Chang UI, Kim CW, et al. The impact of hepatitis B viral load on recurrence after complete necrosis in patients with hepatocellular carcinoma who receive transarterial chemolipiodolization: implications for viral suppression to reduce the risk of cancer recurrence. *Cancer* 2007;110:1760-1767.
94. Iizuka N, Hamamoto Y, Oka M. Predicting individual outcomes in hepatocellular carcinoma. *Lancet* 2004;364:1837-1839.
95. Llovet JM, Wurmback E. Gene expression profiles in hepatocellular carcinoma: not yet there. *J Hepatol* 2004;41:336-339.
96. Amitrano L, Guardascione MA, Brancaccio V, Margaglione M, Manguso F, Iannaccone L, Grandone E, et al. Risk factors and clinical presentation of portal vein thrombosis in patients with liver cirrhosis. *J Hepatol* 2004;40:736-741.
97. Di Bisceglie AM. Hepatitis B and hepatocellular carcinoma. *Hepatology* 2009;49:S56-60.
98. Galli A, Svegliati-Baroni G, Ceni E, Milani S, Ridolfi F, Salzano R, Tarocchi M, et al. Oxidative stress stimulates proliferation and invasiveness of hepatic stellate cells via a MMP2-mediated mechanism. *Hepatology* 2005;41:1074-1084.

99. Iloeje UH, Yang HI, Su J, Jen CL, You SL, Chen CJ. Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. *Gastroenterology* 2006;130:678-686.
100. Yuan HJ, Yuen MF, Ka-Ho Wong D, Sablon E, Lai CL. The relationship between HBV-DNA levels and cirrhosis-related complications in Chinese with chronic hepatitis B. *J Viral Hepat* 2005;12:373-379.
101. Chen G, Lin W, Shen F, Iloeje UH, London WT, Evans AA. Past HBV viral load as predictor of mortality and morbidity from HCC and chronic liver disease in a prospective study. *Am J Gastroenterol* 2006;101:1797-1803.
102. Fattovich G, Olivari N, Pasino M, D'Onofrio M, Martone E, Donato F. Long-term outcome of chronic hepatitis B in Caucasian patients: mortality after 25 years. *Gut* 2008;57:84-90.
103. Iloeje UH, Yang HI, Jen CL, Su J, Wang LY, You SL, Chen CJ. Risk and predictors of mortality associated with chronic hepatitis B infection. *Clin Gastroenterol Hepatol* 2007;5:921-931.
104. Ohkubo K, Kato Y, Ichikawa T, Kajiya Y, Takeda Y, Higashi S, Hamasaki K, et al. Viral load is a significant prognostic factor for hepatitis B virus-associated hepatocellular carcinoma. *Cancer* 2002;94:2663-2668.
105. Yeo W, Mo FK, Chan SL, Leung NW, Hui P, Lam WY, Mok TS, et al. Hepatitis B viral load predicts survival of HCC patients undergoing systemic chemotherapy. *Hepatology* 2007;45:1382-1389.
106. Chen CJ, Yu MW, Liaw YF. Epidemiological characteristics and risk factors of hepatocellular carcinoma. *J Gastroenterol Hepatol* 1997;12:S294-308.
107. Chen CJ, Chen DS. Interaction of hepatitis B virus, chemical carcinogen, and genetic susceptibility: multistage hepatocarcinogenesis with multifactorial etiology. *Hepatology* 2002;36:1046-1049.
108. Calvet X, Bruix J, Gines P, Bru C, Sole M, Vilana R, Rodes J. Prognostic factors of hepatocellular carcinoma in the west: a multivariate analysis in 206 patients. *Hepatology* 1990;12:753-760.
109. Adachi E, Maeda T, Matsumata T, Shirabe K, Kinukawa N, Sugimachi K, Tsuneyoshi M. Risk factors for intrahepatic recurrence in human small hepatocellular carcinoma. *Gastroenterology* 1995;108:768-775.

110. Smalley SR, Moertel CG, Hilton JF, Weiland LH, Weiland HS, Adson MA, Melton LJ, 3rd, et al. Hepatoma in the noncirrhotic liver. *Cancer* 1988;62:1414-1424.
111. Yu MC, Yuan JM, Govindarajan S, Ross RK. Epidemiology of hepatocellular carcinoma. *Can J Gastroenterol* 2000;14:703-709.
112. Bralet MP, Regimbeau JM, Pineau P, Dubois S, Loas G, Degos F, Valla D, et al. Hepatocellular carcinoma occurring in nonfibrotic liver: epidemiologic and histopathologic analysis of 80 French cases. *Hepatology* 2000;32:200-204.
113. Trevisani F, D'Intino PE, Caraceni P, Pizzo M, Stefanini GF, Mazziotti A, Grazi GL, et al. Etiologic factors and clinical presentation of hepatocellular carcinoma. Differences between cirrhotic and noncirrhotic Italian patients. *Cancer* 1995;75:2220-2232.
114. Condeelis J, Segall JE. Intravital imaging of cell movement in tumours. *Nat Rev Cancer* 2003;3:921-930.
115. Johnson PJ. Hepatocellular Carcinoma. In: Gospadarowicz MK, Henson DE, Hutter RVP, Sullivan BO, Sobin LH, Wittekind Ch, editors. *Prognostic factors in cancer*. New York: John Wiley & Sons, 2001.
116. Chevret S, Trinchet JC, Mathieu D, Rached AA, Beaugrand M, Chastang C. A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. Groupe d'Etude et de Traitement du Carcinome Hepatocellulaire. *J Hepatol* 1999;31:133-141.
117. 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.
118. Romani F, Belli LS, Rondinara GF, DeCarlis L, Rimoldi P, Riolo F, Bellati G, et al. The role of transplantation in small hepatocellular carcinoma complicating cirrhosis of the liver. *J Am Coll Surg* 1994;178:379-384.
119. McPeake JR, O'Grady JG, Zaman S, Portmann B, Wight DG, Tan KC, Calne RY, et al. Liver transplantation for primary hepatocellular carcinoma: tumor size and number determine outcome. *J Hepatol* 1993;18:226-234.
120. Gress TM, Hoheisel JD, Lennon GG, Zehetner G, Lehrach H. Hybridization fingerprinting of high-density cDNA-library arrays with cDNA pools derived from whole tissues. *Mamm Genome* 1992;3:609-619.

121. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270:467-470.
122. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet* 1999;21:20-24.
123. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 2002;1:133-143.
124. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-511.
125. Jones MH, Virtanen C, Honjoh D, Miyoshi T, Satoh Y, Okumura S, Nakagawa K, et al. Two prognostically significant subtypes of high-grade lung neuroendocrine tumours independent of small-cell and large-cell neuroendocrine carcinomas identified by gene expression profiles. *Lancet* 2004;363:775-781.
126. Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, Bild A, et al. Gene expression predictors of breast cancer outcomes. *Lancet* 2003;361:1590-1596.
127. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, et al. Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001;412:822-826.
128. Cheung ST, Chen X, Guan XY, Wong SY, Tai LS, Ng IO, So S, et al. Identify metastasis-associated genes in hepatocellular carcinoma through clonality delineation for multinodular tumor. *Cancer Res* 2002;62:4711-4721.
129. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, Gupta S, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* 2008;359:1995-2004.
130. Brown MP, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, Ares M, Jr., et al. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc Natl Acad Sci U S A* 2000;97:262-267.

131. Toure A, Dorseuil O, Morin L, Timmons P, Jegou B, Reibel L, Gacon G. MgcRacGAP, a new human GTPase-activating protein for Rac and Cdc42 similar to *Drosophila* rotundRacGAP gene product, is expressed in male germ cells. *J Biol Chem* 1998;273:6019-6023.
132. Glotzer M. The 3Ms of central spindle assembly: microtubules, motors and MAPs. *Nat Rev Mol Cell Biol* 2009;10:9-20.
133. Wolfe BA, Takaki T, Petronczki M, Glotzer M. Polo-like kinase 1 directs assembly of the HsCyk-4 RhoGAP/Ect2 RhoGEF complex to initiate cleavage furrow formation. *PLoS Biol* 2009;7:e1000110.
134. Jantsch-Plunger V, Gonczy P, Romano A, Schnabel H, Hamill D, Schnabel R, Hyman AA, et al. CYK-4: A Rho family gtpase activating protein (GAP) required for central spindle formation and cytokinesis. *J Cell Biol* 2000;149:1391-1404.
135. Goldstein AY, Jan YN, Luo L. Function and regulation of Tumbleweed (RacGAP50C) in neuroblast proliferation and neuronal morphogenesis. *Proc Natl Acad Sci U S A* 2005;102:3834-3839.
136. Hirose K, Kawashima T, Iwamoto I, Nosaka T, Kitamura T. MgcRacGAP is involved in cytokinesis through associating with mitotic spindle and midbody. *J Biol Chem* 2001;276:5821-5828.
137. Kawashima T, Hirose K, Satoh T, Kaneko A, Ikeda Y, Kaziro Y, Nosaka T, et al. MgcRacGAP is involved in the control of growth and differentiation of hematopoietic cells. *Blood* 2000;96:2116-2124.
138. Canman JC, Lewellyn L, Laband K, Smerdon SJ, Desai A, Bowerman B, Oegema K. Inhibition of Rac by the GAP activity of centralspindlin is essential for cytokinesis. *Science* 2008;322:1543-1546.
139. D'Avino PP, Savoian MS, Glover DM. Mutations in sticky lead to defective organization of the contractile ring during cytokinesis and are enhanced by Rho and suppressed by Rac. *J Cell Biol* 2004;166:61-71.
140. Miller AL, Bement WM. Regulation of cytokinesis by Rho GTPase flux. *Nat Cell Biol* 2009;11:71-77.
141. Glotzer M. The molecular requirements for cytokinesis. *Science* 2005;307:1735-1739.

142. Neumann B, Walter T, Heriche JK, Bulkescher J, Erfle H, Conrad C, Rogers P, et al. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* 2010;464:721-727.
143. Mishima M KS, Glotzer M. Central spindle assembly and cytokinesis require a kinesin-like protein/RhoGAP complex with microtubule bundling activity. *Dev Cell* 2002;2:41-54.
144. Yuce O, Piekny A, Glotzer M. An ECT2-centralspindlin complex regulates the localization and function of RhoA. *J Cell Biol* 2005;170:571-582.
145. Somers WG SR. A RhoGEF and Rho family GTPase-activating protein complex links the contractile ring to cortical microtubules at the onset of cytokinesis. *Dev Cell* 2003;4:29-39.
146. Mishima M, Pavicic V, Gruneberg U, Nigg EA, Glotzer M. Cell cycle regulation of central spindle assembly. *Nature* 2004;430:908-913.
147. Oceguera-Yanez F, Kimura K, Yasuda S, Higashida C, Kitamura T, Hiraoka Y, Haraguchi T, et al. Ect2 and MgcRacGAP regulate the activation and function of Cdc42 in mitosis. *J Cell Biol* 2005;168:221-232.
148. Zhao WM, Fang G. MgcRacGAP controls the assembly of the contractile ring and the initiation of cytokinesis. *Proc Natl Acad Sci U S A* 2005;102:13158-13163.
149. Kamijo K, Ohara N, Abe M, Uchimura T, Hosoya H, Lee JS, Miki T. Dissecting the role of Rho-mediated signaling in contractile ring formation. *Mol Biol Cell* 2006;17:43-55.
150. Burkard ME, Maciejowski J, Rodriguez-Bravo V, Repka M, Lowery DM, Clauser KR, Zhang C, et al. Plk1 self-organization and priming phosphorylation of HsCYK-4 at the spindle midzone regulate the onset of division in human cells. *PLoS Biol* 2009;7:1-16.
151. Gregory SL ES, Milverton J, Jones WM, Bejsovec A, Saint R. Cell division requires a direct link between microtubule-bound RacGAP and Anillin in the contractile ring. *Curr Biol* 2008;18:25-29.
152. Toure A, Mzali R, Liot C, Seguin L, Morin L, Crouin C, Chen-Yang I, et al. Phosphoregulation of MgcRacGAP in mitosis involves Aurora B and Cdk1 protein kinases and the PP2A phosphatase. *FEBS Lett* 2008;582:1182-1188.

153. Zhao WM, Seki A, Fang G. Cep55, a microtubule-bundling protein, associates with centralspindlin to control the midbody integrity and cell abscission during cytokinesis. *Mol Biol Cell* 2006;17:3881-3896.
154. Wang SM, Ooi LL, Hui KM. Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. *Clin Cancer Res* 2007;13:6275-6283.
155. Lu KH, Patterson AP, Wang L, Marquez RT, Atkinson EN, Baggerly KA, Ramoth LR, et al. Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis. *Clin Cancer Res* 2004;10:3291-3300.
156. Ma XJ SR, Tuggle JT, Gaudet J, Enright E, McQuary P, Payette T, Pistone M, Stecker K, Zhang BM, Zhou YX, Varnholt H, Smith B, Gadd M, Chatfield E, Kessler J, Baer TM, Erlander MG, Sgroi DC. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A* 2003;100:5974-5979.
157. Rosty C SM, Tsafrir D, Stransky N, Tsafrir I, Peter M, de Crémoux P, de La Rochefordière A, Salmon R, Dorval T, Thierry JP, Couturier J, Radvanyi F, Domany E, Sastre-Garau X. Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. *Oncogene* 2005;24:7094-7104.
158. Teschendorff AE NA, Barbosa-Morais NL, Pinder SE, Ellis IO, Aparicio S, Brenton JD, Caldas C. A consensus prognostic gene expression classifier for ER positive breast cancer. *Genome Biol* 2006;7:R101.
159. Yu JX SA, Zhang Y, Martens JW, Smid M, Klijn JG, Wang Y, Foekens JA. Pathway analysis of gene signatures predicting metastasis of node-negative primary breast cancer. *BMC Cancer* 2007;7:182.
160. Nilsson R PJ, Björkegren J, Tegnér J. Detecting multivariate differentially expressed genes. *BMC Bioinformatics* 2007;8:150.
161. Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, Heuer D, Becker D, et al. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 2010;463:818-822.
162. König R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, Alamares JG, et al. Human host factors required for influenza virus replication. *Nature* 2010;463:813-817.

163. Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez JC, Yan JX, et al. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* 1996;14:61-65.
164. Tyers M, Mann M. From genomics to proteomics. *Nature* 2003;422:193-197.
165. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;422:198-207.
166. Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem* 2001;70:437-473.
167. Kusmierz JJ, Sumrada R, Desiderio DM. Fast atom bombardment mass spectrometric quantitative analysis of methionine-enkephalin in human pituitary tissues. *Anal Chem* 1990;62:2395-2400.
168. Mirgorodskaya OA, Kozmin YP, Titov MI, Korner R, Sonksen CP, Roepstorff P. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun Mass Spectrom* 2000;14:1226-1232.
169. Yao M, Srinivas NR. Simultaneous quantitation of d7-nefazodone, nefazodone, d7-hydroxynefazodone, hydroxynefazodone, m-chlorophenylpiperazine and triazole-dione in human plasma by liquid chromatographic-mass spectrometry. *Biomed Chromatogr* 2000;14:106-112.
170. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;1:376-386.
171. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004;3:1154-1169.
172. Washburn MP, Wolters D, Yates JR, 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 2001;19:242-247.

173. Lipton MS, Pasa-Tolic L, Anderson GA, Anderson DJ, Auberry DL, Battista JR, Daly MJ, et al. Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags. *Proc Natl Acad Sci U S A* 2002;99:11049-11054.
174. Hanash S. Integrated global profiling of cancer. *Nat Rev Cancer* 2004;4:638-644.
175. Yi X, Luk JM, Lee NP, Peng J, Leng X, Guan XY, Lau GK, et al. Association of mortalin (HSPA9) with liver cancer metastasis and prediction for early tumor recurrence. *Mol Cell Proteomics* 2008;7:315-325.
176. Sun W, Xing B, Sun Y, Du X, Lu M, Hao C, Lu Z, et al. Proteome analysis of hepatocellular carcinoma by two-dimensional difference gel electrophoresis: novel protein markers in hepatocellular carcinoma tissues. *Mol Cell Proteomics* 2007;6:1798-1808.
177. Liang RC, Neo JC, Lo SL, Tan GS, Seow TK, Chung MC. Proteome database of hepatocellular carcinoma. *J Chromatogr B Analyt Technol Biomed Life Sci* 2002;771:303-328.
178. Luk JM, Lam CT, Siu AF, Lam BY, Ng IO, Hu MY, Che CM, et al. Proteomic profiling of hepatocellular carcinoma in Chinese cohort reveals heat-shock proteins (Hsp27, Hsp70, GRP78) up-regulation and their associated prognostic values. *Proteomics* 2006;6:1049-1057.
179. Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, Roberts JR, et al. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 2003;362:433-439.
180. Stanley BA, Gundry RL, Cotter RJ, Van Eyk JE. Heart disease, clinical proteomics and mass spectrometry. *Dis Markers* 2004;20:167-178.
181. Griffin TJ, Gygi SP, Ideker T, Rist B, Eng J, Hood L, Aebersold R. Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol Cell Proteomics* 2002;1:323-333.
182. Katz-Jaffe MG, Schoolcraft WB, Gardner DK. Analysis of protein expression (secretome) by human and mouse preimplantation embryos. *Fertil Steril* 2006;86:678-685.

183. Shiio Y, Donohoe S, Yi EC, Goodlett DR, Aebersold R, Eisenman RN. Quantitative proteomic analysis of Myc oncoprotein function. *EMBO J* 2002;21:5088-5096.
184. Tan TL, Chen WN. A proteomics analysis of cellular proteins associated with HBV genotype-specific HBX: potential in identification of early diagnostic markers for HCC. *J Clin Virol* 2005;33:293-298.
185. Goufman EI, Moshkovskii SA, Tikhonova OV, Lokhov PG, Zgoda VG, Serebryakova MV, Toropygin IY, et al. Two-dimensional electrophoretic proteome study of serum thermostable fraction from patients with various tumor conditions. *Biochemistry (Mosc)* 2006;71:354-360.
186. Huang TC, Chang HY, Hsu CH, Kuo WH, Chang KJ, Juan HF. Targeting therapy for breast carcinoma by ATP synthase inhibitor aurovertin B. *J Proteome Res* 2008;7:1433-1444.
187. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95:14863-14868.
188. Benoit RM, Wilhelm RN, Scherer-Becker D, Ostermeier C. An improved method for fast, robust, and seamless integration of DNA fragments into multiple plasmids. *Protein Expr Purif* 2006;45:66-71.
189. Liu ET, Karuturi KR. Microarrays and clinical investigations. *N Engl J Med* 2004;350:1595-1597.
190. Ntzani EE, Ioannidis JP. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. *Lancet* 2003;362:1439-1444.
191. Xu XR, Huang J, Xu ZG, Qian BZ, Zhu ZD, Yan Q, Cai T, et al. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. *Proc Natl Acad Sci U S A* 2001;98:15089-15094.
192. Kim JW, Ye Q, Forgues M, Chen Y, Budhu A, Sime J, Hofseth LJ, et al. Cancer-associated molecular signature in the tissue samples of patients with cirrhosis. *Hepatology* 2004;39:518-527.
193. Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, Tsunoda T, 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.
194. Zindy P, Andrieux L, Bonnier D, Musso O, Langouet S, Campion JP, Turlin B, et al. Upregulation of DNA repair genes in active cirrhosis associated with hepatocellular carcinoma. *FEBS Lett* 2005;579:95-99.
195. Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, Lai KM, et al. Gene expression patterns in human liver cancers. *Mol Biol Cell* 2002;13:1929-1939.
196. Chuma M, Sakamoto M, Yamazaki K, Ohta T, Ohki M, Asaka M, Hirohashi S. Expression profiling in multistage hepatocarcinogenesis: identification of HSP70 as a molecular marker of early hepatocellular carcinoma. *Hepatology* 2003;37:198-207.
197. Thomas MB, Zhu AX. Hepatocellular carcinoma: the need for progress. *J Clin Oncol* 2005;23:2892-2899.
198. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, Durnez A, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatology* 2004;40:667-676.
199. Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC, Simon R, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nat Med* 2003;9:416-423.
200. Iizuka N, Oka M, Yamada-Okabe H, Nishida M, Maeda Y, Mori N, Takao T, et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. *Lancet* 2003;361:923-929.
201. Kurokawa Y, Matoba R, Takemasa I, Nagano H, Dono K, Nakamori S, Umeshita K, et al. Molecular-based prediction of early recurrence in hepatocellular carcinoma. *J Hepatol* 2004;41:284-291.
202. Hung AY, Sheng M. PDZ domains: structural modules for protein complex assembly. *J Biol Chem* 2002;277:5699-5702.
203. Bitner-Glindzicz M, Lindley KJ, Rutland P, Blaydon D, Smith VV, Milla PJ, Hussain K, et al. A recessive contiguous gene deletion causing infantile hyperinsulinism, enteropathy and deafness identifies the Usher type 1C gene. *Nat Genet* 2000;26:56-60.

204. Caceres M, Guerrero J, Martinez J. Overexpression of RhoA-GTP induces activation of the Epidermal Growth Factor Receptor, dephosphorylation of focal adhesion kinase and increased motility in breast cancer cells. *Exp Cell Res* 2005;309:229-238.
205. Cong F, Schweizer L, Varmus H. Casein kinase Iepsilon modulates the signaling specificities of dishevelled. *Mol Cell Biol* 2004;24:2000-2011.
206. Kunkel M, Reichert TE, Benz P, Lehr HA, Jeong JH, Wieand S, Bartenstein P, et al. Overexpression of Glut-1 and increased glucose metabolism in tumors are associated with a poor prognosis in patients with oral squamous cell carcinoma. *Cancer* 2003;97:1015-1024.
207. Walters KA, Syder AJ, Lederer SL, Diamond DL, Paeper B, Rice CM, Katze MG. Genomic analysis reveals a potential role for cell cycle perturbation in HCV-mediated apoptosis of cultured hepatocytes. *PLoS Pathog* 2009;5:e1000269.
208. Kinoshita M, Miyata M. Underexpression of mRNA in human hepatocellular carcinoma focusing on eight loci. *Hepatology* 2002;36:433-438.
209. Lip GY, Chin BS, Blann AD. Cancer and the prothrombotic state. *Lancet Oncol* 2002;3:27-34.
210. Tang ZY, Sun FX, Tian J, Ye SL, Liu YK, Liu KD, Xue Q, et al. Metastatic human hepatocellular carcinoma models in nude mice and cell line with metastatic potential. *World J Gastroenterol* 2001;7:597-601.
211. Li Y, Tang ZY, Ye SL, Liu YK, Chen J, Xue Q, Gao DM, et al. Establishment of cell clones with different metastatic potential from the metastatic hepatocellular carcinoma cell line MHCC97. *World J Gastroenterol* 2001;7:630-636.
212. Li Y, Tang Y, Ye L, Liu B, Liu K, Chen J, Xue Q. Establishment of a hepatocellular carcinoma cell line with unique metastatic characteristics through in vivo selection and screening for metastasis-related genes through cDNA microarray. *J Cancer Res Clin Oncol* 2003;129:43-51.
213. Rosty C, Sheffer M, Tsafrir D, Stransky N, Tsafrir I, Peter M, de Cremoux P, et al. Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. *Oncogene* 2005;24:7094-7104.

214. Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P, Payette T, et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A* 2003;100:5974-5979.
215. Sakon M UK, Nagano H, Eguchi H, Kishimoto S, Miyamoto A, Ohshima S, Dono K, Nakamori S, Gotoh M, Monden M. Clinical significance of hepatic resection in hepatocellular carcinoma: analysis by disease-free survival curves. *Arch Surg* 2000;135:1456-1459.
216. Yoshioka S TI, Nagano H, Kittaka N, Noda T, Wada H, Kobayashi S, Marubashi S, Takeda Y, Umeshita K, Dono K, Matsubara K, Monden M. Molecular prediction of early recurrence after resection of hepatocellular carcinoma. *Eur J Cancer* 2009;45:881-889.
217. Dephoure N, Zhou C, Villen J, Beausoleil SA, Bakalarski CE, Elledge SJ, Gygi SP. A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A* 2008;105:10762-10767.
218. Raftopoulou M HA. Cell migration: Rho GTPases lead the way. *Dev Biol* 2004;265:23-32.
219. Bustelo XR SV, Berenjeno IM. GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. *Bioessays* 2007;29:356-370.
220. Ching YP, Leong VY, Lee MF, Xu HT, Jin DY, Ng IO. P21-activated protein kinase is overexpressed in hepatocellular carcinoma and enhances cancer metastasis involving c-Jun NH2-terminal kinase activation and paxillin phosphorylation. *Cancer Res* 2007;67:3601-3608.
221. Bokoch GM. Biology of the p21-activated kinases. *Annu Rev Biochem* 2003;72:743-781.
222. Budihardjo I, Oliver H, Lutter M, Luo X, Wang X. Biochemical pathways of caspase activation during apoptosis. *Annu Rev Cell Dev Biol* 1999;15:269-290.
223. Cullen SP, Martin SJ. Caspase activation pathways: some recent progress. *Cell Death Differ* 2009;16:935-938.
224. Rich T, Allen RL, Wyllie AH. Defying death after DNA damage. *Nature* 2000;407:777-783.

225. Inoue S, Browne G, Melino G, Cohen GM. Ordering of caspases in cells undergoing apoptosis by the intrinsic pathway. *Cell Death Differ* 2009;16:1053-1061.
226. Kittler R PL, Heninger AK, Slabicki M, Theis M, Miroslaw L, Poser I, Lawo S, Grabner H, Kozak K, Wagner J, Surendranath V, Richter C, Bowen W, Jackson AL, Habermann B, Hyman AA, Buchholz F. Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat Cell Biol* 2007;9:1401-1412.
227. Roberts AB WL. The two faces of transforming growth factor beta in carcinogenesis. *Proc Natl Acad Sci U S A* 2003;100:8621-8623.
228. Haq R BJ, Takahashi M, Finan D, Finkielstein A, Damaraju S, Rottapel R, Zanke B. Constitutive p38HOG mitogen-activated protein kinase activation induces permanent cell cycle arrest and senescence. *Cancer Res* 2002;62:5076-5082.
229. Qi X TJ, Loesch M, Pohl N, Alkan S, Chen G. p38gamma mitogen-activated protein kinase integrates signaling crosstalk between Ras and estrogen receptor to increase breast cancer invasion. *Cancer Res* 2006;66:7540-7547.
230. Behren A MS, Acuna Sanhueza GA, Schwager C, Plinkert PK, Huber PE, Abdollahi A, Simon C. Phenotype-assisted transcriptome analysis identifies FOXM1 downstream from Ras-MKK3-p38 to regulate in vitro cellular invasion. *Oncogene* 2009;1-12.
231. Bassermann F, Frescas D, Guardavaccaro D, Busino L, Peschiaroli A, Pagano M. The Cdc14B-Cdh1-Plk1 axis controls the G2 DNA-damage-response checkpoint. *Cell* 2008;134:256-267.
232. Pellegrino R, Calvisi DF, Ladu S, Ehemann V, Staniscia T, Evert M, Dombrowski F, et al. Oncogenic and tumor suppressive roles of polo-like kinases in human hepatocellular carcinoma. *Hepatology*;51:857-868.
233. Wang XQ, Zhu YQ, Lui KS, Cai Q, Lu P, Poon RT. Aberrant Polo-like kinase 1-Cdc25A pathway in metastatic hepatocellular carcinoma. *Clin Cancer Res* 2008;14:6813-6820.
234. Park KS, Kim H, Kim NG, Cho SY, Choi KH, Seong JK, Paik YK. Proteomic analysis and molecular characterization of tissue ferritin light chain in hepatocellular carcinoma. *Hepatology* 2002;35:1459-1466.

235. Lim SO, Park SG, Yoo JH, Park YM, Kim HJ, Jang KT, Cho JW, et al. Expression of heat shock proteins (HSP27, HSP60, HSP70, HSP90, GRP78, GRP94) in hepatitis B virus-related hepatocellular carcinomas and dysplastic nodules. *World J Gastroenterol* 2005;11:2072-2079.
236. Takashima M, Kuramitsu Y, Yokoyama Y, Iizuka N, Toda T, Sakaida I, Okita K, et al. Proteomic profiling of heat shock protein 70 family members as biomarkers for hepatitis C virus-related hepatocellular carcinoma. *Proteomics* 2003;3:2487-2493.
237. Blanc JF, Lalanne C, Plomion C, Schmitter JM, Bathany K, Gion JM, Bioulac-Sage P, et al. Proteomic analysis of differentially expressed proteins in hepatocellular carcinoma developed in patients with chronic viral hepatitis C. *Proteomics* 2005;5:3778-3789.
238. Masaki T, Tokuda M, Ohnishi M, Watanabe S, Fujimura T, Miyamoto K, Itano T, et al. Enhanced expression of the protein kinase substrate annexin in human hepatocellular carcinoma. *Hepatology* 1996;24:72-81.
239. de Coupade C, Gillet R, Bennoun M, Briand P, Russo-Marie F, Solito E. Annexin 1 expression and phosphorylation are upregulated during liver regeneration and transformation in antithrombin III SV40 T large antigen transgenic mice. *Hepatology* 2000;31:371-380.
240. Dai Z, Liu YK, Cui JF, Shen HL, Chen J, Sun RX, Zhang Y, et al. Identification and analysis of altered alpha1,6-fucosylated glycoproteins associated with hepatocellular carcinoma metastasis. *Proteomics* 2006;6:5857-5867.
241. Kim BY, Lee JG, Park S, Ahn JY, Ju YJ, Chung JH, Han CJ, et al. Feature genes of hepatitis B virus-positive hepatocellular carcinoma, established by its molecular discrimination approach using prediction analysis of microarray. *Biochim Biophys Acta* 2004;1739:50-61.