INTEGRATIVE TRANSCRIPTOMIC ANALYSIS AND VALIDATION OF NON-CODING RNA IN
LIVER CANCERS

Miss Varunya Virushkul

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Medical Biochemistry
Department of Biochemistry
FACULTY OF MEDICINE
Chulalongkorn University
Academic Year 2021
Copyright of Chulalongkorn University

การวิเคราะห์ข้อมูลทรานสคริปโตมิกส์เชิงบูรณาการและการตรวจสอบระดับการแสดงออกของนอนโค้ดดิ้งอาร์เอ็นเอในโรคมะเร็งตับ

น.ส.วรัญญา วิรัชกุล

Thesis Title	INTEGRATIVE TRANSCRIPTOMIC ANALYSIS AND VALIDATION OF NON-CODING RNA IN LIVER CANCERS
By	Miss Varunya Virushkul
Field of Study	Medical Biochemistry
Thesis Advisor	NATTHAYA CHUAYPEN, Ph.D.
Thesis Co Advisor	Professor Pisit Tangkijvanich, M.D.

Accepted by the FACULTY OF MEDICINE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

.... Dean of the FACULTY OF MEDICINE

(Associate Professor CHANCHAI SITTIPUNT, M.D.)

THESIS COMMITTEE

.... Chairman

(Professor Sittisak Honsawek, M.D., Ph.D)

.... Thesis Advisor

(NATTHAYA CHUAYPEN, Ph.D.)

.... Thesis Co-Advisor

(Professor Pisit Tangkijvanich, M.D.)

.... Examiner

(PORNCHAI KAEWSAPSAK, Ph.D.)

.... External Examiner

(Piroon Jenjaroenpun, Ph.D.)

วรัญญา วิรัชกุล : การวิเคราะห์ข้อมูลทรานสคริปโตมิกส์เชิงบูรณาการและการตรวจสอบระดับการแสดงออกของนอนโค้ดดิ้งอาร์เอ็นเอในโรคมะเร็งตับ. (INTEGRATIVE TRANSCRIPTOMIC ANALYSIS AND VALIDATION OF NON-CODING RNA IN LIVER CANCERS) อ.ที่ปรึกษาหลัก : อ. ดร.ณัฐธยาน์ ช่วยเพ็ญ, อ.ที่ปรึกษาร่วม : ศ. นพ.พิสิฐ ตั้งกิจวานิชย์

ในปัจจุบันมะเร็งตับชนิด Hepatocellular carcinoma (HCC) และมะเร็งตับในท่อน้ำดี (intrahepatic cholangiocarcinoma; iCCA) พบบ่อยและก่อให้เกิดอัตราการเสียชีวิตสูง เนื่องจากตัวบ่งชี้ทางชีวภาพที่ใช้ในการวินิจฉัยโรค เช่น AFP, CA19-9 และ CEAยังมีความไวและความจำเพาะไม่เพียงพอ ดังนั้นงานวิจัยนี้จึงมีจุดประสงค์ในการหาตัวบ่งชี้ทางชีวภาพชนิดใหม่สำหรับการวินิจฉัยมะเร็งตับชนิด HCC และ iCCA โดยเริ่มต้นจากการเก็บข้อมูล RNA sequencing และ small RNA sequencing จากฐานข้อมูล (GEO database) ซึ่งเป็นข้อมูลจากชิ้นเนื้อบริเวณที่เป็นมะเร็ง HCC และชิ้นเนื้อใกล้เคียง จำนวน 992 ราย และ ชิ้นเนื้อ iCCA และชิ้นเนื้อใกล้เคียง จำนวน 116 ราย จากนั้นจะวิเคราะห์การแสดงออกที่แตกต่างกันระหว่างชิ้นเนื้อมะเร็งกับชิ้นเนื้อปกติใกล้เคียงอย่างมีนัยสำคัญทางสถิติ จากนั้นจะนำข้อมูลดังกล่าวมาสร้าง เป็น lncRNA-miRNA-mRNA network จากนั้น miRNA ที่มาจาก network จะนำมาเปรียบเทียบกับ miRNA ที่มาจากฐานข้อมูล TCGA และวิเคราะห์ความแตกต่างอย่างมีนัยสำคัญทางสถิติของ miRNA ระหว่างชิ้นเนื้อมะเร็งและชิ้นเนื้อปกติใกล้เคียงอีกครั้ง โดยมีความแตกต่างตั้งแต่ 2 fold change ขึ้นไป และ P-value <0.05 จากนั้นนำ candidate miRNAs ที่ได้มาตรวจสอบเลือดของผู้ป่วย ผลการศึกษาพบว่าระดับของmiR-122-5p, miR-182-5p และ miR-199b-3p ในกลุ่มผู้ป่วย HCC มีการแสดงออกมากกว่าในกลุ่มผู้ป่วย CHB นอกจากนี้ miR-139-3p, miR-148a-3p, miR-221-3p, และ miR-222-3p มีการแสดงออกที่เพิ่มมากขึ้นในกลุ่ม iCCA อย่างมีนัยสำคัญทางสถิติเมื่อเทียบกับกลุ่มคนปกติ จากนั้นวิเคราะห์ประสิทธิภาพของ miRNAs ในการวินิจฉัยแยกระหว่างผู้ป่วยมะเร็งออกจากผู้ป่วยที่ไม่เป็นมะเร็งโดยการวิเคราะห์ ROC curve พบว่า miR-199b-3p, miR-122-5p, miR-182-5p มีพื้นที่ใต้กราฟ (AUC) เท่ากับ 0.67, 0.62 และ 0.54 ตามลำดับสำหรับกลุ่ม HCC และ AUC เท่ากับ 0.99, 0.90, 0.87 และ 0.86 สำหรับ miR-222-3p, miR-221-3p, miR-148a-3p และ miR-139-3p ตามลำดับในกลุ่ม iCCA ดังนั้นจึงสามารถสรุปได้ว่าmiR-122-5p และ miR-199b-3pอาจจะสามารถใช้เป็นตัวบ่งชี้ทางชีวภาพในการวินิจฉัยโรคมะเร็งชนิด HCC ส่วนmiR-139-3p, miR-148a-3p, miR-221-3p, และ miR-222-3p อาจจะสามารถใช้เป็นตัวบ่งชี้สำหรับมะเร็งตับชนิด iCCA ได้

สาขาวิชา	ชีวเคมีทางการแพทย์	ลายมือชื่อนิสิต ....
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม ....

6370101030 : MAJOR MEDICAL BIOCHEMISTRY

KEYWORD: Hepatocellular carcinoma (HCC), Intrahepatic cholangiocarcinoma (iCCA), RNA-sequencing, miRNA, lncRNA-miRNA-mRNA network

Varunya Virushkul : INTEGRATIVE TRANSCRIPTOMIC ANALYSIS AND VALIDATION OF NON-CODING RNA IN LIVER CANCERS . Advisor: NATTHAYA CHUAYPEN, Ph.D. Co-advisor: Prof. Pisit Tangkijvanich, M.D.

Hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (iCCA) are the most common and cause cancer-related deaths. Due to the lack of sensitivity and specificity of conventional markers, such as AFP, CA19-9 and CEA. Thus, this study aimed to identify miRNAs as a diagnostic biomarker for HCC and iCCA. First, in discovery set, RNA sequencing and small RNA sequencing data of 992 HCC and adjacent tissues and 116 iCCA and adjacent tissues were downloaded from the GEO database. After that, the lncRNA-miRNA-mRNA network was constructed and analyzed for the differential expression of lncRNA (DElncRNA), miRNA (DEmiRNA), and mRNA (DEmRNA). Then, the significant DEmiRNAs from the network were compared to miRNA data from the TCGA database. The significant differences in miRNA expression of tumor tissues compared to adjacent tissues from the TCGA database were selected according to the inclusion criteria including fold change and P-value <0.05. In validation set, there were three serum miRNAs in HCC, including miR-122-5p, miR-182-5p, and miR-199b-3p were found to be upregulated in HCC when compared to the CHB group. In addition, there were four serum miRNAs in iCCA, including miR-139-3p, miR-148a-3p, miR-221-3p, and miR-222-3p, were significantly higher in iCCA patients when compared to healthy controls. Moreover, ROC curves of candidate miRNAs were analyzed. The result showed area under the curve (AUC) of 0.67, 0.62 and 0.54 for miR-199b-3p, miR-122-5p, miR-182-5p, respectively. In addition, in iCCA, AUC was 0.99, 0.90, 0.87 and 0.86 for miR-222-3p, miR-221-3p, miR-148a-3p and miR-139-3p, respectively. In conclusion, serum miR-122-5p, and miR-199b-3p could potentially serve as diagnostic biomarkers for HCC. And serum miR-139-3p, miR-148a-3p, miR-221-3p, and miR-222-3p could potentially serve as diagnostic biomarkers for the iCCA.

Field of Study:	Medical Biochemistry	Student's Signature	
Academic Year:	2021	Advisor's Signature	
		Co-advisor's Signature	

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

LIST OF TABLES

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1   Background and rational

Liver cancer is the world's seventh most prevalent cancer and the second leading cause of cancer-related mortality (1). Liver cancer is classified into several types based on their histological features. Hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (iCCA) are the most prevalent types of liver cancer. The main factors for the development of HCC include hepatitis B virus (HBV) or hepatitis C virus (HCV) infection, alcoholism, and non-alcoholic fatty liver disease (NAFLD) (2). On the other hand, primary sclerosing cholangitis (PSC), biliary cirrhosis (PBC), hepatolithiasis, parasitic biliary infestation with flukes, as well as HBV or HCV infection are involved in the development of iCCA (3). However, curative treatments, such as surgical resection, liver transplantation, and local ablation, are required for the early stages of cancer (4, 5). Conventional markers, such as alpha-fetoprotein (AFP) for HCC detection, carbohydrate antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA) for iCCA detection, currently have limited sensitivity and specificity, especially in the initial stages of the disease. (6, 7). So, identifying a novel diagnostic biomarker that might potentially be used in clinical practice for detecting HCC and iCCA is required.

A meta-analysis from several RNA datasets showed a significant difference in 935 gene expressions between HCC and nontumor tissues (8). Besides gene expressions, non-coding RNAs such as microRNAs (miRNAs) and long non-coding RNA (lncRNA) are promising biomarkers in several cancers (9). MiRNAs are small non-coding molecules that contain about 18-20 nucleotides (nt) and play a crucial role in mRNA regulation and act as regulatory molecules in cancer biology pathways in humans (10). LncRNAs contain nucleotides longer than 200 nt and act as a pseudo miRNA sponge, competing with actual targets for interaction with the miRNAs of interest (11). They can regulate gene expression, which is involved in cancer development and is associated with tumor growth, invasion, and metastasis (9). As a consequence, it is possible to claim that lncRNA can connect with mRNA via miRNA

(12)and microRNA have an important role in regulatory in gene expression. Furthermore, earlier research has shown that miRNAs have a contribution to HCC and iCCA. For example, inhibiting miR-142-5p enhances cell proliferation, which has been related to the development of HCC (13). Koduru *et al.* (14) demonstrated that miRNA may serve as biomarkers to diagnose liver cancer in each phenotypic of the liver. When compared to normal liver tissues, miR-101 was consistently overexpressed in all stages of liver cancer. MiR-22 was shown to be considerably overexpressed in HCC, but miR-23a was found to downregulate interferon regulatory factor-1 expression. When compared to cirrhosis, MiR-7704 was shown to be substantially overexpressed in HCC. According to Wang *et al.* (15), the finding revealed that key miRNAs associated with HCC were miR-199a-3p, miR-199b-3p, miR-139-5p, miR-139-3p, miR-424-3p, miR-1269b, and miR-1269a. The overexpression of lncRNAs, named *HOTAIR* and *ANRIL,* promoted the carcinogenic activity of HCC (16) and iCCA cells (17), respectively.

Nowadays, RNA sequencing (RNA-Seq) is a transcriptomic sequencing approach based on next-generation sequencing (NGS) technologies and is widely used in several perspectives of medical and biological researches, and biomarker discovery (18). Small-RNA sequencing (sRNA-seq) is a type of RNA-seq which is used to analyze and discover small RNAs (19) such as miRNAs. A previous study showed that RNA and sRNA sequencing is the best high-throughput platform that can be used to analyze and investigate the transcriptomic profile (20). Therefore, understanding of the regulatory roles of the miRNAs, lncRNAs and mRNAs may provide the clinical relevance of new diagnostic and prognostic approaches for patients with liver cancers. The main goal of this study is to discover miRNAs via a bioinformatics approach that may be utilized as a novel biomarker for HCC and iCCA diagnosis. The regulatory network of lncRNA-miRNA-mRNA of both liver cancer types will be investigated to study the etiology and mechanism of liver cancers. Moreover, this study aimed to validate the expression of candidate miRNAs and its correlation with clinical outcomes in patients with chronic hepatitis B virus infection and liver cancers.

1.2   Objectives

1.2.1   To identify and validate the potential miRNA from the lncRNA-miRNA-mRNA network that can serve as a novel biomarker for HCC and iCCA.

1.2.2   To investigate the lncRNA-miRNA-mRNA network in HCC and iCCA by retrieving RNA-seq and small RNA-seq data.

1.2.3.   To discover the common molecular characteristics between HCC and iCCA from lncRNA-miRNA-mRNA network.

1.3   Hypothesis

1.3.1   Regulatory networks of lncRNA-miRNA-mRNA for HCC and iCCA could be constructed from retrieving RNA-seq and small RNA-seq data.

1.3.2   There might be potentially detectable associations between HCC and iCCA regulatory networks.

1.3.3   A novel non-coding RNA could serve as a biomarker for HCC and iCCA.

1.4   Conceptual framework

Problem:
Conventional markers, such as AFP*, CA19-9**, and CEA**, have been limited in sensitivity for detecting HCC and iCCA. Therefore, new biomarkers that can be used in clinical practice are required.

↓

To solve problem:
Non-coding RNA such as miRNAs play a crucial role in regulatory molecules and lead to cancer generation which can be used as a biomarker for early detection.

Bioinformatics tools are applied to analyze the high-throughput sequencing data

Processing:
Raw sequencing data collection → Data analysis → Biomarker identification

↓

Outcome:
miRNA could be used as diagnostic biomarkers for HCC and iCCA.

↓

Advantages:
Patients with liver cancers would be detected and diagnosed early. The survival rate of patients with HCC and iCCA could be increased, which might lead to a decrease in cancer-related deaths.

Figure 1   Conceptual framework in this study. *AFP, or alpha-fetoprotein, is a standard marker used to diagnose HCC. **CA19-9 (Carbohydrate Antigen 19-9) and CEA (Carcinoembryonic Antigen) are common markers that are used for iCCA detection. HCC = Hepatocellular Carcinoma, iCCA = Intrahepatic cholangiocarcinoma.

1.5   Research workflow
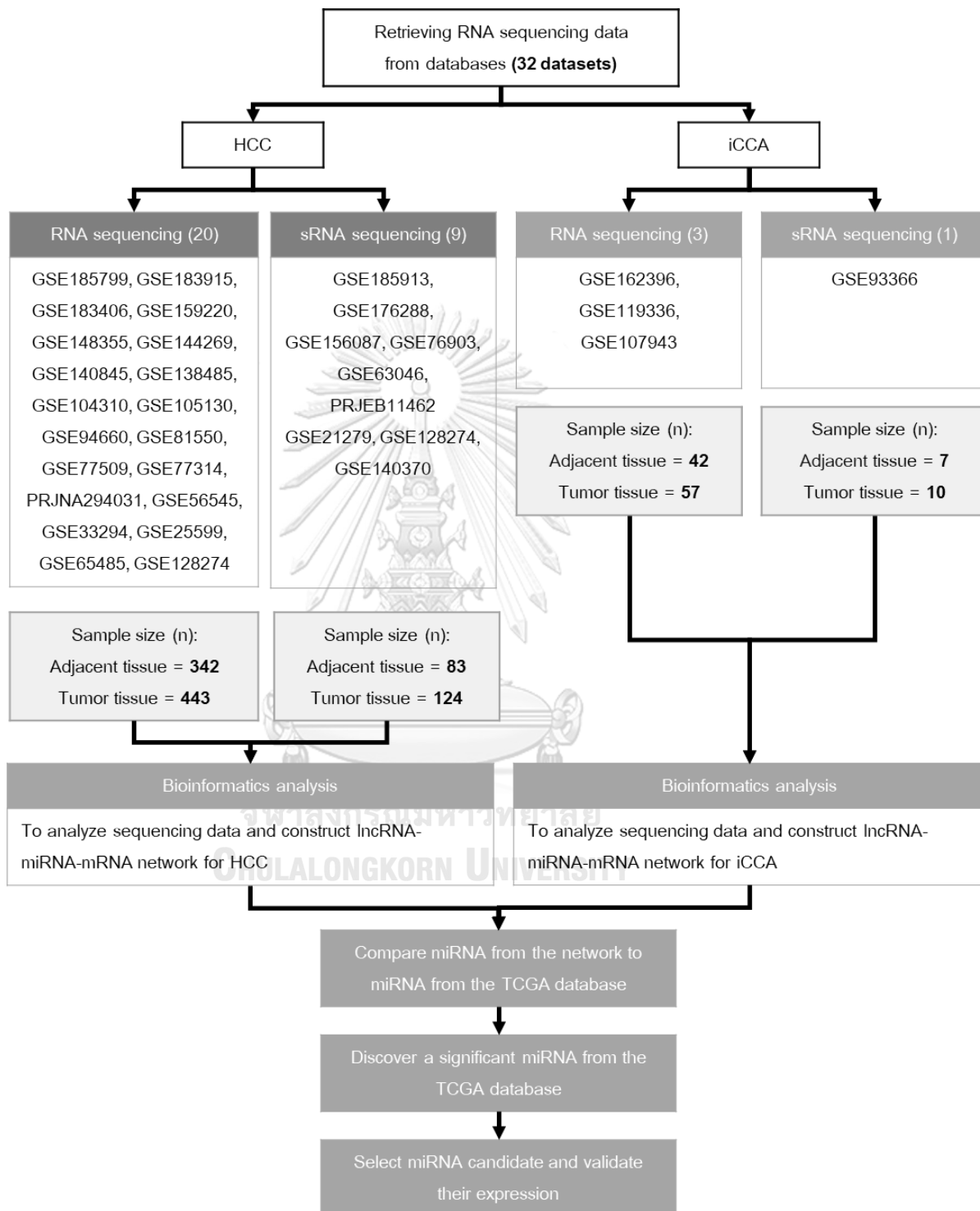
Figure 2   Research workflow for identifying a novel biomarker in liver cancers.

1.6   Expected benefits and applications

To the best of our knowledge, there is no previous study that used bioinformatics tools to explore the non-coding gene interactions of hepatocellular carcinoma (HCC), especially in intrahepatic cholangiocarcinoma (iCCA). Moreover, the results from conducting bioinformatics analysis were validated in serum samples from patients and healthy individuals. The expected outcome for the identification of miRNA biomarkers in liver cancers, including HCC and iCCA would be potentially useful, especially in clinical practice. The expected benefits include improving the accuracy of diagnosis and prognosis of liver cancers, and the overall survival rate of patients.

CHAPTER 2: LITERATURE REVIEW

2.1   Liver cancer

Liver cancer is the seventh most prevalent cancer and the second leading cause of cancer-related mortality (1). Globally, Asia and Africa have the largest rates of occurrence as shown in Figure 3. There are several types of liver cancer, based on histological characteristics, such as hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (iCCA), mixed hepatocellular cholangiocarcinoma (HCC-CCA), fibrolamellar HCC (FLC), and the pediatric neoplasm hepatoblastoma (2). In addition, HCC and iCCA are the most common types of liver cancer. This is because there is over 80 and 20 percent of who are diagnosed with HCC and iCCA from liver cancer cases worldwide, respectively (21). Other kinds of liver cancer, such as combined HCC-CCA and neoplastic hepatoblastoma tumors, account for less than 1 percent of cases. (2). As a result, HCC and iCCA are regarded as a public health concern in people with liver cancer.
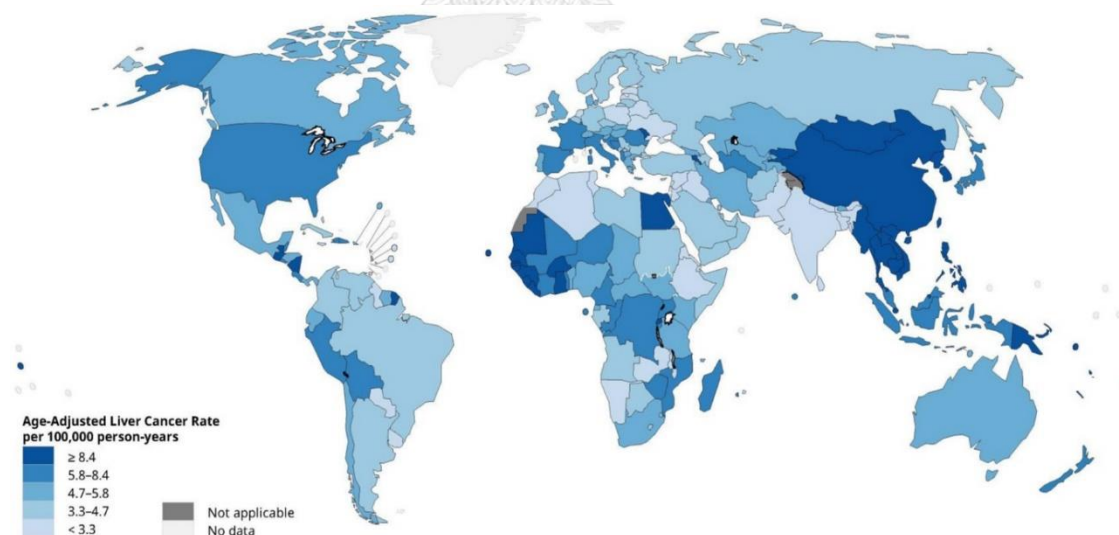
Figure 3   The global incidence rate of liver cancer.

Source: McGlynn KA *et al.*, 2021 (1)

Hepatocellular carcinoma (HCC) is the most common presenting type of liver cancer in the world. Southeast Asia and Sub-Saharan Africa have the strongest incidence, accounting for 90 percent of all liver cancer cases (2). Viral hepatitis B and/or C infection (HBV/HCV), alcoholism, and non-alcoholic fatty liver disease (NAFLD) disease in those with metabolic syndrome and diabetes are the primary risk factors for HCC. In addition, the other common risk factors also include aflatoxin B1 and tobacco (2).

Cholangiocarcinoma (CCA) is divided into three subtypes based on the anatomical location of the tumor. CCA comprises intrahepatic (iCCA), perihilar (pCCA) and distal (dCCA) cholangiocarcinoma. After HCC, the most common type of liver cancer is intrahepatic cholangiocarcinoma (iCCA). The incidence of iCCA is uncommon, with fewer than six cases per 100,000 in most countries, such as in European countries. On the other hand, the incidence of iCCA is largest in East Asia, including China and the Republic of Korea, and Southeast Asia, especially in Thailand (22) (Fig 4). There are many factors that contribute to iCCA development, such as cholangitis or primary sclerosing cholangitis (PSC), hepatolithiasis, parasitic biliary infestation with flukes (Clonorchis sinensis infection), biliary cirrhosis (PBC), as well as HBV or HCV infection. The highest incidence of iCCA in Northeast Thailand in those who eat raw or undercooked fish is commonly related to liver flukes (3).
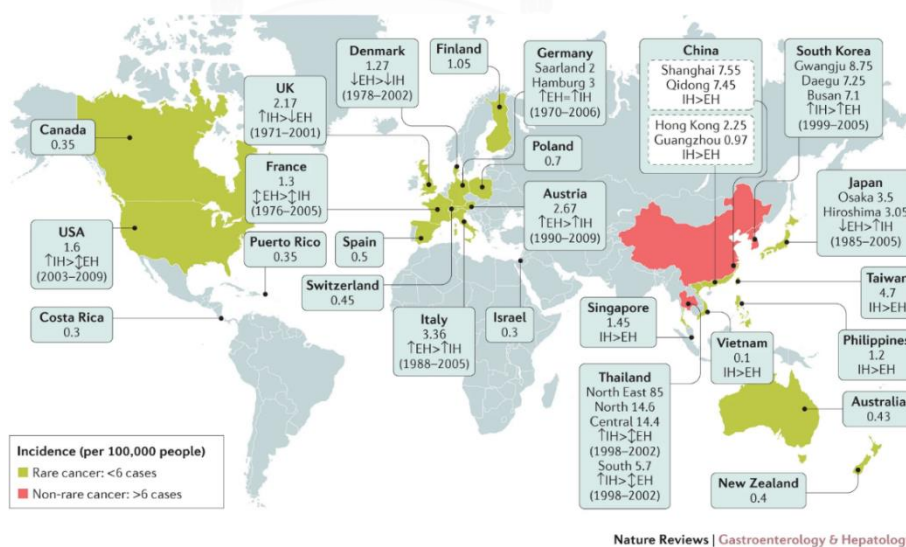
Figure 4   The global incidence rate of cholangiocarcinoma (CCA). Green color shows the rare incidence which is lower than six cases in 100,000 people. Pink

color presents a high incidence which is more than 6/100,000. The incidence of intrahepatic (IH) and extrahepatic (EH) CCA, and also provided the temporal pattern of incidence (↑increasing; ↕stable; ↓decreasing).

Source: Banales JM *et al.*, 2016 (23)

2.2   Cell of origin

The liver consists of two major cell types which are parenchymal and nonparenchymal cells. Parenchymal cells include hepatocytes and cholangiocytes, while others, such as endothelial cells, stellate cells, Kupffer cells and fibroblasts, are defined as nonparenchymal cells. Hepatocytes account for more than 80% of the total liver mass, with the remainder being bile duct epithelial cells or ductular cells. The canal of Hering is used to generate the ductular cells and is considered to be the liver progenitor cells also known as the hepatic progenitor cells (HPCs) (24) (Fig. 5). The presence of stem cells in the adult liver is still under discussion. However, a previous study indicates that HCCs and iCCAs contain stem cell characteristics in some subtypes (2), and hepatocytes and HPCs play a role in liver tumor heterogeneity (24). The discoveries have resulted in a number of ideas concerning the cells that cause liver cancer. Because hepatocytes and cholangiocytes both originate from a shared progenitor (HPCs) during liver development, the HPCs may form primary liver cancers or hepatoblasts. On the other hand, HCC and iCCA might be generated from mature hepatocytes and cholangiocytes, respectively (2) (Fig.6).
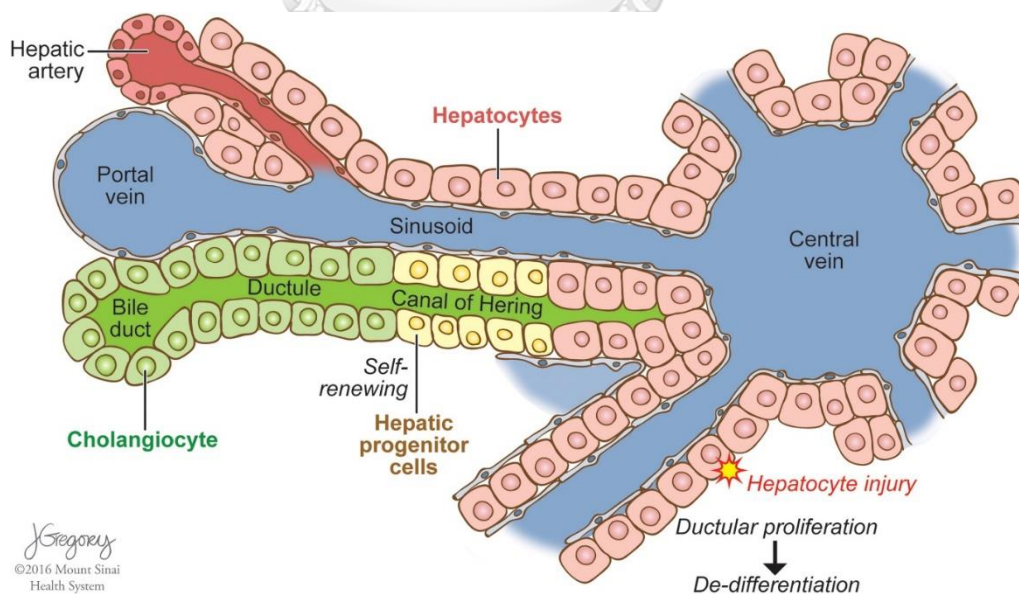
Figure 5   *The liver's cell structure.*Source: Sia D *et al.*, 2017 (2)

According to previous studies (2, 24-26), they also suggest that progenitor cells might be the source of liver cancer. For example, the proteins that are involved in stem cell renewal and cell differentiation are MET50, which mainly stimulates hepatocyte differentiation, and EGFR and NOTCH, which enhance cholangiocyte specification. There are also many factors that are involved in HCC and iCCA development, such as DNA methylation patterns, chromosomal aberrations, as well as epigenetic features, which include the expression patterns of regulatory molecules (2). Therefore, HPCs are involved in the generation of the two most commonly observed primary liver tumors, according to these research studies.
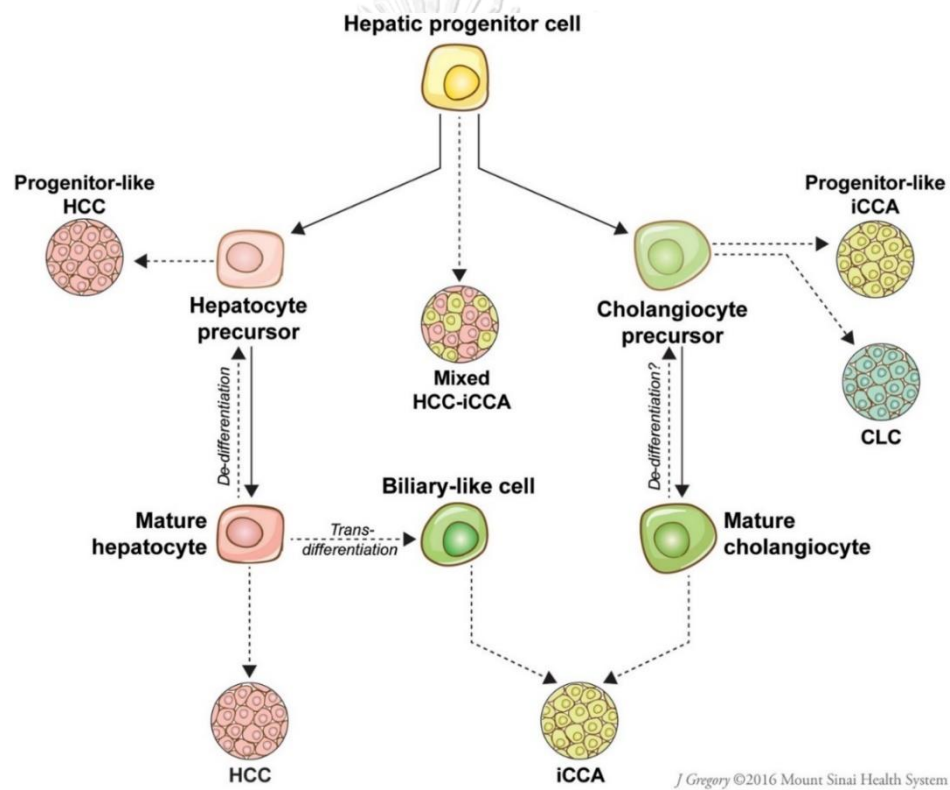
Figure 6   ***Multiple cells of origin****.* HPCs can be differentiated into two cell types which are hepatocytes and cholangiocytes. In the case of mixed HCC-iCCA, it might also progress to liver cancer. Both mature hepatocytes and cholangiocytes can be developed into HCC and iCCA, respectively. However, damaged or injured mature hepatocytes can trans-differentiate into biliary-like cells, which subsequently progress into iCCA.

Source: Sia D *et al*., 2017 (2)

2.3 Diagnosis and treatment

To detect HCC, there are many types of biomarkers that are used nowadays, such as Alpha-fetoprotein (AFP), Lens culinaris agglutinin-reactive AFP (AFP-L3), Des-gamma carboxy-prothrombin, Golgi protein 73 (GP73), Aldo-keto reductase family 1 member 10 (AKR1B10), Dickkopf 1 (DKK1) and High mobility group box 3 (HMGB3). The most common conventional marker used worldwide is to determine the level of AFP in the serum (27). AFP concentrations of more than 400-500 ng/mL will be considered as a diagnosis for HCC, while the AFP level in normal people is presented in low concentrations (28). However, the sensitivity and specificity of the AFP-based approach to diagnosis, especially in early detection, are quite low. According to a recent research, the sensitivity and specificity of their serum AFP-based technique to classifying HCC and liver cirrhosis in a Sicilian population were 65 percent and 89 percent, respectively, at a cutoff value of 30 ng/mL (6). On the other hand, the most common biomarkers for detecting iCCA are carbohydrate antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA). They have been reported to have a wide range of sensitivity and specificity of 50 to 90% and 54 to 98%, respectively (29). However, iCCA is usually asymptomatic during early stages and conventional markers are potentially less effective in early detection (7). Imaging technologies such as ultrasonography, contrast-enhanced ultrasonography (CEUS), CT and MRI, play an essential part in the diagnosis, staging, follow-up, treatment response evaluation as well as decision to be the HCC (28) or iCCA (30).

The treatment of HCC is classified into five stages according to the size of the nodule or child-pugh scores by the Barcelona Clinic Liver Cancer (BCLC) staging system (31). As shown in Figure 7, patients who have a diagnosis of HCC at an initial stage are considered for treatment with resection, transplantation, local ablation, as well as chemoembolization. These stages are considered as curative treatments (4), and there is the highest overall survival rate of patients (31). Patients with intermediate stages and advanced stages of HCC are treated by chemoembolization or receive systemic treatments (31).
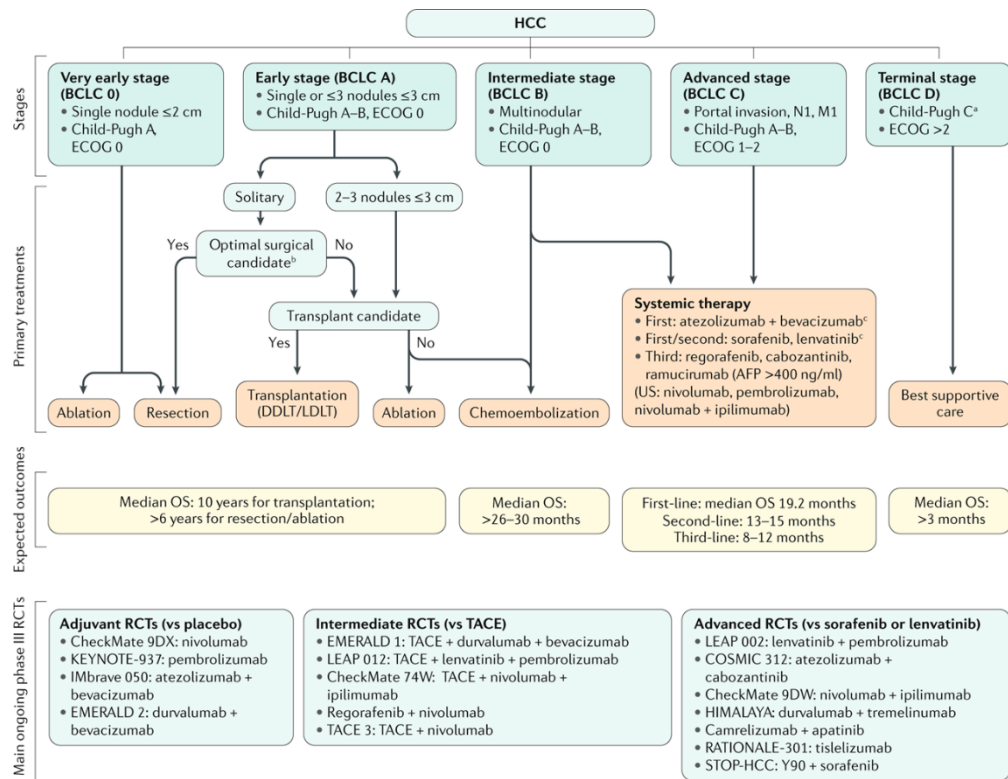
Figure 7   HCC treatment classification by Barcelona Clinic Liver Cancer (BCLC) staging system. DDLT = deceased-donor liver transplantation; ECOG = Eastern Cooperative Oncology Group; HCC = hepatocellular carcinoma; LDLT = living-donor liver transplantation; M1 = distant metastasis; N1 = lymph node metastasis; OS = overall survival; RCT = randomized controlled trial; TACE = transarterial chemoembolization.

Source: Llovet JM *et al*., 2021 (31)

The American Joint Committee on Cancer (AJCC) TNM system is used to classify the treatment of iCCA (30). Patients in the early stages of iCCA, defined as TNM stages I and II, are resected, whereas those in TNM stages III and IV are unresectable. However, there are about 30%-40% of patients who are resectable. This might be because patients with iCCA frequently have massive and locally advanced tumors that require technically difficult and challenging surgeries. Systemic

therapy drugs will be used on patients in TNM stages III and IV (5). The recommended therapy protocol for iCCA patients is shown in Figure 8.
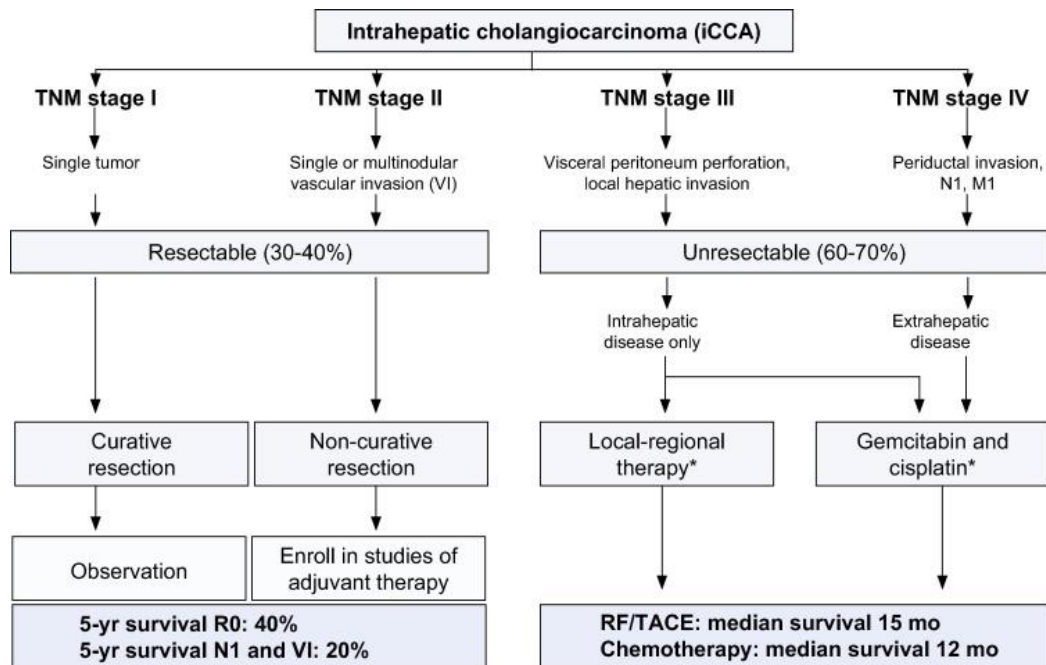
Figure 8 *iCCA treatment classification by the American Joint Committee on Cancer (AJCC) TNM system.* * These are guidelines for standard practices. Larger and more relevant research are needed to provide evidence for the standard of care guidelines.

Source: Bridgewater J *et al.*, 2014 (5)

2.4   Non-coding RNA (ncRNA)

About 20,000 protein-coding genes are found in the human genome, accounting for less than 2% of the entire genome sequence. However, non-coding genes are contained in at least 90% of the total genome sequence. The transcriptome of humans has been revealed to be more complicated than a group of protein-coding genes, and non-coding genes may play a significant biological role in cellular development and metabolism. For example, non-coding RNA (ncRNA) genes are increasingly being used as therapeutic targets in cancer. Developing successful RNA-based anticancer treatments requires an improved comprehension of the complex and context-dependent nature of ncRNA interactions (32). NcRNA is classified into two main groups according to their transcript size which are small ncRNA and long ncRNA (lncRNA). Small ncRNAs are composed of many transcript types which have a nucleotide (nt) length of 18 to 200, such as microRNAs (miRNAs), transfer RNAs (tRNAs), ribosomal 5S and 5.8S RNAs (rRNAs), piwi interacting RNAs (piRNAs), and etc. On the other hand, lncRNAs are non-protein-coding RNA species that are longer than 200 nt (33). Nowadays, Therefore, in this review, the focus is on miRNA.

Micro RNA (miRNA) is one type of small RNA, which has a length of about 18-20 nt. MiRNA is essential in gene expression regulation and is associated with a range of biological processes. In addition, the dysfunctional expression of miRNA is involved with human disease (11). A previous study, for example, found that overexpression of miR-142-5p reduces cell viability and promotes apoptosis through regulating forkhead box O1 (FOXO1) genes. In contrast, inhibiting miR-142-5p enhances cell proliferation through the overexpression of FOXO1, which has been associated with the development of HCC (13). In addition, iCCA tissues showed overexpression of glucose transporter 1 (GLUT1) genes, which was directly regulated by miR-148a. The results showed that lowering miR-148a levels induces GLUT1 overexpression in iCCA, which leads to cancer development (34). Nowadays, some miRNAs are applied in clinical trials, such as miR-21 and miR-122, which can be used as biomarkers for detection of HCC (35) and iCCA (36).

2.5   lncRNA-miRNA-mRNA interaction

As a miRNA biogenesis pathway (Fig. 9), first in the nucleus, RNA Polymerase II and transcription factors transcribed miRNA genes into hairpin miRNA also known as primary miRNAs (pri-miRNAs). Then pri-miRNAs are turned into precursor miRNAs (pre-miRNAs) which are processed by the RNase III enzyme complex, which is Drosha and DGCR8, and then exported into the cytoplasm. The complex RNase III enzyme named Dicer and TRBP turned pre-miRNAs into a duplex miRNA which consist of a guide strand (miRNA) and passenger star strand (miRNA*). Argonaute (AGO) protein will react with duplex miRNA to yield mature miRNA. Finally, mature miRNA is incorporated into the RNA-induced silencing complex (RISC) and serves as a guide strand to recognize target mRNAs through sequence-based complementarity. Therefore, if mature miRNA is perfectly matched with target mRNA, it will result in target mRNA cleavage or mRNA degradation. On the other hand, if mature miRNA is partially matched with target mRNA, then the translation will be repressed (37). This is the reason why miRNAs play an important role in gene expression regulation and are also involved in human diseases.
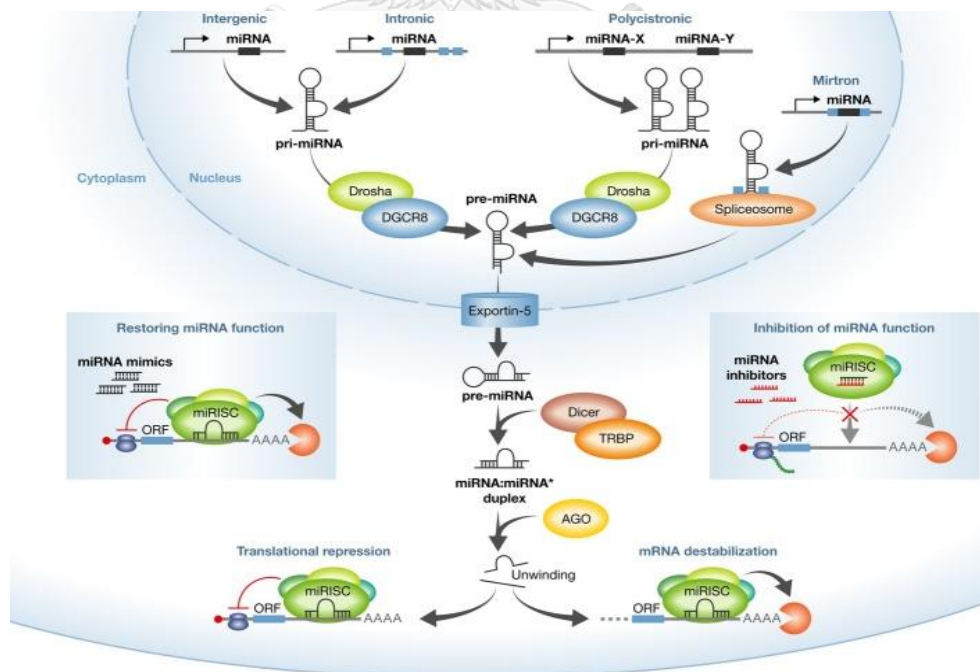
Figure 9   MiRNA biogenesis pathway.

Source: van Rooij E *et al*., 2014 (38)

As previous study, lncRNAs typically interact with mRNA, DNA, protein and miRNA, and as a result, they influence gene expression in multiple ways such as epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels (39). However, most gene expression interaction occurs between lncRNA and miRNA (40). First of all, lncRNA, in both linear and circular form, can act as miRNA sponges (Fig. 10A) by binding to the miRNA sequences and then affecting the interaction between miRNA-target mRNA in the gene regulation. From Figure 10B, lncRNA can serve as a miRNA precursor which produces mature miRNAs after Dicer and/or Drosha cleavage. Nevertheless, miRNA can interact with lncRNA and result in lncRNA degradation (Fig. 10C), which is the same as the mechanism of miRNA that is perfectly matched with target mRNA. Finally, as shown in Figure 10D, lncRNA can compete with miRNA for binding to the mRNA target. Due to the competition between lncRNA and miRNA, they are also known as endogenous competitor RNAs (ceRNAs) (40). Therefore, the lncRNA-miRNA-mRNA interaction may be a crucial mechanism key to understanding the etiology of HCC and iCCA.
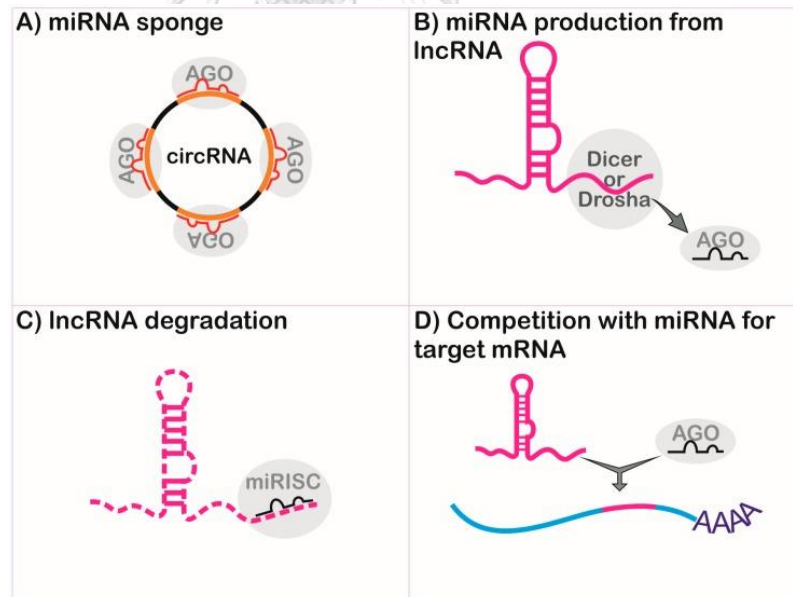
Figure 10   Interaction of lncRNA and miRNA. circRNA = circular lncRNA; AGO = argonaute proteins; miRISC = miRNA-induced silencing complex; Dicer or Drosha = endoribonuclease.

Source: Fernandes JCR *et al.,* 2015 (40)

2.6   Next generation sequencing

The principle of next-generation sequencing (NGS) is the integration of fluorescently tagged deoxyribonucleotide triphosphates (dNTPs) into a DNA template strand, which processes across millions of DNA fragments, and the nucleotides are identified by fluorophore excitation. There are four basic steps, as follows: (1) Library preparation is the method whereby the adapters are added to both ends of DNA fragments. The ligation of adapters and fragments are amplified and then purified (Fig. 11A). (2) Cluster generation is the process in which each fragment is amplified. The library from step (1) is loaded into the flow cell, which contains the oligos that are complementary with the fragments. After that, the fragments are complimentary bound with oligo and perform double strand bridges (Fig. 11B).
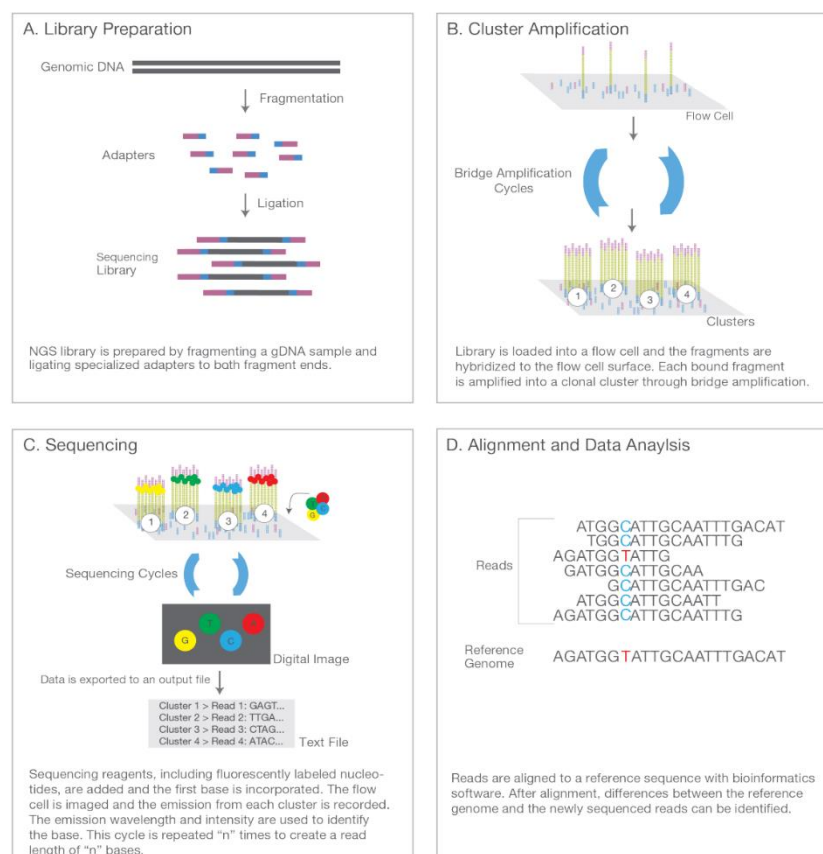
Figure 11   Overview of four basic steps of next generation sequence.

Source: Illumina, lnc. 2017 (41)

(3) Sequencing is the process in which fluorescently tagged nucleotides are incorporated into DNA template strands, then the characteristics of each fluorescent tag are emitted. This process is called sequencing by synthesis (SBS) (Fig. 11C). (4) Data analysis is the technique that can be used to analyze the sequencing read data. The reading data are mapped to the reference sequence during analysis. (Fig. 11D) (41).

In terms of application areas, NGS is widely used in a variety of areas of medical and biological research. RNA (RNA-Seq) and small RNA sequencing (sRNA-Seq) are transcriptomic sequencing approaches that are based on using NGS technologies to study the whole transcriptome. Both have several important characteristics, including the ability to identify transcripts with low expression levels and to analyze transcripts with or without a reference sequence. Furthermore, the two techniques have been shown to be strong techniques for reliably determining expression levels (18).

2.7   Bioinformatics analysis

Bioinformatics is a study of biological systems that integrates many fields of study, including biology, mathematics, computer science, and statistics. Bioinformatics tools can be used to analyze biological data, such as for gene characterization and analysis of DNA of protein sequences, protein 3D structure prediction, physiochemical characteristics determination, phylogenetic analysis, drug design, and simulations to understand how biomolecules interact in the living cell (Fig. 12) (42). This study is focused on sequencing analysis.

Sequencing analysis is used to understand different characteristics of a biomolecule such as the sequence of a nucleic acid or a protein. Sequencing data can be retrieved from public databases, which contain a lot of information (42). RNA or sRNA sequencing is the best high-throughput platform that can be used to analyze and investigate the transcriptome. The purpose of a bioinformatics analysis of RNA/sRNA sequencing data is to identify and quantify these sequences. Quantitative analysis of gene expression, such as the differential gene expression between a group, is an important way to better understand diseases. The workflow of RNA/sRNA sequencing via a computational approach includes three main steps: (1) Raw sequencing read alignment to a reference genome or transcriptome; (2) differential gene expression analysis; and (3) identification of the specific gene with differential expression between two groups of samples. The result will be used for further analysis, such as functional and pathway analysis (20).
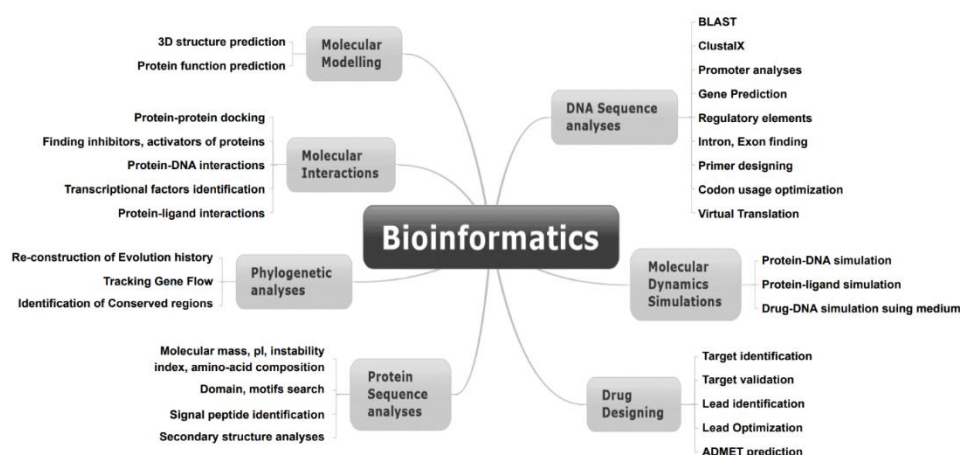
Figure 12   Bioinformatics can be applied in various areas of studies. ADMET = Absorption, Distribution, Metabolism, Excretion, and Toxicity. BAST = The Basic Local Alignment Search Tool.

Source: Mehmood M *et al.*, 2014 (42)

The Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) is supported by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM). This database is an international public repository that aims to store high-throughput gene expression data provided by the scientific community, such as microarrays, next-generation sequencing, and other kinds of high-throughput (43). The raw RNA/sRNA sequencing reads can be downloaded from these databases, then analyzed by using bioinformatics tools. For example, software packages in R programs such as DESeq2, edgeR, limma and Bowtie can be used to analyze the differential of gene expression between two groups of samples (44). WGCNA is a software tool that is used to build a co-expression module based on similar gene expression patterns in order to expose modules (clusters) of highly correlated genes and identify modules that are connected to one another and to external sample traits (45). Moreover, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) tools (46) can be used to analyze the functional and pathway of gene enrichment, respectively.

CHAPTER 3: MATERIALS AND METHODS

3.1   Discovery set

3.1.1   Retrieving data from GEO database

RNA sequencing and small RNA sequencing data of both HCC and iCCA were collected from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The two keywords were used to explore the sequencing dataset in HCC and iCCA which are "Hepatocellular carcinoma" and "Intrahepatic cholangiocarcinoma", respectively (Fig. 13). However, this study focused on RNA-seq in Homo sapiens organisms. As a result, there were 28 HCC datasets that can be divided into three groups based on sequencing type: RNA-seq for 19 datasets, sRNA-seq for 8 datasets, and one dataset that used both RNA- and sRNA-seq. Moreover, there were four datasets for iCCA which were subgrouped into RNA-seq for 3 datasets and another for sRNA-seq (Fig. 13). So, these data included RNA-seq data from a total of 785 tumor (HCC) and adjacent tissues and small RNA-seq data from 207 HCC and adjacent tissues. For iCCA, tumor and adjacent tissues, a total of 99 and 17 tissues were used for the RNA-seq and small RNA-seq analyses, respectively.
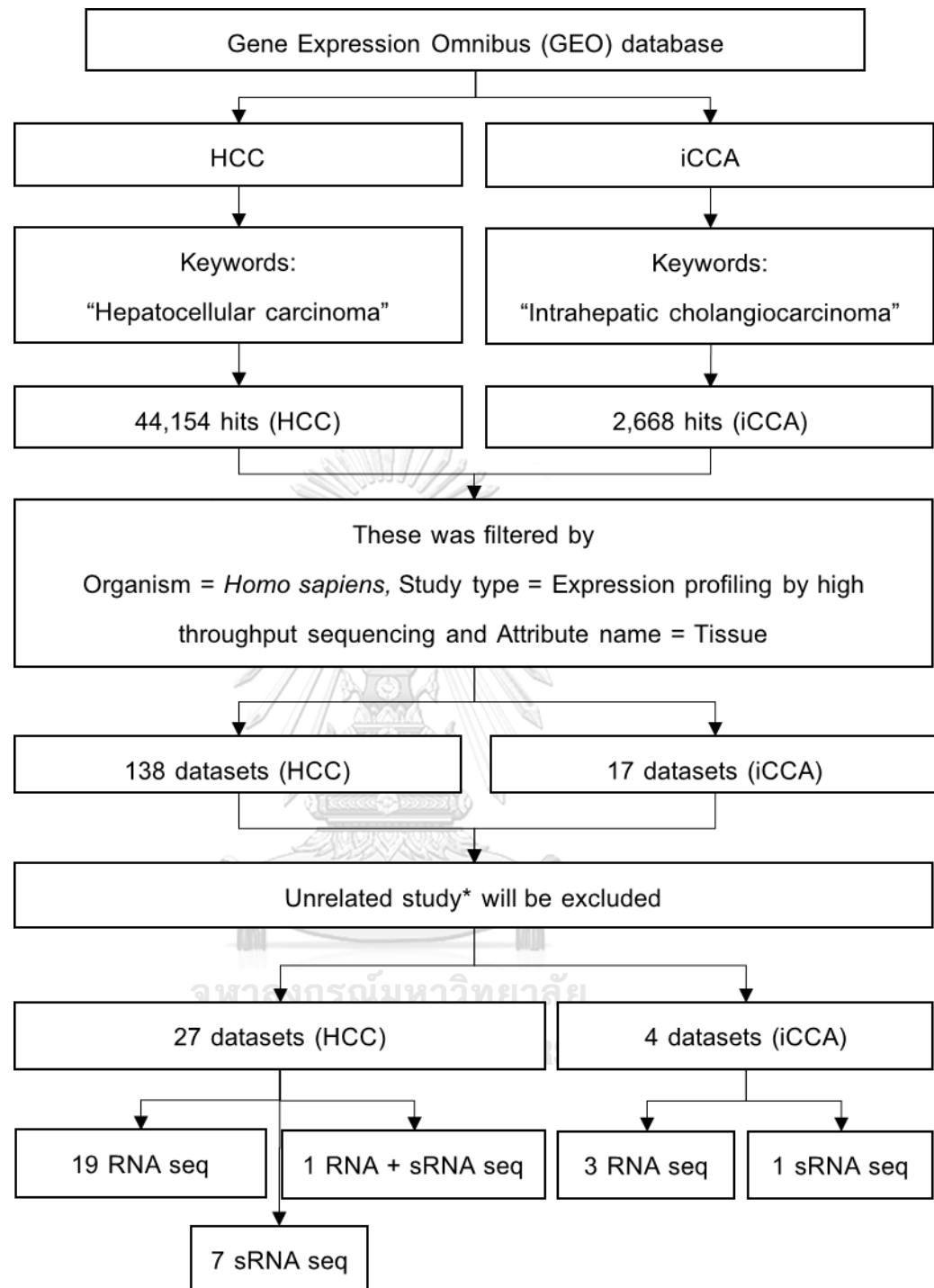
Figure 13   The flow chart of data collection from the Gene Expression Omnibus (GEO). *Unrelated studies such as expression profile of protein coding gene, circRNA, piRNA and snRNA-seq in liver cancers will be excluded.

3.1.2   Pre-processing data and quality control (QC)

Quality control and preprocessing of the raw sequencing data are required to yield clean data (47). The raw reads were filtered out with quality scores lower than 30 (Q30), and the detectable adaptors were trimmed. Fastp was used in this study to prepare clean data, including quality control, read filtering, and base correction for FASTQ data (47).

3.1.3   Differential expression (DE) analysis

The clean RNA-seq data were aligned to human reference genome from the genome reference consortium human build 38 (GRCh38) for the long non-coding RNAs (lncRNAs) and messenger RNAs (mRNAs) expression profiles using HISAT2 (48). Sequencing reads were assembled and merged. In this step, the StringTie (49, 50) was used as an assembler of the sequence read to provide a potential transcript (51). The data was normalized with Fragments Per Kilobase Million (FPKM) units. The DESeq2 package (52) in the R program was used to measure the differential gene expression of lncRNAs and mRNAs. As DE analysis criteria, $|Log_2 (FC)| > 1.0$ and adjusted p-value 0.05 were used. The significant lncRNAs (DElncRNAs) and mRNAs (DEmRNAs) were analyzed in the next step.

For the micro RNAs (miRNAs) expression profile, the Nextflow pipeline (53) for small RNA-sequencing was presented to analyze the sequencing data, including aligning to miRBase mature miRNA and the human reference genome using Bowtie1, discovering known and novel miRNAs with the mirdeep2 module, and interpreting the results. The DESeq2 package (52) in R program was used to measure the differential gene expression of miRNAs. As a DE analysis criterion, $|Log_2 (FC)| > 1.0$ and an adjusted p-value of 0.05 were used. The significant miRNAs (DEmiRNAs) were analyzed in the next step.

3.1.4   Target prediction

A total of significant DElncRNAs were used to predict the miRNAs target by using the miRcode database. Then all predicted miRNAs were compared to the significant DEmiRNAs. The intersection between predicted miRNAs and the significant DEmiRNAs was used for discovering mRNA targets which was conducted using the multiMIR package version 2.3.0 in the R program. The package was estimated according to the 8 databases which contained DIANA-microT-CDS, ElMMo, MicroCosm, miRanda, miRDB, PicTar, PITA and TargetScan database (54). Subsequently, the lncRNA-miRNA-mRNA network was constructed and visualized using Cytoscape software (55). The process overview of target prediction is shown in Figure 14.
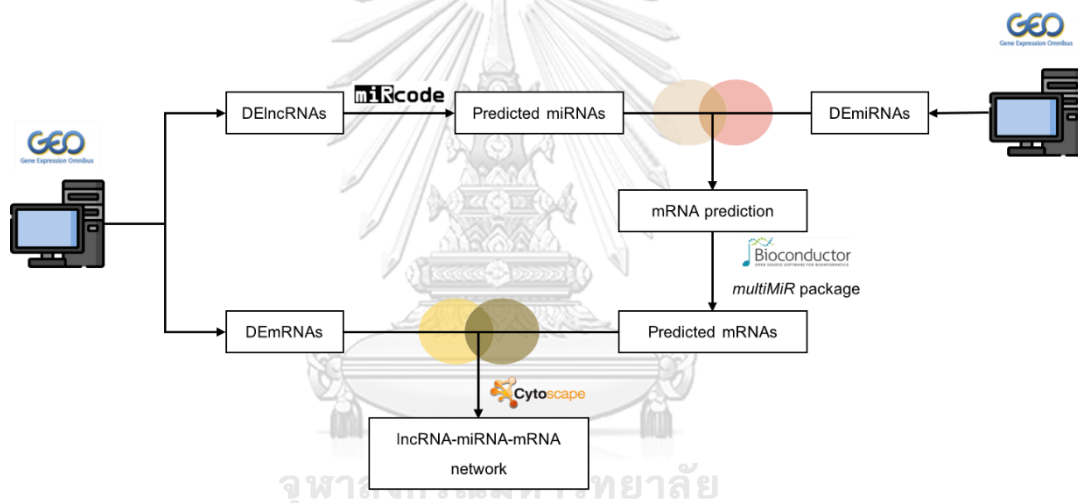
Figure 14   The overview steps to predict target non-coding and coding genes. The miRcode database was used to predict miRNA targets. Subsequently, the intersect results between predicted miRNA and DEmiRNA (differential expression miRNA) will be further used to predict mRNA targets via the R program with the multiMiR package.

Source: This pipeline was adapted from Zhang H. *et al.*, 2021 (56).

3.1.5   LncRNA-miRNA-mRNA networks construction

The lncRNA-miRNA-mRNA networks in HCC and iCCA were constructed and visualized into an interaction network using the Cytoscape software (version 3.9.0) (55). This study was focused on miRNAs due to the fact that miRNAs can be perfectly or partially matched with target mRNAs which leads to the dysfunction of mRNA, so miRNA has an important role in gene expression regulation and is also involved in human diseases (37). Therefore, all of the miRNAs in the network were further analyzed.

3.1.6 The comparison between miRNAs from the network and miRNAs from TCGA database

All of the miRNAs from the network were aligned with miRNAs which were retrieved from the TCGA database (https://portal.gdc.cancer.gov/). The data from the TCGA database contains 372 HCC with 50 adjacent tissues, and 36 iCCA with 9 adjacent tissues. Moreover, there are 2,563 miRNAs that were found in the TCGA database. Following that, the list of interesting miRNAs was subjected to the significant differential expression of miRNAs from the overlapping. The candidate miRNAs were selected according to the criteria in Table 1 for HCC and Table 2 for iCCA. After filtering, there were three candidate miRNAs in HCC, including miR-122-5p, miR-182-5p, and miR-199b-3p, as well as four candidate miRNAs in iCCA, including miR-139-3p, miR-148a-3p, miR-221-3p, and miR-222-3p.

Table 1   Criteria for selecting miRNA candidate in HCC

Conditions	Criteria for HCC
Area under the curve (AUC)	>0.75
Log_2 fold change (for up regulation)	>2.0
Log_2 fold change (for down-regulation)	<-1.5
P adjust value	<0.05
Regulation in two data (from the network and TCGA database)	Need to be the same

Table 2   Criteria for selecting miRNA candidate in iCCA

Conditions	Criteria for iCCA
Area under the curve (AUC)	>0.95
Log$_2$ fold change (for up regulation)	>2.0
Log$_2$ fold change (for down-regulation)	<-1.5
P adjust value	<0.05
Regulation in two data (from the network and TCGA database)	Need to be the same

3.2   Validation set

3.2.1   Sample collection and miRNA isolation

The sample size is calculated by using the G*power program version 3.1.9.7 with conditions as follows:

Test family: F tests   **Statistical test:** ANOVA: Fixed effects, omnibus, one-way

Analysis: A priori: Compute required sample size

Input: Effect size f: 0.25

α err prob: 0.05

Power (1-β err prob):   0.90

Number of groups: 4

Output: Noncentrality parameter λ: 14.5000000

Critical F: 2.6441945

Numerator df: 3

Denominator df: 228

Total sample size: 232

Actual power: 0.9018055

From the calculation, total sample size was 232 samples which contained 58 samples in each healthy, CHB, HBV-related HCC and iCCA groups. However, in this study, the leftover 190 specimens from IRB No. 196/64 were used which were divided into three groups, which contained 60 healthy, 60 chronic hepatitis B virus, and 70 HCC serum samples. On the other hand, there were only 18 leftover iCCA and 18 healthy specimens from IRB No. 196/64 that were used in this study. A Total serum samples were obtained from the medical record of King Chulalongkorn Memorial Hospital with the inclusion and exclusion criterias as shown in Table 3.

Table 3   *The inclusion and exclusion criterias for collecting the samples*
from HCC and iCCA patients.

Inclusion criteria	Exclusion criteria
• Patients identified and confirmed positive for liver cancers (HCC, iCCA) by typical imaging studies and/or histopathology in accordance with the American Association for the Study of Liver Diseases (AASLD) guideline • Thai male and female patients with age between 35-75 years old	• Other nations • Patients with a history of liver cancer treatment such as Radiofrequency ablation (RFA) and transarterial chemoembolization (TACE) • Patients who have a history of Hepatitis C Virus (HCV) and/or Human Immune deficiency Virus (HIV) infections • Heavy alcohol consumption history

        3.2.2   miRNA isolation and validation of candidate miRNAs using qRT-PCR

The total miRNA was isolated by using the miRNeasy Serum/Plasma Kit (QIAGEN, Germany) according to the manufacturer's protocol. The RNA concentration and purification (A260/280) were measured using a NanoDrops spectrophotometer. The viral miRNA was reversed to cDNA by adding Poly U (New England Biolabs, USA) first. Then the first-strand cDNA was synthesized from purified total RNA using SL-poly (A) sequence (GTCGTATCCAGTGCAGGGTCCGAGGTATTCGC ACTGGATACGACAAAAAAAAAAAAAAAAAAAVN) and RevertAid First Strand cDNA Synthesis (Thermo Scientific, Carlsbad, CA, USA), respectively. The expression level of candidate miRNAs was evaluated using qRT-PCR with SYBR Green. (QPCR Green Master Mix HRox, 4x, Biotechrabbit). Specific primers are provided in Table 4. The relative quantitation of expression levels was determined by the $2^{-\Delta\Delta Ct}$ method. The results were normalized to the miR-21 (57) and U6 spliceosomal RNA (U6) in HCC and iCCA, respectively. All qRT-PCR reactions were conducted in duplicate.

Table 4   *Specific primers of candidate miRNAs*

Group	Name	Sequences (5' to 3')
HCC	miR-122-5p (forward)	TGGAGTGTGACAATGGTGTTTG
HCC	miR-182-5p (forward)	TTTGGCAATGGTAGAACTCACACT
HCC	miR-199b-3p (forward)	ACAGTAGTCTGCACATTGGTTA
iCCA	miR-139-3p (forward)	TGGAGACGCGGCCCTGTTGGAGT
iCCA	miR-148a-3p (forward)	TCAGTGCACTACAGAACTTTGT
iCCA	miR-221-3p (forward)	AGCTACATTGTCTGCTGGGTTTC
iCCA	miR-222-3p (forward)	AGCTACATCTGGCTACTGGGT
-	Universal reverse	GCAGGGTCCGAGGTATTCG

3.2.3   Bioinformatics and statistical analysis

For the discovery set, the operating system Ubuntu on Window version 18.04 and Rstudio program version 4.1.2 were used to analyze the data in this study. First of all, retrieving sequencing data and pre-processing analysis were conducted by using Ubuntu software. Differential expression analysis was calculated by using Rstudio program with the DESeq2 package. The volcano plot, which is used to interpret the differential expression result, was constructed by the EnhancedVolcano package. Receiver operating characteristic (ROC) was used to analyze whether the how well of sensitivity and specificity of interested miRNAs in diagnostic of HCC and iCCA. The area under the ROC curve (AUC) was calculated using ROCR package. For the validation set, the Statistical Package for the Social Sciences (SPSS) program version 28 (SPSS Inc., United States) was used for statistical analysis. Unpaired t-test with nonparametric Mann Whitney test was used to analyze the differential relative expression between two groups of the tumor and non-tumor groups. The relative expression data was visualized using GraphPad Prism 9.0 (GraphPad software, CA, USA) Quantitative data in the baseline characteristics was shown as median (IQR) which was analyzed by the nonparametric Kruskal-Wallis test for three groups (healthy, CHB, and HCC) of patients, and Mann-Whitney test for comparing between groups of patients. Categorical data in the baseline characteristics was shown as

number (percentage) (n (%)) which was calculated by using Chi-square test. ROC analysis was calculated by SPSS software and visualized in GraphPad software. The statistical significance was determined to be p-value < 0.05.

CHAPTER 4: RESULTS

4.1   Differential expression analysis

From the differential expression analysis of 443 HCC and 342 adjacent tissues, the result showed that 222 lncRNAs (Fig. 15A) and 361 mRNAs (Fig. 15C) were found to be differentially expressed between tumor and non-tumor. In addition, 125 miRNAs (Fig. 15B) were found from the differential expression analysis of 121 HCC and 80 adjacent tissues. According to the criteria, with |log2 (FC)| > 1.0 and an adjusted p-value of 0.05, there were 186 DElncRNAs, 37 DEmRNAs and 82 DEmiRNAs that showed significant differences in expression between tumor and non-tumor. On the other hand, the results showed that there were 182 DElncRNAs (Fig. 15D) and 1,005 DEmRNAs (Fig. 15F) were found in the differential expression analysis of 57 iCCA and 42 adjacent tissues, while 338 DEmiRNAs (Fig. 15F) were identified in the differential expression analysis of 10 iCCA and 7 adjacent tissues. However, there were 108 DElncRNAs, 145 DEmRNAs and 58 DEmiRNAs that showed significant different expression. A total of significant genes in both HCC and iCCA were further used in the next step. A list of all significant genes is provided in Appendix.
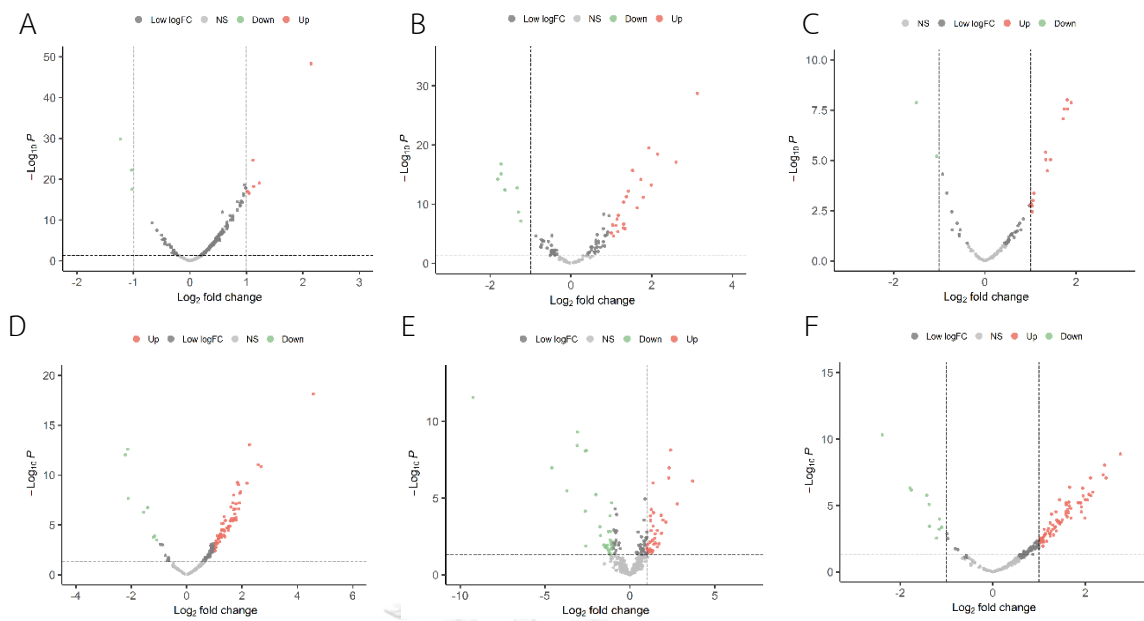
Figure 15   Differential expression analysis. A Volcano plot is used to visualize the finding in HCC analysis of (A) differential expression in lncRNA (DElncRNA), (B) DEmiRNA and (C) DEmRNA, as well as in iCCA analysis of (D) DElncRNA, (E) DEmiRNA and (F) DEmRNA. Red and green color dots represent upregulation and downregulation, respectively. The genes that showed no significant difference between tumor and non-tumor are represented in light gray. The genes with a |$\log_2$(FC)| value of > 1.0 and an adjusted p-value of 0.05 are indicated as significant genes.

4.2   lncRNA-miRNA-mRNA network

The significant genes of both HCC and iCCA from the previous results were used to construct the network. The mircode database was used to predict miRNA targets from lncRNA. In addition, prediction DEmRNA targets from DEmiRNA were from 8 predictor tools (DIANA-microT-CDS, ElMMo, MicroCosm, miRanda, miRDB, PicTar, PITA and TargetScan database), which were included in the multiMir package from R. From the significant genes in HCC, there were only 92 DElncRNAs, 38 DEmiRNAs and 11 DEmRNAs involved in the lncRNA-miRNA-mRNA network (Fig. 16). At the same time, 50 DElncRNAs, 17 DEmiRNAs and 30 DEmRNAs are involved in the lncRNA-miRNA-mRNA iCCA network as shown in Figure 17. So, these networks

showed that the dysregulation of coding genes, such as mRNAs, can be affected by non-coding RNAs, such as lncRNAs and miRNAs. For example, the dysfunction of lncRNA OSTM1-AS1 in the HCC network affected many miRNAs, such as miR-101-3p, miR-122-5p, miR-143-3p, miR-182-5p, miR-193b-3p, miR-19a-3p, miR-23a-3p, miR-24-3p, miR-27b-3p, hsa-miR-30b-5p and hsa-miR-34a-5p, as well as lncRNA SLC26A4-AS1 in the iCCA network affected the miR-101-3p, miR-139-3p, miR-142-3p, miR-221-3p, miR-222-3p, miR-375-3p, miR-454-3p. Consequently, these miRNAs disturb the expression of mRNAs and could lead to abnormal gene function and cancer development. As mentioned above, it can be concluded that lncRNA can communicate with mRNA via miRNA. In this study, miRNAs from the networks were used in further analysis.
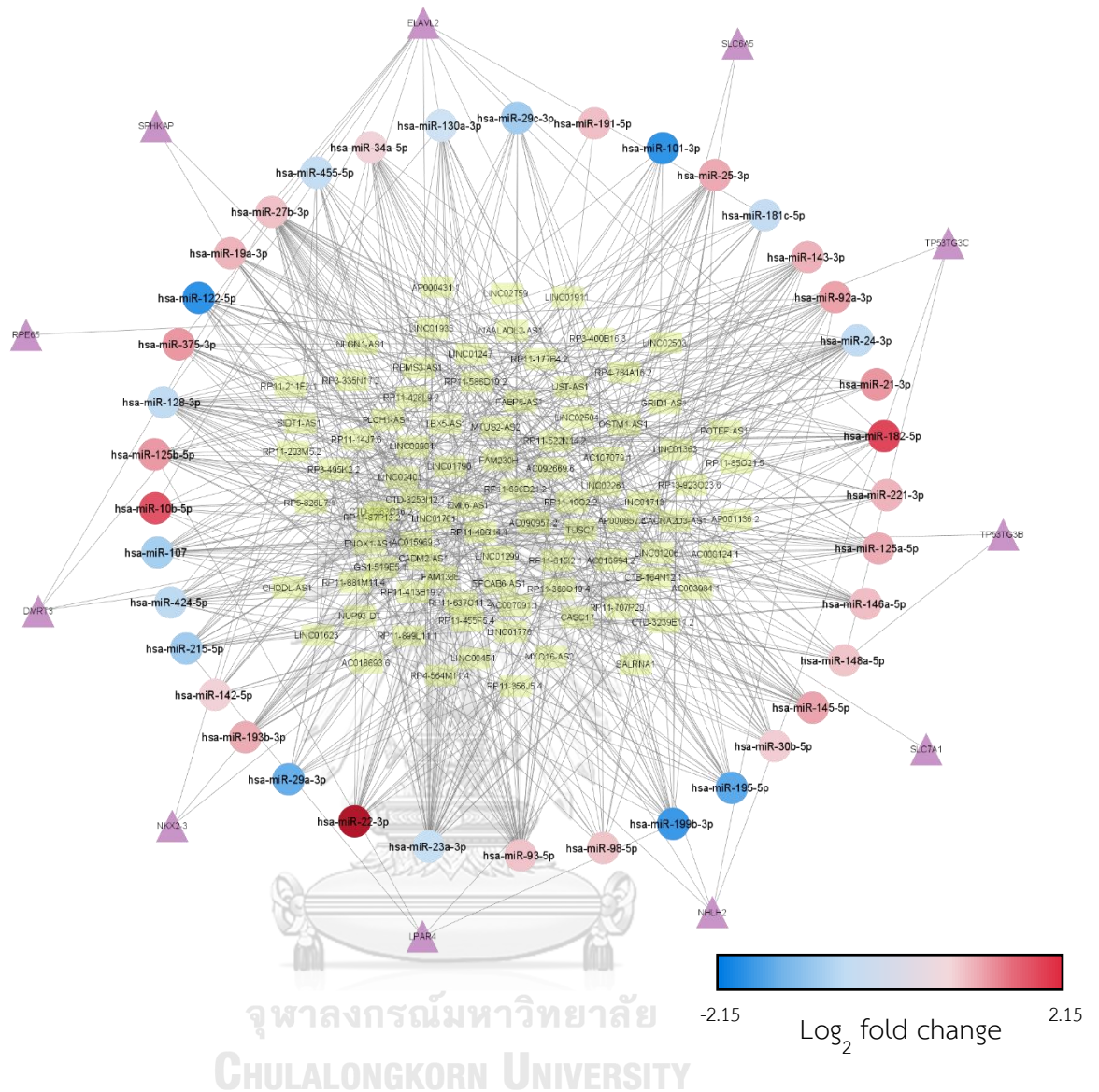
Figure 16 ***lncRNA-miRNA-mRNA network of HCC.*** A total of DElncRNAs are represented in green inside the network, as well as DEmRNAs are shown as a triangle in purple. A total of DEmiRNAs are represented in round shapes and colored according to their expression.
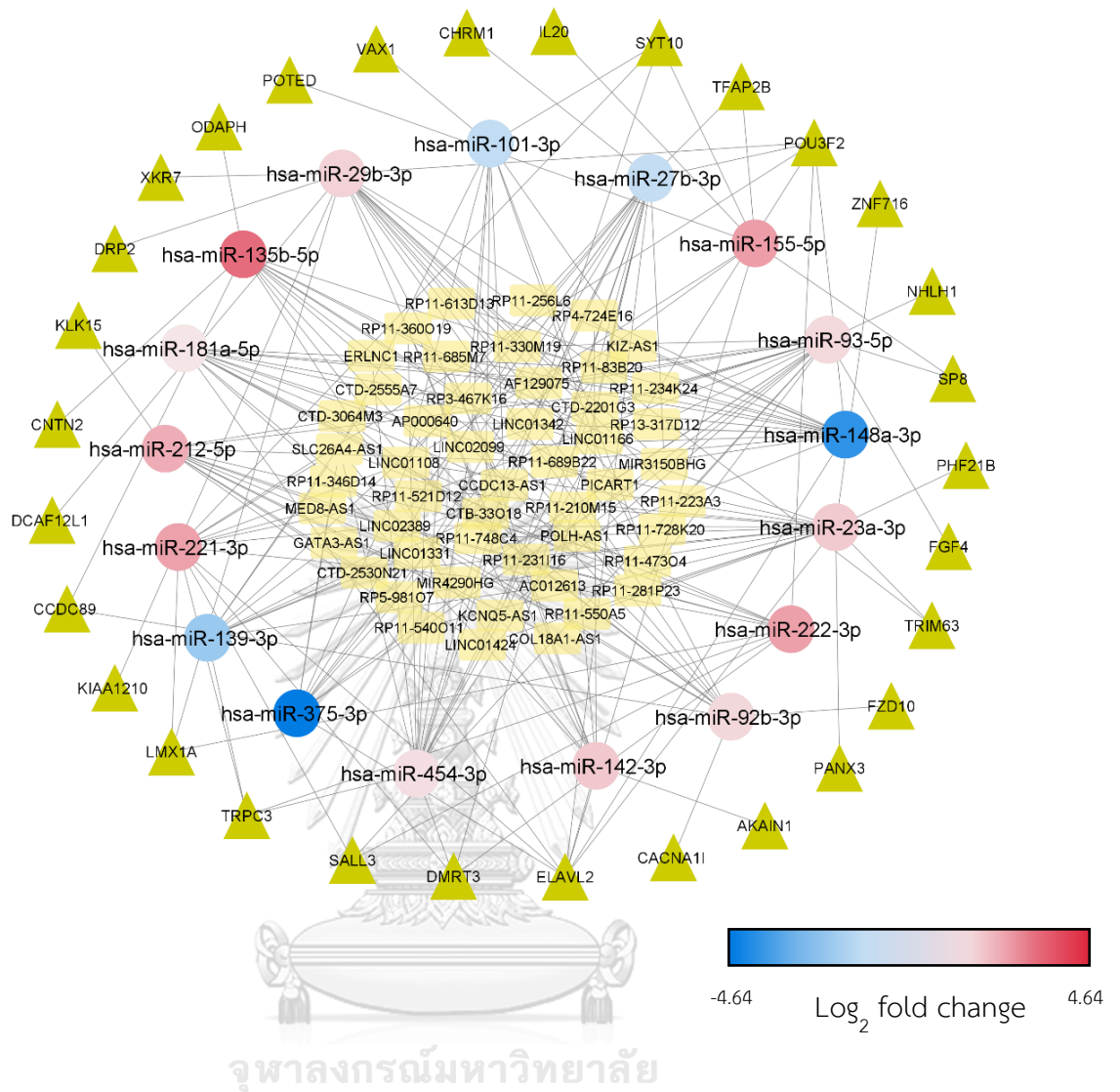
Figure 17   *lncRNA-miRNA-mRNA network of iCCA.* A total of DElncRNAs are represented in yellow inside the network, as well as DEmRNAs are shown as a triangle in green. A total of DEmiRNAs are represented in round shapes and colored according to their expression.

4.3   Compare DEmiRNAs from the network to miRNAs from the TCGA database

There were 38 and 17 DEmiRNAs that was obtained from the HCC and iCCA network, respectively. According to 372 HCC with 50 adjacent tissues, and 36 iCCA with 9 adjacent tissues from the TCGA database, there are 2,563 miRNAs that were found. After comparing, the finding indicated that there were 30 miRNAs, from the TCGA database, showed significant different between HCC and adjacent tissue (Fig. 18A, 18B), while 12 miRNAs showed significant different in iCCA and adjacent tissue (Fig. 18C, 18D).
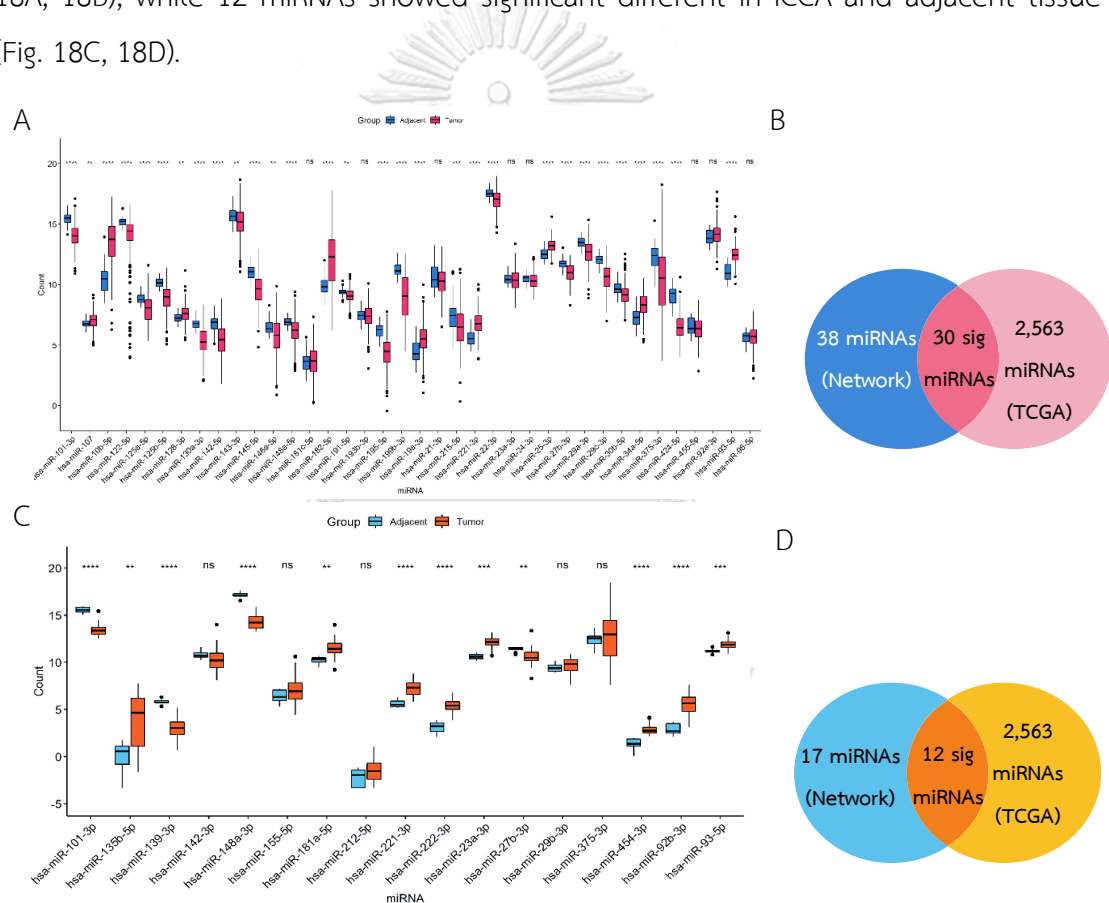
Figure 18   *Comparison of miRNA from the network and miRNA from the TCGA database and their expression.* Multiple box plots are used to represent miRNA expression, from the TCGA database, in both (A) HCC and (C) iCCA. The overlapping between significant miRNA in HCC (A) and in iCCA (C) which obtained from the network and obtained from the TCGA database showed in Venn diagram.

Mann Whitney test was used to analyze in statistic with p-value < 0.05. **** P-value < 0.001, *** 0.001 ≤ P-value < 0.01, ** 0.01 ≤ P-value < 0.05, * P-value < 0.05.

4.4   Selection and validation of miRNA candidate using RT-qPCR

The significant miRNAs from the TCGA database (Table 5 and Table 6) were further merged with differential expression data from Result 4.1. The area under the curve (AUC) was analyzed to determine whether the candidates could distinguish tumor tissue from adjacent tissue. The criteria were mentioned in Table 1 and Table 2. After filtering, the result showed that three miRNAs in HCC, including miR-122-5p, miR-182-5p, and miR-199b-3p, as well as four miRNAs in iCCA, including miR-139-3p, miR-148a-3p, miR-221-3p, and miR-222-3p, were identified as candidate miRNAs. Then, all selected miRNAs were validated in a serum sample that contained 60 healthy controls, 60 chronic hepatitis B virus (CHB) and 70 HCC patients, as well as 18 patients in each healthy and iCCA group. Baseline characteristics in the HCC and iCCA cohorts were shown in Table. 7 and Table. 8, respectively.

Relative expression level of miRNA was conducted by RT-qPCR. The findings revealed that miR-122-5p (Fig. 19A) was found to be significantly high expression (p-value = 0.0001) in HCC when compared to healthy controls. In addition, the expression level of miR-122-5p (Fig. 19A) in HCC group was significantly higher than CHB group (p-value = 0.0186). Moreover, the expression level in CHB group between was increased when compared to healthy controls (p-value = 0.0212). For the miR-182-5p expression level (Fig. 19B), CHB and HCC group were significantly elevated when compared to a healthy group (p-value = 0.0024 and p-value = 0.0023). In contrast, there was no differential expression in HCC compared to CHB (p-value = 0.4512). The expression level of miR-199b-3p (Fig. 19C) was obviously increased in HCC when compared to non-HCC patients (p-value = 0.0009). Similarly, miR-199b-3p expression in CHB (Fig. 19C) was significantly lower than a healthy controls (p-value < 0.0001). However, the expression of miR-199b-3p was not different between HCC patients and healthy controls. On the other hand, all relative expression levels of

miR-139-3p (Fig. 19D), miR-148a-3p (Fig. 19E), miR-221-3p (Fig. 19F), and miR-222-3p (Fig. 19G) in iCCA were shown to be significantly higher expressed than in a healthy group with p-value = 0.0002 (Fig. 19D) and p-value < 0.0001 (Fig. 19E, 19F, 19G).

Table 5   *A list of the significant miRNAs in HCC from the TCGA database*

No.	miRNA	Regulation	AUC	Log$_2$ fold change	*P*-adj
1	hsa-miR-424-5p	Down	0.973253	-0.55083	0.006874
2	hsa-miR-93-5p	Up	0.929866	0.615957	0.000476
3	hsa-miR-101-3p	Down	0.92578	-1.47632	1.14E-12
4	hsa-miR-195-5p	Down	0.907796	-1.37386	8.04E-13
5	hsa-miR-10b-5p	Up	0.894194	2.035803	1.77E-13
6	hsa-miR-130a-3p	Down	0.885188	-0.33848	0.036419
7	hsa-miR-125b-5p	Down	0.884597	0.920317	0.00000798
8	hsa-miR-29c-3p	Down	0.884328	-0.75338	0.000698
9	hsa-miR-199b-3p	Down	0.86172	-1.74465	1.17E-13
10	hsa-miR-145-5p	Down	0.851505	0.835439	0.00027
11	hsa-miR-221-3p	Up	0.840995	0.669439	0.000103
12	hsa-miR-25-3p	Up	0.821828	0.946785	5.24E-08
13	hsa-miR-142-5p	Down	0.819704	0.419541	0.022037
14	hsa-miR-34a-5p	Up	0.812419	0.336135	0.056347
15	hsa-miR-122-5p	Down	0.806909	-1.76561	5.22E-15
16	hsa-miR-27b-3p	Down	0.802823	0.663123	0.002163
17	hsa-miR-29a-3p	Down	0.778602	-1.14384	0.00000279
18	hsa-miR-182-5p	Up	0.776909	2.185984	9.72E-18
19	hsa-miR-19a-3p	Up	0.768495	0.717712	0.0016
20	hsa-miR-125a-5p	Down	0.742151	0.721137	0.000289
21	hsa-miR-148a-5p	Down	0.739005	0.448066	0.021112
22	hsa-miR-375-3p	Down	0.735726	0.845915	0.001288
23	hsa-miR-22-3p	Down	0.730645	2.640444	4.18E-17
24	hsa-miR-30b-5p	Down	0.704435	0.485686	0.010268
25	hsa-miR-215-5p	Down	0.677608	-0.46525	0.052709
26	hsa-miR-191-5p	Down	0.672043	0.723017	0.007136
27	hsa-miR-128-3p	Up	0.666183	-0.50236	0.010827
28	hsa-miR-107	Up	0.655591	-0.74394	0.000392

No.	miRNA	Regulation	AUC	Log$_2$ fold change	P-adj
29	hsa-miR-146a-5p	Down	0.653683	0.511385	0.006254
30	hsa-miR-143-3p	Down	0.642769	0.756258	0.004027
31	hsa-miR-24-3p	-	0.618414	-0.47209	0.0000708
32	hsa-miR-92a-3p	-	0.592984	1.030786	0.00000153
33	hsa-miR-21-3p	-	0.584866	1.173702	4.22E-08
34	hsa-miR-455-5p	-	0.557258	-0.432	0.02573367
35	hsa-miR-23a-3p	-	0.548629	-0.40967	0.01223236
36	hsa-miR-181c-5p	-	0.522339	-0.44623	0.01794913
37	hsa-miR-193b-3p	-	0.507903	0.903419	0.0000501
38	hsa-miR-98-5p	-	0.480349	0.615998	0.0011144

Table 6   A list of the significant miRNAs in iCCA from the TCGA database

No.	miRNA	Regulation	AUC	Log$_2$ fold change	P-adj
1	hsa-miR-139-3p	Down	1	-1.74575	0.01977232
2	hsa-miR-148a-3p	Down	1	-3.76018	0.0000745
3	hsa-miR-222-3p	Up	1	2.299301	0.00000518
4	hsa-miR-454-3p	Up	1	0.671113	0.02669631
5	hsa-miR-101-3p	Down	0.990741	-1.01052	0.04366867
6	hsa-miR-221-3p	Up	0.984568	2.277658	0.0000266
7	hsa-miR-92b-3p	Up	0.984568	0.834595	0.00435353
8	hsa-miR-23a-3p	Up	0.978395	1.189697	0.04366867
9	hsa-miR-181a-5p	Up	0.916667	0.475185	0.02052435
10	hsa-miR-93-5p	Up	0.904321	0.814769	0.04272699
11	hsa-miR-27b-3p	Down	0.87963	-0.96634	0.01189852
12	hsa-miR-135b-5p	Up	0.867284	3.693336	0.0000284
13	hsa-miR-142-3p	-	0.70679	1.253145	0.00088871
14	hsa-miR-212-5p	-	0.688272	1.952208	0.00354087
15	hsa-miR-155-5p	-	0.660494	2.395681	0.00000042
16	hsa-miR-29b-3p	-	0.617284	0.982879	0.02994311
17	hsa-miR-375-3p	-	0.540123	-4.64156	0.00000473

Table 7 *Baseline characteristic for validation in HCC cohorts*

Data are presented as median (IQR) or n (%).

Baseline characteristics	Healthy controls (n = 60)	Patients without HCC (n = 60)	Patients with HCC (n = 70)	P
Age (years)	54 (49.0-58.0)	54 (45.0-63.0)	58.5 (49.0-68.0)	0.020*
Gender (Male)	11 (18.3)	24 (43.6)	59 (84.3)	<0.001*
Total bilirubin (mg/dL)		0.8 (0.6-1.1)	1.0 (0.6-1.4)	0.621
Serum albumin (g/dL)		4.4 (3.9-4.5)	3.6 (3.2-4.0)	<0.001*
Aspartate aminotransferase (IU/L)	19.5 (16.0-24.0)	22 (16.3.0-29.0)	54.5 (35.5-119.3)	<0.001*
Alanine aminotransferase (IU/L)	16 (12.3-21.0)	23 (16.0-40.0)	42 (27.0-65.8)	<0.001*
Alkaline phosphatase (IU/L)	64 (53.0-77.0)	63 (54.0-76.5)	119 (82.5-174.0)	<0.001*
Platelet count (10^9/L)	258 (229.0-306.0)	239 (198.0-277.0)	170 (114.0-254.3)	<0.001*
Alpha fetoprotein (IU/mL)		2.5 (2.1-3.7)	105.3 (7.7-2979.4)	<0.001*
Presence of cirrhosis		0 (0.0)	56 (84.8)	<0.001*
BCLC stage (0-A/B/C-D)			18 (27.7)/24 (36.9)/23 (35.4)	

Table 8   Baseline characteristic for validation in iCCA cohorts

Baseline characteristics	Healthy controls (n = 18)	Patients with iCCA (n = 18)	P
Age (years)	52 (48.8-57.5)	64 (58.8-77.3)	<0.001*
Gender (Male)	9 (50)	11 (61.1)	0.502
Total bilirubin (mg/dL)		0.8 (0.5-1.3)	
Serum albumin (g/dL)		3.5 (3.1-4.1)	
Aspartate aminotransferase (IU/L)	18.5 (15.0-23.3)	162.5 (29.5-426.5)	<0.001*
Alanine aminotransferase (IU/L)	17 (12.0-21.0)	93.5 (35.0-338.0)	<0.001*
Alkaline phosphatase (IU/L)	60 (51.0-84.0)	90 (68.5-110.8)	0.003*
Platelet count (10^9/L)	247.5 (234.0-290.5)	200.5 (150.0-242.8)	0.001*
Alpha fetoprotein (IU/mL)		3.2 (2.2-5.6)	
CA19-9 (U/mL)		14.8 (5.5-538.0)	
CEA (ng/mL)		4.1 (2.7-7.2)	
Presence of cirrhosis		2 (25.0)	

Data are presented as median (IQR) or n (%)

REFERENCES

1. McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of Hepatocellular Carcinoma. 2021;73(S1):4-13.

2. Sia D, Villanueva A, Friedman SL, Llovet JM. Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. Gastroenterology. 2017;152(4):745-61.

3. Aukkanimart R, Boonmars T, Sriraj P, Sripan P, Songsri J, Ratanasuwan P, et al. Carcinogenic Liver Fluke and Others Contaminated in Pickled Fish of Northeastern Thailand. Asian Pac J Cancer Prev. 2017;18(2):529-33.

4. Dasse KD, Lander MJ, Novelli PM. Chemoembolization With Drug-Eluting Beads for the Treatment of Hepatocellular Carcinoma. J Adv Pract Oncol. 2016;7(7):764-78.

5. Bridgewater J, Galle PR, Khan SA, Llovet JM, Park J-W, Patel T, et al. Guidelines for the diagnosis and management of intrahepatic cholangiocarcinoma. Journal of Hepatology. 2014;60(6):1268-89.

6. Lou J, Zhang L, Lv S, Zhang C, Jiang S. Biomarkers for Hepatocellular Carcinoma. Biomark Cancer. 2017;9:1-9.

7. Rodrigues PM, Vogel A, Arrese M, Balderramo DC, Valle JW, Banales JM. Next-Generation Biomarkers for Cholangiocarcinoma. 2021;13(13):3222.

8. Allain C, Angenard G, Clément B, Coulouarn C. Integrative Genomic Analysis Identifies the Core Transcriptional Hallmarks of Human Hepatocellular Carcinoma. 2016;76(21):6374-81.

9. Ratti M, Lampis A, Ghidini M, Salati M, Mirchev MB, Valeri N, et al. MicroRNAs (miRNAs) and Long Non-Coding RNAs (lncRNAs) as New Tools for Cancer Therapy: First Steps from Bench to Bedside. Targeted oncology. 2020;15(3):261-78.

10. Peng Y, Croce CM. The role of MicroRNAs in human cancer. Signal Transduction and Targeted Therapy. 2016;1(1):15004.

11. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. 2018;9(402).

12. Chan JJ, Tay Y. Noncoding RNA:RNA Regulatory Networks in Cancer. International journal of molecular sciences. 2018;19(5).

13. Lou K, Chen N, Li Z, Zhang B, Wang X, Chen Y, et al. MicroRNA-142-5p Overexpression Inhibits Cell Growth and Induces Apoptosis by Regulating FOXO in Hepatocellular Carcinoma Cells. Oncol Res. 2017;25(1):65-73.

14. Koduru SV, Leberfinger AN, Kawasawa YI, Mahajan M, Gusani NJ, Sanyal AJ, et al. Non-coding RNAs in Various Stages of Liver Disease Leading to Hepatocellular Carcinoma: Differential Expression of miRNAs, piRNAs, lncRNAs, circRNAs, and sno/mt-RNaS. Scientific Reports. 2018;8(1):7967.

15. Wang X, Gao J, Zhou B, Xie J, Zhou G, Chen Y. Identification of prognostic markers for hepatocellular carcinoma based on miRNA expression profiles. Life Sciences. 2019;232:116596.

16. Yang Z, Zhou L, Wu L-M, Lai M-C, Xie H-Y, Zhang F, et al. Overexpression of Long Non-coding RNA HOTAIR Predicts Tumor Recurrence in Hepatocellular Carcinoma Patients Following Liver Transplantation. Annals of Surgical Oncology. 2011;18(5):1243-50.

17. Angenard G, Merdrignac A, Louis C, Edeline J, Coulouarn C. Expression of long non-coding RNA ANRIL predicts a poor prognosis in intrahepatic cholangiocarcinoma. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2019;51(9):1337-43.

18. Ahmad KRaS. PRINCIPLE, ANALYSIS, APPLICATION AND CHALLENGES OF NEXT-GENERATION SEQUENCING: A REVIEW. 2016.

19. contributors W. Small RNA sequencing: Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia; 2021 [updated 11 July 2021. Available from: https://en.wikipedia.org/w/index.php?title=Small_RNA_sequencing&oldid=1033017891.

20. Ji F, Sadreyev RI. RNA-seq: Basic Bioinformatics Analysis. Current protocols in molecular biology. 2018;124(1):e68.

21. Mak D, Kramvis A. Epidemiology and aetiology of hepatocellular carcinoma in Sub-Saharan Africa. Hepatoma Research. 2021;7:39.

22. Cardinale V, Bragazzi MC, Carpino G, Matteo SD, Overi D, Nevi L, et al. Intrahepatic cholangiocarcinoma: review and update. Hepatoma Research. 2018;4:20.

23. Banales JM, Cardinale V, Carpino G, Marzioni M, Andersen JB, Invernizzi P, et al. Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). Nature Reviews Gastroenterology & Hepatology. 2016;13(5):261-80.

24. Tummala KS, Brandt M, Teijeiro A, Graña O, Schwabe RF, Perna C, et al. Hepatocellular Carcinomas Originate Predominantly from Hepatocytes and Benign Lesions from Hepatic Progenitor Cells. Cell Reports. 2017;19(3):584-600.

25. He G, Dhar D, Nakagawa H, Font-Burgada J, Ogata H, Jiang Y, et al. Identification of Liver Cancer Progenitors Whose Malignant Progression Depends on Autocrine IL-6 Signaling. Cell. 2013;155(2):384-96.

26. Dang HT, Budhu A, Wang XW. The origin of cancer stem cells. Journal of Hepatology. 2014;60(6):1304-5.

27. Singh G, Yoshida EM, Rathi S, Marquez V, Kim P, Erb SR, et al. Biomarkers for hepatocellular cancer. World J Hepatol. 2020;12(9):558-73.

28. Bialecki ES, Di Bisceglie AM. Diagnosis of hepatocellular carcinoma. HPB (Oxford). 2005;7(1):26-34.

29. Tshering G, Dorji PW, Chaijaroenkul W, Na-Bangchang K. Biomarkers for the Diagnosis of Cholangiocarcinoma: A Systematic Review. Am J Trop Med Hyg. 2018;98(6):1788-97.

30. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nature Reviews Gastroenterology & Hepatology. 2020;17(9):557-88.

31. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. Nature Reviews Disease Primers. 2021;7(1):6.

32. Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. Nature Reviews Cancer. 2018;18(1):5-18.

33. Gibb EA, Brown CJ, Lam WL. The functional role of long non-coding RNA in human carcinomas. Molecular Cancer. 2011;10(1):38.

34. Tiemin P, Peng X, Qingfu l, Yan W, Junlin X, Zhefeng H, et al. Dysregulation of the miR-148a–GLUT1 axis promotes the progression and chemoresistance of human intrahepatic cholangiocarcinoma. Oncogenesis. 2020;9(2):19.

35. Bharali D, Banerjee BD, Bharadwaj M, Husain SA, Kar P. Expression Analysis of MicroRNA-21 and MicroRNA-122 in Hepatocellular Carcinoma. J Clin Exp Hepatol. 2019;9(3):294-301.

36. Hu J, Wang Y-N, Song D-J, Tan J-P, Cao Y, Fan J, et al. A High-Accuracy Model Based on Plasma miRNAs Diagnoses Intrahepatic Cholangiocarcinoma: A Single Center with 1001 Samples. Diagnostics (Basel). 2021;11(4):610.

37. Gurtan AM, Sharp PA. The role of miRNAs in regulating gene expression networks. J Mol Biol. 2013;425(19):3582-600.

38. van Rooij E, Kauppinen S. Development of microRNA therapeutics is coming of age. EMBO Mol Med. 2014;6(7):851-64.

39. Zhang X, Wang W, Zhu W, Dong J, Cheng Y, Yin Z, et al. Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. International journal of molecular sciences. 2019;20(22):5573.

40. Fernandes JCR, Acuña SM, Aoki JI, Floeter-Winter LM, Muxel SM. Long Non-Coding RNAs in the Regulation of Gene Expression: Physiology and Disease. Noncoding RNA. 2019;5(1):17.

41. Illumina I. An introduction to next-generation sequencing technology. 2017.

42. Mehmood M. Use of Bioinformatics Tools in Different Spheres of Life Sciences. Journal of Data Mining in Genomics & Proteomics. 2014;5.

43. Clough E, Barrett T. The Gene Expression Omnibus Database. Methods in molecular biology (Clifton, NJ). 2016;1418:93-110.

44. Assefa AT, De Paepe K, Everaert C, Mestdagh P, Thas O, Vandesompele J. Differential gene expression analysis tools exhibit substandard performance for long non-coding RNA-sequencing data. Genome Biology. 2018;19(1):96.

45. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):559.

46. Tipney H, Hunter L. An introduction to effective use of enrichment analysis software. Human genomics. 2010;4(3):202-6.

47. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884-i90.

48. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment

and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology. 2019;37(8):907-15.

49. Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biology. 2019;20(1):278.

50. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nature Protocols. 2016;11(9):1650-67.

51. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology. 2015;33(3):290-5.

52. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15(12):550.

53. Federico A, Karagiannis T, Karri K, Kishore D, Koga Y, Campbell JD, et al. Pipeliner: A Nextflow-Based Framework for the Definition of Sequencing Data Processing Pipelines. 2019;10(614).

54. Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, et al. The multiMiR R package and database: integration of microRNA–target interactions along with their disease and drug associations. Nucleic Acids Research. 2014;42(17):e133-e.

55. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-504.

56. Zhang H, Bian C, Tu S, Yin F, Guo P, Zhang J, et al. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in human aortic dissection. BMC Genomics. 2021;22(1):724.

57. Tang G, Shen X, Lv K, Wu Y, Bi J, Shen Q. Different normalization strategies might cause inconsistent variation in circulating microRNAs in patients with hepatocellular carcinoma. Medical science monitor : international medical journal of experimental and clinical research. 2015;21:617-24.

58. Yoon S-H, Choi S-W, Nam SW, Lee KB, Nam J-W. Preoperative immune landscape

predisposes adverse outcomes in hepatocellular carcinoma patients with liver transplantation. npj Precision Oncology. 2021;5(1):27.

59. Candia J, Bayarsaikhan E, Tandon M, Budhu A, Forgues M, Tovuu L-O, et al. The genomic landscape of Mongolian hepatocellular carcinoma. Nature Communications. 2020;11(1):4383.

60. Hu X, Bao M, Huang J, Zhou L, Zheng S. Identification and Validation of Novel Biomarkers for Diagnosis and Prognosis of Hepatocellular Carcinoma. 2020;10(1942).

61. Jin Y, Lee WY, Toh ST, Tennakoon C, Toh HC, Chow PK-H, et al. Comprehensive analysis of transcriptome profiles in hepatocellular carcinoma. Journal of Translational Medicine. 2019;17(1):273.

62. Yoo S, Wang W, Wang Q, Fiel MI, Lee E, Hiotis SP, et al. A pilot systematic genomic comparison of recurrence risks of hepatitis B virus-associated hepatocellular carcinoma with low- and high-degree liver fibrosis. BMC Medicine. 2017;15(1):214.

63. Zhang J, Baddoo M, Han C, Strong MJ, Cvitanovic J, Moroz K, et al. Gene network analysis reveals a novel 22-gene signature of carbon metabolism in hepatocellular carcinoma. Oncotarget. 2016;7(31):49232-45.

64. Yang Y, Chen L, Gu J, Zhang H, Yuan J, Lian Q, et al. Recurrently deregulated lncRNAs in hepatocellular carcinoma. Nature Communications. 2017;8(1):14421.

65. Liu G, Hou G, Li L, Li Y, Zhou W, Liu L. Potential diagnostic and prognostic marker dimethylglycine dehydrogenase (DMGDH) suppresses hepatocellular carcinoma metastasis in vitro and in vivo. 2016;7(22).

66. Chan THM, Lin CH, Qi L, Fei J, Li Y, Yong KJ, et al. A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma. 2014;63(5):832-43.

67. Huang Q, Lin B, Liu H, Ma X, Mo F, Yu W, et al. RNA-Seq Analyses Generate Comprehensive Transcriptomic Landscape and Reveal Complex Transcript Patterns in Hepatocellular Carcinoma. PLOS ONE. 2011;6(10):e26168.

68. Dong H, Zhang L, Qian Z, Zhu X, Zhu G, Chen Y, et al. Identification of HBV-MLL4 Integration and Its Molecular Basis in Chinese Hepatocellular Carcinoma. PLOS

ONE. 2015;10(4):e0123175.

69. Varghese RS, Barefoot ME, Jain S, Chen Y, Zhang Y, Alley A, et al. Integrative Analysis of DNA Methylation and microRNA Expression Reveals Mechanisms of Racial Heterogeneity in Hepatocellular Carcinoma. 2021;12(1575).

70. Han J, Thurnherr T, Chung AYF, Goh BKP, Chow PKH, Chan CY, et al. Clinicopathological-Associated Regulatory Network of Deregulated circRNAs in Hepatocellular Carcinoma. 2021;13(11):2772.

71. Wojcicka A, Swierniak M, Kornasiewicz O, Gierlikowski W, Maciag M, Kolanowska M, et al. Next generation sequencing reveals microRNA isoforms in liver cirrhosis and hepatocellular carcinoma. The International Journal of Biochemistry & Cell Biology. 2014;53:208-17.

72. Hou J, Lin L, Zhou W, Wang Z, Ding G, Dong Q, et al. Identification of miRNomes in Human Liver and Hepatocellular Carcinoma Reveals miR-199a/b-3p as Therapeutic Target for Hepatocellular Carcinoma. Cancer Cell. 2011;19(2):232-43.

73. Sheng Z, Wang X, Xu G, Shan G, Chen L. Analyses of a Panel of Transcripts Identified From a Small Sample Size and Construction of RNA Networks in Hepatocellular Carcinoma. 2019;10(431).

74. Kim H-D, Jeong S, Park S, Lee YJ, Ju YS, Kim D, et al. Implication of CD69+CD103+ tissue-resident-like CD8+ T cells as a potential immunotherapeutic target for cholangiocarcinoma. 2021;41(4):764-76.

75. Ahn KS, O'Brien D, Kang YN, Mounajjed T, Kim YH, Kim T-S, et al. Prognostic subclass of intrahepatic cholangiocarcinoma by integrative molecular–clinical analysis and potential targeted approach. Hepatology International. 2019;13(4):490-500.

76. Li H, Zhou Z-Q, Yang Z-R, Tong D-N, Guan J, Shi B-J, et al. MicroRNA-191 acts as a tumor promoter by modulating the TET1–p53 pathway in intrahepatic cholangiocarcinoma. 2017;66(1):136-51.

VITA

NAME	Varunya Virushkul
DATE OF BIRTH	2 May 1998
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Chulalonhkorn University

APPENDIX

Appendix A   list of significant DElncRNA in HCC

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
1	AC003984.1	2.127316843	2.37E-45	Up
2	RP3-335N17.2	-1.231087059	3.52E-28	Down
3	CTA-14H9.7	1.122657147	4.67E-23	Up
4	RP11-176P14.2	-1.036354006	6.10E-21	Down
5	SALRNA1	0.978990859	1.51E-17	Down
6	RP11-2P2.2	1.211592542	2.91E-17	Up
7	RP11-384P7.10	1.006473909	5.77E-17	Up
8	GS1-146J4.4	-1.043839852	5.77E-17	Down
9	RP11-749H20.5	1.099718539	3.89E-16	Up
10	RP11-129O7.3	0.964826071	1.04E-15	Down
11	LINC01938	1.041438081	1.54E-15	Up
12	RP4-785P20.2	1.001156214	1.54E-15	Up
13	RP11-363F12.1	0.937772842	2.05E-14	Down
14	RP11-637O11.2	0.840380378	8.48E-14	Down
15	RP5-1139B12.5	0.933230089	1.08E-13	Down
16	RP11-146G7.5	0.897246717	1.65E-13	Down
17	LINC01790	0.895069963	6.27E-13	Down
18	GRID1-AS1	0.874967832	7.80E-13	Down
19	RP11-360O19.4	0.894406566	7.87E-13	Down
20	RP11-522N14.2	0.860804286	3.15E-12	Down
21	RP11-517A5.5	0.58469033	1.79E-11	Down
22	RP11-275B14.1	0.743629162	5.87E-11	Down
23	LINC01776	0.749302574	4.61E-10	Down
24	ENOX1-AS1	0.788741066	5.86E-10	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
25	RP11-881M11.4	0.757838496	6.87E-10	Down
26	LINC01206	0.741606374	2.54E-09	Down
27	RP11-575N24.1	0.702043828	3.70E-09	Down
28	RP11-430L3.2	-0.678939857	4.49E-09	Down
29	CTD-3253I12.1	0.656125928	7.76E-09	Down
30	RP4-790G17.7	0.658041364	2.43E-08	Down
31	RP11-286O18.2	0.649023174	2.80E-08	Down
32	CTD-2339L15.3	0.564897623	9.66E-08	Down
33	TUSC7	0.656147482	2.04E-07	Down
34	LINC01623	0.606078047	2.14E-07	Down
35	RP11-812F18.1	-0.593371976	2.82E-07	Down
36	MYO16-AS2	-0.51277856	3.04E-07	Down
37	AC000124.1	-0.525452475	3.99E-07	Down
38	LINC01911	0.590557125	4.33E-07	Down
39	PLCH1-AS1	0.565277763	4.35E-07	Down
40	RP11-498F10.1	0.578662385	5.17E-07	Down
41	CTB-180A7.6	0.551752922	1.35E-06	Down
42	NUP93-DT	0.55272749	1.45E-06	Down
43	RP11-349F21.6	0.556044935	1.77E-06	Down
44	RP11-383D22.3	0.532808975	2.24E-06	Down
45	RP11-19O2.2	0.557349138	2.93E-06	Down
46	CADM2-AS1	0.577677001	2.95E-06	Down
47	RP11-546B8.6	0.555934767	3.00E-06	Down
48	LINC01713	0.507070721	3.32E-06	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
49	RP11-25K24.3	0.540348472	3.32E-06	Down
50	CTD-2336O2.5	-0.502868101	4.89E-06	Down
51	RP1-33L1.2	0.523202949	5.61E-06	Down
52	RP11-1046B16.4	0.512921195	5.61E-06	Down
53	RP4-564M11.4	0.509905653	7.76E-06	Down
54	RP13-126P21.3	0.500984658	8.08E-06	Down
55	RP11-362H12.2	0.507925079	9.77E-06	Down
56	LINC01761	0.476442662	1.73E-05	Down
57	RP11-14J7.6	0.504000924	1.73E-05	Down
58	CTD-2311B13.5	0.509319729	2.27E-05	Down
59	RP11-1134I14.9	0.476630419	2.54E-05	Down
60	RP5-1141E15.2	0.452660576	2.54E-05	Down
61	RP11-835E18.6	-0.449907043	3.33E-05	Down
62	RBMS3-AS1	-0.424423071	4.15E-05	Down
63	MGC15885	0.484478852	4.31E-05	Down
64	RP11-164H16.3	0.543202077	4.31E-05	Down
65	AF064866.1	-0.42071755	5.02E-05	Down
66	POTEF-AS1	0.45865633	5.05E-05	Down
67	RP11-342A17.2	-0.437210368	5.05E-05	Down
68	RP11-563N6.7	-0.483516354	6.65E-05	Down
69	FAM138E	0.466118564	7.41E-05	Down
70	AP000857.3	-0.411023522	7.41E-05	Down
71	RP11-784K9.1	0.433712645	0.00010668	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
72	AC015969.3	0.44757832	0.00011764	Down
73	AC006538.9	0.450957924	0.00015973	Down
74	RP11-713K18.1	-0.438611029	0.00017794	Down
75	AP001136.2	0.440629764	0.00019483	Down
76	RP11-283G6.7	-0.4341344	0.00019511	Down
77	RP11-586D19.2	0.44586106	0.00024418	Down
78	RP11-49C20.2	0.418518259	0.00024418	Down
79	RP11-4F5.3	-0.465291851	0.00024418	Down
80	AC016994.2	-0.381499832	0.00027166	Down
81	TBX5-AS1	0.467383419	0.00033568	Down
82	NLGN1-AS1	0.427366304	0.00038616	Down
83	GS1-519E5.1	0.401395253	0.00042151	Down
84	RP11-899L11.4	-0.424352736	0.00044793	Down
85	RP11-615I2.1	0.402325423	0.00049175	Down
86	LINC01555	0.385088187	0.00050182	Down
87	CASC17	0.386879735	0.00050182	Down
88	CTD-3239E11.2	0.380821363	0.00052983	Down
89	RP11-26A3.1	0.411884312	0.00052983	Down
90	RP11-899L11.1	-0.361982375	0.00055228	Down
91	LINC02759	0.393415763	0.00059416	Down
92	RP11-621L6.2	0.371307691	0.0006134	Down
93	RP11-492O8.3	-0.357854858	0.00079	Down
94	RP3-495K2.2	0.39781384	0.00079112	Down
95	RP11-459E5.1	0.407270818	0.00080135	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
96	EML6-AS1	0.372302139	0.00090049	Down
97	NAALADL2-AS1	-0.351875743	0.00091611	Down
98	RP13-923O23.6	0.363641582	0.00096659	Down
99	CTD-2547L24.5	-0.345189111	0.00132857	Down
100	RP11-97F8.3	0.355745264	0.0015589	Down
101	LINC02338	0.35442345	0.0016383	Down
102	CTA-396D5.2	0.348772948	0.00167005	Down
103	RP11-85O21.5	-0.338401938	0.00192018	Down
104	FAM230J	0.342080877	0.00217361	Down
105	AP000431.1	0.34832911	0.00222762	Down
106	RP11-237N19.3	0.361996493	0.00239779	Down
107	RP11-109D24.2	0.356178125	0.00242945	Down
108	CHODL-AS1	0.338912196	0.00264478	Down
109	RP11-313C15.1	0.34401884	0.00264478	Down
110	LINC01299	0.345433446	0.00264478	Down
111	SIDT1-AS1	-0.32020943	0.00267636	Down
112	RP5-970A17.1	0.3424287	0.00267636	Down
113	LINC02503	0.319640024	0.00284579	Down
114	LINC02401	-0.324964191	0.00346434	Down
115	LINC02264	-0.321114198	0.00359424	Down
116	RP11-696D21.2	0.335887011	0.00408647	Down
117	RP11-335K21.1	0.339922226	0.00408647	Down
118	RP11-816B4.3	0.334997814	0.00416504	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
119	RP11-18O11.3	0.339392162	0.0046098	Down
120	LINC00901	0.338201546	0.00509875	Down
121	CTB-164N12.1	0.319111834	0.00559085	Down
122	CTC-268N12.3	0.318350513	0.0056375	Down
123	CTA-150C2.23	0.321722152	0.00676976	Down
124	FAM230H	0.303411894	0.00710867	Down
125	RP11-527D7.4	0.327383982	0.00710867	Down
126	RP5-1067M6.6	0.30855239	0.00741236	Down
127	AC007091.1	0.303384047	0.00939987	Down
128	CTD-2363C16.2	0.297633388	0.01008819	Down
129	RP11-162K11.4	0.292482964	0.01083617	Down
130	LINC02423	0.30419107	0.01158509	Down
131	RP11-80N9.1	-0.278807612	0.01158509	Down
132	RP11-172F4.8	0.299194527	0.01158509	Down
133	RP11-179G8.2	0.289591529	0.01173449	Down
134	LINC01241	0.292653936	0.01196726	Down
135	RP11-445J9.2	0.29345119	0.01196726	Down
136	RP11-379L12.1	0.290075222	0.01242766	Down
137	RP11-707P20.1	0.292196389	0.0136343	Down
138	RP11-403D15.2	-0.273429576	0.01398532	Down
139	RP11-413B19.2	-0.278260827	0.01406477	Down
140	UST-AS1	-0.280617456	0.01419736	Down
141	RP11-27G24.3	-0.278175803	0.01451576	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
142	AC090957.2	-0.263595161	0.01456423	Down
143	CACNA2D3-AS1	0.305350689	0.01456423	Down
144	RP11-418B12.1	0.285707541	0.01463969	Down
145	FABP6-AS1	0.285290996	0.01463969	Down
146	AC083864.4	0.277191759	0.01575595	Down
147	AC107079.1	0.282896421	0.0159448	Down
148	AC002551.1	0.270948009	0.0159448	Down
149	RP5-879K22.3	0.307302473	0.0159448	Down
150	MTUS2-AS2	-0.269085322	0.01606234	Down
151	RP11-428L9.2	0.282387891	0.01649728	Down
152	RP5-1166F10.2	0.295930638	0.01664043	Down
153	RP11-25J23.4	0.277824628	0.01765327	Down
154	RP11-125C10.1	0.267987906	0.01901549	Down
155	AC092669.6	0.270105455	0.0219993	Down
156	RP11-33M22.4	0.264963979	0.02572839	Down
157	RP11-455F5.4	0.250565577	0.02637991	Down
158	RP4-753M9.1	0.260951274	0.02869327	Down
159	RP11-87P13.2	0.260864291	0.02875351	Down
160	RP11-356J5.4	-0.244292065	0.02885163	Down
161	AC018693.6	0.242388131	0.03008414	Down
162	AP000708.1	0.247777735	0.03027282	Down
163	OSTM1-AS1	0.256854449	0.03057383	Down
164	RP4-672J20.3	0.255296801	0.03181852	Down
165	LINC01363	0.251795746	0.03352129	Down

Appendix A   list of significant DElncRNA in HCC (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
166	L34079.3	0.24278051	0.03398585	Down
167	RP11-56I23.2	-0.244549564	0.03433352	Down
168	RP11-528A4.5	0.248798925	0.03510835	Down
169	LINC02504	0.244787728	0.03565922	Down
170	RP5-826L7.1	-0.247688598	0.03662527	Down
171	RP11-47G4.2	0.244869581	0.03689331	Down
172	RP11-655G22.3	0.249548454	0.03689331	Down
173	RP11-211F2.1	0.248029471	0.03696105	Down
174	RP4-784A16.2	0.243968721	0.03735001	Down
175	LINC00454	0.247771716	0.03767757	Down
176	LINC01247	0.248115207	0.03849348	Down
177	RP3-400B16.3	0.250815691	0.04057665	Down
178	RP11-203M5.2	0.242634454	0.04057665	Down
179	RP11-629D22.1	0.253536319	0.04084442	Down
180	RP11-720L2.2	0.240702938	0.04174429	Down
181	RP11-376P6.4	-0.305287583	0.0418844	Down
182	CTD-3104H21.1	-0.217968766	0.04350057	Down
183	RP11-393E10.1	0.232916489	0.04484395	Down
184	RP11-177B4.2	0.234386878	0.04704924	Down
185	EFCAB6-AS1	0.258039524	0.04772097	Down
186	RP11-406H4.1	-0.227757202	0.0493441	Down

Appendix B   list of significant DEmiRNA in HCC

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
1	hsa-miR-127-3p	3.131505	2.28E-27	Up
2	hsa-miR-1260a	1.931292	1.78E-18	Up
3	hsa-miR-182-5p	2.148816	1.46E-17	Up
4	hsa-miR-22-3p	2.605001	2.44E-16	Up
5	hsa-miR-101-3p	-1.73251	3.77E-16	Down
6	hsa-miR-151a-3p	1.526867	4.11E-15	Up
7	hsa-miR-122-5p	-1.7359	1.29E-14	Down
8	hsa-miR-122b-3p	-1.81713	8.82E-14	Down
9	hsa-miR-186-5p	1.735237	8.82E-14	Up
10	hsa-miR-10b-5p	1.987331	7.21E-13	Up
11	hsa-miR-195-5p	-1.3369	1.80E-12	Down
12	hsa-miR-199a-3p	-1.63787	3.78E-12	Down
13	hsa-miR-199b-3p	-1.63787	3.78E-12	Down
14	hsa-miR-19b-3p	1.418404	5.64E-12	Up
15	hsa-miR-423-3p	1.37641	4.28E-11	Up
16	hsa-miR-320a-3p	1.79129	5.43E-11	Up
17	hsa-miR-1260b	1.309061	3.12E-10	Up
18	hsa-miR-769-5p	1.639227	2.80E-09	Up
19	hsa-let-7c-5p	-1.29789	1.40E-08	Down
20	hsa-let-7i-5p	0.811911	3.07E-08	Down
21	hsa-miR-21-3p	1.173702	4.22E-08	Up
22	hsa-miR-25-3p	0.941156	4.92E-08	Down

Appendix B   list of significant DEmiRNA in HCC (Cont.)

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
23	hsa-miR-125b-5p	1.144882	1.74E-07	Up
24	hsa-miR-29a-3p	-1.2463	3.71E-07	Down
25	hsa-miR-486-5p	1.308454	1.06E-06	Up
26	hsa-miR-92a-3p	1.030786	1.53E-06	Up
27	hsa-miR-21-5p	1.110654	1.81E-06	Up
28	hsa-miR-1307-5p	1.29129	4.67E-06	Up
29	hsa-miR-409-3p	1.34092	6.05E-06	Up
30	hsa-miR-221-3p	0.804316	6.06E-06	Down
31	hsa-miR-375-3p	1.160423	1.77E-05	Up
32	hsa-miR-125a-5p	0.913062	2.10E-05	Down
33	hsa-miR-145-5p	1.000698	2.67E-05	Up
34	hsa-miR-193b-3p	0.903419	5.01E-05	Down
35	hsa-miR-340-5p	0.647839	5.17E-05	Down
36	hsa-miR-24-3p	-0.47209	7.08E-05	Down
37	hsa-miR-3184-3p	0.847108	7.12E-05	Down
38	hsa-miR-423-5p	0.847108	7.12E-05	Down
39	hsa-miR-378a-3p	1.052515	7.12E-05	Up
40	hsa-let-7d-5p	-0.87129	7.36E-05	Down
41	hsa-miR-374a-5p	-0.74146	0.000293	Down
42	hsa-miR-93-5p	0.619664	0.000362	Down
43	hsa-let-7e-5p	-0.68511	0.000411	Down
44	hsa-miR-107	-0.70749	0.000464	Down
45	hsa-miR-532-5p	0.68058	0.000472	Down
46	hsa-miR-342-3p	-0.67169	0.000521	Down

Appendix B   list of significant DEmiRNA in HCC (Cont.)

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
47	hsa-miR-122b-5p	-0.55776	0.000552	Down
48	hsa-miR-19a-3p	0.802931	0.000618	Down
49	hsa-miR-146a-5p	0.679492	0.00095	Down
50	hsa-miR-98-5p	0.615998	0.001114	Down
51	hsa-miR-29c-3p	-0.72889	0.001679	Down
52	hsa-miR-28-5p	-0.48419	0.002277	Down
53	hsa-miR-3184-5p	0.792073	0.002417	Down
54	hsa-miR-143-3p	0.835992	0.002792	Down
55	hsa-miR-215-5p	-0.75147	0.003913	Down
56	hsa-miR-28-3p	0.521457	0.003942	Down
57	hsa-miR-27b-3p	0.644771	0.003974	Down
58	hsa-miR-151a-5p	0.496259	0.004132	Down
59	hsa-let-7a-5p	-0.74624	0.004132	Down
60	hsa-miR-361-3p	0.736127	0.004216	Down
61	hsa-miR-148a-5p	0.600555	0.005553	Down
62	hsa-miR-374a-3p	-0.59167	0.005559	Down
63	hsa-miR-30b-5p	0.509604	0.009668	Down
64	hsa-miR-30e-3p	-0.39178	0.011018	Down
65	hsa-miR-3074-5p	-0.35503	0.011163	Down
66	hsa-let-7b-5p	-0.4593	0.01176	Down
67	hsa-miR-23a-3p	-0.40967	0.012232	Down
68	hsa-miR-574-3p	-0.48264	0.012979	Down
69	hsa-miR-191-5p	0.673743	0.017333	Down

Appendix B   list of significant DEmiRNA in HCC (Cont.)

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
70	hsa-miR-100-5p	0.602049	0.017726	Down
71	hsa-miR-181c-5p	-0.44623	0.017949	Down
72	hsa-miR-361-5p	-0.45942	0.020183	Down
73	hsa-miR-126-5p	0.574943	0.021365	Down
74	hsa-miR-142-5p	0.450286	0.022878	Down
75	hsa-miR-424-5p	-0.48444	0.023621	Down
76	hsa-miR-128-3p	-0.46482	0.024539	Down
77	hsa-miR-455-5p	-0.432	0.025734	Down
78	hsa-miR-34a-5p	0.43075	0.025734	Down
79	hsa-miR-181b-5p	-0.40653	0.027054	Down
80	hsa-miR-378c	0.598797	0.028279	Down
81	hsa-let-7f-5p	-0.48818	0.041149	Down
82	hsa-miR-130a-3p	-0.34415	0.042094	Down

Appendix C   list of significant DEmRNA in HCC

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
1	ELAVL2	1.805541651	9.76E-09	Up
2	CCDC160	-1.501733931	1.32E-08	Down
3	TP53TG3C	1.891524554	1.32E-08	Up
4	TP53TG3B	1.806158889	2.78E-08	Up
5	TP53TG3E	1.744428021	2.78E-08	Up
6	TP53TG3F	1.714320542	8.47E-08	Up
7	OR51E2	1.327717236	3.86E-06	Up
8	SPHKAP	-1.058453432	6.20E-06	Down
9	AGTR2	1.438419838	9.00E-06	Up
10	FOXI2	1.334473432	9.16E-06	Up
11	MAGEB3	1.366477082	3.20E-05	Up
12	TEX51	-0.92742945	4.72E-05	Down
13	VWC2	-0.834449469	0.000421655	Down
14	RPE65	1.07501462	0.000432587	Up
15	DYNAP	1.052985363	0.000990102	Up
16	HELT	1.012945863	0.00145589	Up
17	RAX	0.974617575	0.001767231	Down
18	SLC6A5	1.032083703	0.001833554	Up
19	ZNF479	1.002225475	0.001833554	Up
20	NKX2-3	1.038846133	0.003604025	Up
21	DMRTC1B	-0.727607856	0.003604025	Down
22	LPAR4	0.843976387	0.008167448	Down

Appendix C   list of significant DEmRNA in HCC (Cont.)

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
23	OR5A2	-0.607911109	0.01320536	Down
24	IL36B	0.762038436	0.013396388	Down
25	DMRT3	0.726525943	0.017509255	Down
26	OR10G4	-0.539031093	0.028013331	Down
27	NHLH2	0.786529555	0.028094728	Down
28	FAM181B	-0.717882323	0.028094728	Down
29	IL1F10	0.706454033	0.032403164	Down
30	ESX1	0.683519478	0.03447302	Down
31	CER1	0.672504336	0.036397786	Down
32	LY6D	0.668736286	0.037231834	Down
33	OR5M11	0.599138454	0.042992861	Down
34	KRT31	0.587106838	0.04919869	Down
35	MYL10	0.571691562	0.04919869	Down
36	DMRTC1	-0.563566824	0.04919869	Down
37	KRT25	0.6003679	0.04973637	Down

Appendix D   list of significant DElncRNA in iCCA

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
1	RP11-613D13.8	4.578335184	1.31E-16	Up
2	RP11-234K24.3	2.274201465	8.50E-12	Up
3	AC012613.2	-2.122062607	1.57E-11	Down
4	MIR4290HG	-2.206622719	4.17E-11	Down
5	RP11-419K12.2	2.587964036	3.40E-10	Up
6	RP11-20O13.1	2.687388008	4.32E-10	Up
7	RP11-540O11.1	1.838834025	1.41E-08	Up
8	RP11-346D14.1	2.183374809	1.47E-08	Up
9	RP11-231I16.1	1.87390877	1.90E-08	Up
10	SPACA6P-AS	1.953652422	9.11E-08	Up
11	LINC01342	1.92046015	1.16E-07	Up
12	RP11-268J15.6	1.692858663	1.51E-07	Up
13	RP1-81D8.7	-2.103926213	3.05E-07	Down
14	RP11-685M7.3	1.672991722	7.52E-07	Up
15	AC003099.2	1.902562216	7.52E-07	Up
16	RP13-516M14.2	1.780654266	8.73E-07	Up
17	LINC01424	1.597102058	1.69E-06	Up
18	LINC02078	-1.400739414	1.86E-06	Down
19	MIR3150BHG	1.768971415	2.29E-06	Up
20	LINC01910	1.879721629	2.29E-06	Up
21	RP11-438D8.6	1.693838131	2.42E-06	Up
22	RP11-256L6.2	-1.558900934	4.49E-06	Down
23	RP11-142O6.1	1.741346771	6.46E-06	Up

Appendix D   list of significant DElncRNA in iCCA (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
24	RP11-490F3.3	1.357667141	1.15E-05	Up
25	RP5-981O7.2	1.78098707	1.40E-05	Up
26	RP11-333J10.3	1.693502459	1.53E-05	Up
27	CTC-215C12.2	1.608635875	1.65E-05	Up
28	RP11-472H7.2	1.721410819	2.13E-05	Up
29	CTC-273B12.13	1.79562569	2.13E-05	Up
30	AF129075.5	1.408590333	2.19E-05	Up
31	LINC01331	1.577225123	2.19E-05	Up
32	RP11-66D17.15	1.657725983	2.22E-05	Up
33	POLH-AS1	1.325571987	3.78E-05	Up
34	RP11-394J1.3	1.269930871	3.94E-05	Up
35	SLC26A4-AS1	1.570743855	6.77E-05	Up
36	RP11-473O4.5	1.515253317	7.51E-05	Up
37	CTD-2555A7.2	1.518866816	7.60E-05	Up
38	COL18A1-AS1	1.449130032	9.54E-05	Up
39	RP11-609L3.3	1.495005638	0.00010215	Up
40	RP11-674P19.3	1.274151062	0.00013864	Up
41	RP3-404F18.5	1.319301396	0.00013907	Up
42	RP4-644L1.2	1.51430095	0.00015461	Up
43	CTD-2587H24.16	1.274058479	0.00015751	Up
44	RP3-467K16.4	1.381941005	0.00016562	Up
45	RP1-274L7.5	1.23303424	0.00033622	Up
46	LINC02099	1.228566574	0.00033969	Up
47	RP11-1146N6.3	1.21421819	0.00034951	Up

Appendix D   list of significant DElncRNA in iCCA (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
48	RP11-748C4.1	1.343739311	0.00039677	Up
49	RP11-565O12.1	1.11139929	0.0004342	Up
50	RP11-424G14.2	-1.151331719	0.00047494	Down
51	RP11-306M1.1	1.166786405	0.00055405	Up
52	RP11-521D12.5	1.380043774	0.00055405	Up
53	ERLNC1	-1.195917831	0.00057522	Down
54	LINC02389	1.221993012	0.00057522	Up
55	C1orf147	1.262925355	0.00057522	Up
56	RP11-281P23.1	1.141656112	0.00070176	Up
57	KIZ-AS1	1.155389616	0.00078627	Up
58	RP11-793H4.2	-1.082492766	0.00100914	Down
59	RP11-1D19.1	1.066269982	0.00122626	Up
60	MED8-AS1	1.19059887	0.00141892	Up
61	CTD-3064M3.4	1.083849583	0.0018725	Up
62	CTD-2530N21.4	0.968011822	0.00202594	Down
63	RP11-360O19.4	1.084901936	0.00205013	Up
64	RP11-210M15.1	1.110644981	0.00211141	Up
65	CTD-2301A4.6	0.898190722	0.00243535	Down
66	RP11-550A5.2	1.053095165	0.00243535	Up
67	CTB-33O18.3	-0.953835997	0.0024583	Down
68	RP11-887P2.6	-0.871477482	0.00268291	Down
69	RP11-407G23.8	0.878499094	0.00304345	Down
70	XXbac-B476C20.13	0.964850889	0.00389692	Down

Appendix D   list of significant DElncRNA in iCCA (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
71	RP11-83B20.1	0.923338891	0.00431469	Down
72	RP11-325J6.2	0.94741879	0.00431469	Down
73	AP000640.10	0.90557089	0.00435192	Down
74	RP11-25H12.2	1.052787341	0.00438591	Up
75	RP3-359N14.3	0.990609073	0.00463442	Down
76	RP11-629N8.6	0.94232283	0.00517513	Down
77	AD000684.2	0.894041478	0.00562565	Down
78	RP13-317D12.3	0.955963702	0.00562565	Down
79	RP11-446E24.3	0.883697357	0.00620014	Down
80	RP4-621B10.8	0.89686777	0.00649189	Down
81	RP11-13F20.1	0.940846591	0.00649189	Down
82	RP11-223A3.1	0.957568974	0.00649189	Down
83	RP11-261P9.4	0.894021168	0.00848477	Down
84	LINC01166	0.923985649	0.00848477	Down
85	RP11-21K20.4	1.029907562	0.00848477	Up
86	RP11-275G7.2	0.83192204	0.00877944	Down
87	RP4-724E16.2	0.907101686	0.00972662	Down
88	LINC01108	0.968458138	0.0100068	Down
89	PICART1	0.886002983	0.01354042	Down
90	RP3-333H23.10	0.866961794	0.01391901	Down
91	CCDC13-AS1	0.789445564	0.01800748	Down
92	RP11-689B22.2	0.776973679	0.01946312	Down
93	LINC01338	0.832259311	0.0199654	Down
94	CTD-2201G3.1	-0.7634214	0.02167467	Down

Appendix D   list of significant DElncRNA in iCCA (Cont.)

No.	lncRNA	Log$_2$ fold change	P-adj	Regulation
95	GATA3-AS1	0.834229913	0.02277903	Down
96	CTD-2366F13.3	0.800787725	0.02351712	Down
97	AC005481.5	0.776000218	0.02365706	Down
98	RP11-466A19.1	0.759022996	0.02940477	Down
99	VIPR1-AS1	0.833838179	0.02988004	Down
100	KCNQ5-AS1	0.759861311	0.03047295	Down
101	RP11-371C18.3	0.706631997	0.03164402	Down
102	AC100830.3	-0.654529313	0.03325504	Down
103	RP11-728K20.2	0.666774627	0.03325504	Down
104	RP11-330M19.1	0.689410487	0.03583037	Down
105	RP11-429A24.1	0.743439808	0.03690722	Down
106	RP11-489G11.6	0.69446464	0.04408345	Down
107	RP1-68O2.5	0.678236831	0.04556222	Down
108	RP11-495F22.1	0.694373311	0.04819314	Down

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

Appendix E   list of significant DEmiRNA in iCCA

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
1	hsa-miR-217-5p	-9.301631976	7.02E-10	Down
2	hsa-miR-451a	-3.109505005	7.88E-08	Down
3	hsa-miR-144-5p	-3.115882729	3.55E-07	Down
4	hsa-miR-155-5p	2.395680592	4.20E-07	Up
5	hsa-miR-486-5p	-2.598671707	4.20E-07	Down
6	hsa-miR-486-3p	-2.657270174	4.20E-07	Down
7	hsa-miR-375-3p	-4.641555082	4.73E-06	Down
8	hsa-miR-222-3p	2.299301444	5.18E-06	Up
9	hsa-miR-221-3p	2.277657658	2.66E-05	Up
10	hsa-miR-135b-5p	3.693336284	2.84E-05	Up
11	hsa-miR-221-5p	1.359766139	5.08E-05	Up
12	hsa-miR-148a-3p	-3.760179864	7.45E-05	Down
13	hsa-miR-3065-5p	-2.031564759	0.000164337	Down
14	hsa-miR-664a-3p	-1.107531582	0.00039778	Down
15	hsa-miR-1246	2.807668479	0.000454061	Up
16	hsa-miR-181b-5p	0.87864933	0.000493002	Down
17	hsa-miR-30c-5p	-0.884853634	0.000888713	Down
18	hsa-miR-142-3p	1.253144837	0.000888713	Up
19	hsa-miR-148a-5p	-2.682655809	0.000942588	Down
20	hsa-miR-30e-3p	-0.792125771	0.000942588	Down
21	hsa-miR-629-5p	1.372747594	0.001459108	Up
22	hsa-miR-142-5p	1.204469996	0.001804602	Up
23	hsa-miR-421	1.826989504	0.002032641	Up

Appendix E   list of significant DEmiRNA in iCCA (Cont.)

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
24	hsa-miR-664a-5p	-1.183108715	0.002172624	Down
25	hsa-miR-212-5p	1.95220772	0.003540874	Up
26	hsa-miR-132-3p	1.206379465	0.004162262	Up
27	hsa-miR-210-3p	2.136720796	0.00435353	Up
28	hsa-miR-92b-3p	0.834595179	0.00435353	Down
29	hsa-miR-30b-5p	-1.783774706	0.007140017	Down
30	hsa-miR-21-5p	1.366022952	0.008542955	Up
31	hsa-miR-27b-3p	-0.966340409	0.011898517	Down
32	hsa-miR-30e-5p	-0.807244282	0.011898517	Down
33	hsa-miR-3615	1.2380542	0.013561568	Up
34	hsa-miR-1304-3p	1.60053687	0.013798964	Up
35	hsa-let-7b-5p	-1.288725771	0.014022387	Down
36	hsa-miR-30a-5p	-1.163400743	0.014033503	Down
37	hsa-miR-30c-1-3p	-0.846117289	0.014033503	Down
38	hsa-miR-652-3p	0.990969832	0.014033503	Down
39	hsa-miR-708-5p	1.847166424	0.018582689	Up
40	hsa-miR-21-3p	1.486497148	0.019579462	Up
41	hsa-miR-139-3p	-1.74575071	0.019772318	Down
42	hsa-miR-181a-5p	0.475185184	0.020524351	Down
43	hsa-miR-454-3p	0.671113363	0.02669631	Down
44	hsa-miR-29b-3p	0.982879128	0.02994311	Down
45	hsa-miR-106b-3p	0.950983191	0.02994311	Down
46	hsa-miR-452-5p	1.48285674	0.02994311	Up
47	hsa-miR-29c-5p	-1.181992672	0.033891093	Down

Appendix E   list of significant DEmiRNA in iCCA (Cont.)

No.	miRNA	Log$_2$ fold change	P-adj	Regulation
48	hsa-miR-671-3p	0.926741239	0.036970112	Down
49	hsa-miR-576-3p	-0.817232009	0.037253878	Down
50	hsa-miR-493-5p	1.331400772	0.038599121	Up
51	hsa-miR-30a-3p	-1.067042205	0.042108752	Down
52	hsa-miR-93-5p	0.814769436	0.042726987	Down
53	hsa-miR-23a-3p	1.189696905	0.043668666	Up
54	hsa-miR-101-3p	-1.010521038	0.043668666	Down
55	hsa-miR-148b-3p	-1.152889833	0.043668666	Down
56	hsa-miR-455-5p	-0.893990002	0.043668666	Down
57	hsa-miR-548k	0.862378982	0.043668666	Down
58	hsa-let-7a-5p	-0.655700296	0.043940058	Down

จุฬาลงกรณ์มหาวิทยาลัย

CHULALONGKORN UNIVERSITY

Appendix F   list of significant DEmRNA in iCCA

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
1	TMEM252	-2.388954606	4.84E-11	Down
2	FAM71D	2.756619599	1.32E-09	Up
3	CHRM1	2.417680731	9.29E-09	Up
4	KCNV1	2.386944401	4.91E-08	Up
5	HTR2C	2.448235537	8.48E-08	Up
6	CENPVL2	2.116086424	8.48E-08	Up
7	KIF4B	1.660561138	4.29E-07	Up
8	OPRPN	-1.789032694	4.95E-07	Down
9	CENPVL1	1.942306044	4.95E-07	Up
10	CXorf66	-1.75905064	6.77E-07	Down
11	XKR7	2.16339179	1.01E-06	Up
12	ANKK1	2.079172315	1.49E-06	Up
13	LCE2D	-1.428634884	1.72E-06	Down
14	DISP3	2.119446371	1.84E-06	Up
15	CACNG1	1.955236239	1.93E-06	Up
16	TAF1L	1.442425459	3.77E-06	Up
17	DCAF8L2	2.044755217	3.77E-06	Up
18	FOXI3	2.001244826	3.77E-06	Up
19	SNCB	1.643267536	5.69E-06	Up
20	SP8	1.844236222	6.30E-06	Up
21	XIRP1	1.869851834	6.30E-06	Up
22	TCEAL6	1.587253113	6.30E-06	Up
23	CDHR4	1.900807085	6.47E-06	Up

Appendix F   list of significant DEmRNA in iCCA (Cont.)

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
24	KCNE1B	-1.373486754	8.43E-06	Down
25	NHLH1	1.590668608	9.21E-06	Up
26	ASTL	1.658248661	1.08E-05	Up
27	TRPC3	1.914801629	1.31E-05	Up
28	FOXD4L1	1.684124345	1.69E-05	Up
29	TRPV5	1.619831723	2.04E-05	Up
30	CACNA1I	1.915061457	2.95E-05	Up
31	KRT2	1.57775118	2.95E-05	Up
32	H2AC4	1.364899294	3.04E-05	Up
33	ELAVL2	1.636077954	3.38E-05	Up
34	FOXD4L3	1.267775077	4.44E-05	Up
35	FAM181B	1.579656813	5.14E-05	Up
36	SLC6A3	1.610841077	7.04E-05	Up
37	DRP2	1.991077543	8.83E-05	Up
38	AVPR1B	1.689646216	9.29E-05	Up
39	ATOH7	-1.151497892	0.000107985	Down
40	HMX2	1.502953605	0.000125224	Up
41	CENPVL3	1.378022009	0.000129046	Up
42	TCF24	1.41464731	0.000141432	Up
43	PHF21B	1.443827371	0.000180072	Up
44	DRD2	1.473720039	0.000190962	Up
45	SLC22A13	1.40752667	0.000194244	Up
46	H2BC3	1.241159353	0.000201701	Up
47	C1QTNF8	1.455780943	0.000202886	Up

Appendix F   list of significant DEmRNA in iCCA (Cont.)

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
48	VCX3A	1.448462198	0.0002366	Up
49	FZD10	1.290435002	0.00029748	Up
50	SYT10	-1.372489709	0.000365516	Down
51	OR2L13	1.351044943	0.000418879	Up
52	ODAPH	1.35486171	0.000419511	Up
53	LMX1A	-1.103396084	0.000440901	Down
54	RBP3	1.207902355	0.000541222	Up
55	QRFPR	1.305714365	0.000615296	Up
56	AKAIN1	-1.158394158	0.000615296	Down
57	POU3F2	1.110099637	0.000652936	Up
58	MS4A15	1.18691345	0.000786472	Up
59	DCAF12L1	1.376298576	0.000805158	Up
60	KLF1	1.259342686	0.000829363	Up
61	DLX2	1.13487678	0.001279986	Up
62	OR8A1	1.232454682	0.001279986	Up
63	FCRL4	1.240695592	0.001310032	Up
64	PRR32	-0.99057009	0.00134973	Down
65	LRRC10B	1.140102426	0.001434575	Up
66	SMIM43	1.31517421	0.001660077	Up
67	FOXL2NB	1.119166756	0.001805253	Up
68	PRDM12	1.1123941	0.001882554	Up
69	CHIA	1.244404899	0.002045851	Up
70	GGTLC3	1.093170025	0.002314792	Up
71	LINGO4	-1.217008084	0.002832844	Down

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
72	RIMBP3B	1.035735501	0.002832844	Up
73	KLK15	1.062758436	0.002875736	Up
74	FAM9A	-0.971220437	0.003067811	Down
75	MC5R	1.020304975	0.003394718	Up
76	SALL3	1.054231712	0.003581966	Up
77	OR10H1	1.073069635	0.004211009	Up
78	KIAA1210	1.017274604	0.004297737	Up
79	TP53TG3F	1.03334058	0.004297737	Up
80	MOBP	1.096673507	0.004694345	Up
81	TP53TG3C	1.021614579	0.004788575	Up
82	CACNG5	1.16745305	0.004916481	Up
83	POM121L2	0.955303699	0.004922991	Down
84	MMP27	1.096159137	0.005501868	Up
85	UCN3	1.056343969	0.005793247	Up
86	DMRT3	0.930045627	0.006218357	Down
87	BRS3	0.938869259	0.006218357	Down
88	KRTAP5-5	0.970053669	0.006218357	Down
89	TGM5	1.019566306	0.006889477	Up
90	CTRB2	0.93341352	0.006889477	Down
91	LCE1E	0.905562318	0.006889477	Down
92	TP53TG3B	0.942434756	0.006986242	Down
93	OR3A2	0.936899128	0.007019622	Down
94	USP26	0.979332815	0.008078624	Down
95	FEV	0.954707899	0.008078624	Down

Appendix F   list of significant DEmRNA in iCCA (Cont.)

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
96	GDF10	1.088051077	0.011441345	Up
97	TP53TG3E	0.873737556	0.011609063	Down
98	LGI3	0.99232712	0.012199634	Down
99	UBL4B	0.914736765	0.012589208	Down
100	NKX6-1	0.880236488	0.01278753	Down
101	VAX1	0.965809275	0.013095561	Down
102	VWA5B1	0.86618776	0.013095561	Down
103	TFAP2B	0.911084686	0.01313452	Down
104	CNTN2	0.995610782	0.01313452	Down
105	MRGPRE	0.932645293	0.01313452	Down
106	C1orf158	0.956794272	0.013915453	Down
107	TKTL2	0.818235968	0.013960153	Down
108	SLC22A16	0.8668507	0.014449181	Down
109	KCNT1	0.942461141	0.014983631	Down
110	GPR31	0.901130599	0.015093934	Down
111	RP11-286H14.4	0.871156184	0.015093934	Down
112	ZNF716	0.840392105	0.015853239	Down
113	TRIM63	0.991835459	0.016234168	Down
114	ZNF157	0.838602872	0.017391587	Down
115	CDC14C	-0.810550671	0.017475285	Down
116	IL36RN	0.918230769	0.017760778	Down
117	LHFPL1	0.816647176	0.018278899	Down
118	LIPK	0.788950432	0.018278899	Down

Appendix F   list of significant DEmRNA in iCCA (Cont.)

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
119	LLCFC1	0.8317042	0.020669835	Down
120	SMIM34B	-0.755298365	0.021171129	Down
121	CALML5	0.804243993	0.021660888	Down
122	FEZF2	0.852982955	0.023128484	Down
123	OR51E2	0.809854901	0.023788638	Down
124	SLC12A3	0.789995239	0.023803423	Down
125	TCHHL1	0.825682232	0.023803423	Down
126	CCDC89	0.754057478	0.026875543	Down
127	PRR20G	0.682018184	0.029689389	Down
128	PPIAL4F	0.663975172	0.029689389	Down
129	IFNB1	0.677257489	0.030506287	Down
130	KRT77	0.724478553	0.030506287	Down
131	OR5M11	0.762116686	0.031890079	Down
132	TSGA10IP	0.736872978	0.033814359	Down
133	FGF4	0.755010464	0.037466586	Down
134	MEIOSIN	0.681468895	0.037585762	Down
135	GLB1L3	0.757630939	0.038513863	Down
136	PANX3	0.696735459	0.040563694	Down
137	CTAGE1	0.72965842	0.040976716	Down
138	NKX2-5	0.716367472	0.044987998	Down
139	AMER3	0.658425389	0.045806893	Down
140	IL20	0.645813588	0.046102199	Down
141	OR7D2	0.860961145	0.047329066	Down

Appendix F   list of significant DEmRNA in iCCA (Cont.)

No.	mRNA	Log$_2$ fold change	P-adj	Regulation
142	SLC22A14	0.74400806	0.048108887	Down
143	LENEP	0.610005296	0.048720639	Down
144	OR2L2	0.583814318	0.048720639	Down
145	POTED	0.719572919	0.048966507	Down

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Appendix G   The list of available previous studies which deposited the RNA sequencing data in GEO database

GEO dataset	Title	Cancer type	Sequencing type	Journal	Country
GSE148355	Preoperative immune landscape predisposes adverse outcomes in hepatocellular carcinoma patients with liver transplantation (58)	HCC	RNA-seq	npj Precision Oncology	South Korea
GSE144269	The genomic landscape of Mongolian hepatocellular carcinoma (59)	HCC	RNA-seq	Nature Communications	USA
GSE138485	Identification and Validation of Novel Biomarkers for Diagnosis and Prognosis of Hepatocellular Carcinoma (60)	HCC	RNA-seq	Frontiers in Oncology	China
GSE105130	Comprehensive analysis of transcriptome profiles in hepatocellular carcinoma (61)	HCC	RNA-seq	Journal of Translational Medicine	Singapore

Appendix G The list of available previous studies which deposited the RNA sequencing data in GEO database (Cont.)

GEO dataset	Title	Cancer type	Sequencing type	Journal	Country
GSE94660	A pilot systematic genomic comparison of recurrence risks of hepatitis B virus-associated hepatocellular carcinoma with low- and high-degree liver fibrosis (62)	HCC	RNA-seq	BMC Medicine	USA
GSE81550	Gene network analysis reveals a novel 22-gene signature of carbon metabolism in hepatocellular carcinoma (63)	HCC	RNA-seq	Oncotarget	USA
GSE77509	Recurrently deregulated lncRNAs in hepatocellular carcinoma (64)	HCC	RNA-seq	Nature Communications	China
GSE77314	Potential diagnostic and prognostic marker dimethylglycine dehydrogenase (DMGDH) suppress hepatocellular carcinoma metastasis in vitro and in vivo (65)	HCC	RNA-seq	Oncotarget	China

Appendix G The list of available previous studies which deposited the RNA sequencing data in GEO database (Cont.)

GEO dataset	Title	Cancer type	Sequencing type	Journal	Country
GSE33294	A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma (66)	HCC	RNA-seq	Gut	Hong Kong
GSE25599	RNA-Seq Analyses Generate Comprehensive Transcriptomic Landscape and Reveal Complex Transcript Patterns in Hepatocellular Carcinoma (67)	HCC	RNA-seq	Plos One	China
GSE65485	Identification of HBV-MLL4 Integration and Its Molecular Basis in Chinese Hepatocellular Carcinoma (68)	HCC	RNA-seq	Plos One	China
GSE176288	Integrative analysis of DNA methylation and microRNA expression reveals mechanisms of disparity in hepatocellular carcinoma (69)	HCC	sRNA-seq	Frontiers in Genetics	USA

Appendix G   The list of available previous studies which deposited the RNA sequencing data in GEO database (Cont.)

GEO dataset	Title	Cancer type	Sequencing type	Journal	Country
GSE156087	Clinicopathological-Associated Regulatory Network of Deregulated circRNAs in Hepatocellular Carcinoma (70)	HCC	sRNA-seq	Cancers	Singapore
GSE63046	Next generation sequencing reveals microRNA isoforms in liver cirrhosis and hepatocellular carcinoma (71)	HCC	sRNA-seq	The International Journal of Biochemistry & Cell Biology	Poland
GSE21279	Identification of miRNomes in Human Liver and Hepatocellular Carcinoma Reveals miR-199a/b-3p as Therapeutic Target for Hepatocellular Carcinoma (72)	HCC	sRNA-seq	Cancer Cell	China
GSE128274	Analyses of a Panel of Transcripts Identified from a Small Sample Size and Construction of RNA Networks in Hepatocellular Carcinoma (73)	HCC	RNA- and sRNA-seq	Frontiers in Genetics	China

Appendix G The list of available previous studies which deposited the RNA sequencing data in GEO database (Cont.)

GEO dataset	Title	Cancer type	Sequencing type	Journal	Country
GSE162396	Implication of CD69+CD103+ tissue-resident-like CD8+ T cells as a potential immunotherapeutic target for cholangiocarcinoma (74)	iCCA	RNA-seq	Liver International	South Korea
GSE107943	Prognostic subclass of intrahepatic cholangiocarcinoma by integrative molecular–clinical analysis and potential targeted approach (75)	iCCA	RNA-seq	Hepatology International	South Korea
GSE93366	MicroRNA-191 acts as a tumor promoter by modulating the TET1–p53 pathway in intrahepatic cholangiocarcinoma (76)	iCCA	sRNA-seq	Hepatology	China