CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Genomic Insights Into Liver Cancer: An *Insilico* Approach

by

Arifa Saddiqa

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Health and Life Sciences
Department of Bioinformatics and Biosciences

2025

I would like to dedicate this thesis to my Beloved Father Abdul Khaliq (late).

CERTIFICATE OF APPROVAL

Genomic Insights Into Liver Cancer: An *Insilico* Approach

by

Arifa Saddiqa

(MBS231001)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Muhammad Muddassar	CUI, Islamabad
(b)	Internal Examiner	Dr. Syeda Marriam Bakhtiar	CUST, Islamabad
(c)	Supervisor	Dr. Shaukat Iqbal Malik	CUST, Islamabad

Dr. Shaukat Iqbal Malik
Thesis Supervisor
May, 2025

Dr. Syeda Marriam Bakhtiar
Head
Dept. of BI and BS
May, 2025

Dr. Sahar Fazal
Dean
Faculty of Health and Life Sciences
May, 2025

Author's Declaration

I, **Arifa Saddiqa** hereby state that my MS thesis titled "**Genomic Insights Into Liver Cancer: An *Insilico* Approach**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Arifa Saddiqa)

Registration No: MBS231001

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Genomic Insights Into Liver Cancer: An *Insilico* Approach**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Arifa Saddiqa)

Registration No: MBS231001

Acknowledgement

Alhamdulillah! All praise and gratitude belong to Allah, the Lord of the worlds. I would like to express my heartfelt gratitude to my supervisor, Dr. Shaukat Iqbal Malik, for his invaluable guidance throughout my thesis journey. I am deeply thankful to my family for their unwavering love and support—my mother, sisters Noor Zeeshan, Shavana Adil, and Samya Nisa, as well as my brothers Omer Farooq and Atiq Rehman, have been my constant pillars of strength. I also want to acknowledge my beloved friend, Jaweria Malik, for her companionship and encouragement during challenging times. Additionally, I am grateful to senior Sadaf Saeed for her generous help and guidance, which have been instrumental in my progress. Thank you all for being an essential part of this journey.

(Arifa Saddiqa)

Abstract

Liver cancer is a life-threatening malignancy with high global mortality, necessitating advanced research to improve early detection and treatment strategies. This study aims to predict and analyze genes associated with liver cancer using an *insilico* bioinformatics approach. By leveraging biomedical text mining and computational analysis, potential genes were identified and functionally annotated to gain insights into their roles in liver cancer-related pathways. Understanding these molecular interactions is crucial for uncovering key biomarkers and therapeutic targets, ultimately contributing to the development of more effective diagnostic and treatment strategies. Biomedical text mining techniques, facilitated by COREMINE Medical, enabled the extraction of crucial biological entities such as genes, proteins, metabolic pathways, and disease-related terms. Gene clustering and interaction network analysis were performed using the STRING database, revealing complex molecular relationships potentially involved in liver cancer progression. Pathway enrichment analysis using KEGG identified several key pathways significantly associated with the disease. To refine the findings, functional annotation was conducted using the DAVID tool, ensuring specificity to human liver cancer. Additionally, protein-protein interaction analysis was performed using the Protein Prompt server to validate the functional significance of the identified proteins. This integrative bioinformatics approach facilitated a deeper understanding of the molecular mechanisms underlying liver cancer. The findings contribute to the growing body of knowledge in liver cancer research and provide a foundation for future studies aimed at improving early detection and developing targeted therapeutic strategies.

Contents

List of Figures

List of Tables

Abbreviations

DAVID	Database for Interaction, Visualization and Integrated Discovery
DNMTs	DNA Methyl Transferases
HCC	Hepatocellular Carcinoma
NAFLD	Non-Alcoholic Fatty Liver Disease
STRING	Search Tool for Retrieval of Interacting Genes
TACE	Trans-arterial Chemoembolization
TERT	Telomerase Reverse Transcriptase

Chapter 1

Introduction

An aberrant cell's Unrestrained growth and dispersed across body is defining characteristic of a class of diseases known as cancer. Eventually, via a process known as metastasis, these cancerous cells may migrate to other areas of body after invading and destroying healthy tissues. From skin to important organs including the liver, brain, and lungs, cancer can form in any kind of tissue [1, 2].

Cancer is a fatal illness often caused by genetic disorder aggregation and a variety of pathological changes. Liver cancer is a cancer that starts in the liver, is one of the most prevalent kinds. Hepatocellular carcinoma, the most common kind, cholangiocarcinoma, angiosarcoma, and hepatoblast are among the several forms of liver cancer [1].

Ninety percent of instances of liver cancer are thought to be of the most prevalent primary form, hepatocellular carcinoma. Hepatocytes, the primary kind of liver cells, are the source of hepatocellular carcinoma, a particular kind of liver cancer. Alcohol-related liver disease, non-alcoholic fatty liver disease (NAFLD), and chronic hepatitis B and C infections are the factors that lead to liver cancer, which is largely hepatocellular carcinoma (HCC), a dangerous and frequently fatal condition. It is the third leading cause of cancer-related fatalities worldwide. Because early signs are generally absent, the disease is frequently discovered late, which complicates therapy and highlights the importance of routine screening in high-risk population [3, 4].

Hepatocellular carcinoma, which originates from hepatocytes, and cholangiocarcinoma, which originates from the epithelial lining of the intrahepatic bile ducts, are the two primary histological forms of adult primary liver malignancies. In many developing nations, hepatocellular cancer is a common tumor in people. This illness has several important risk factors, such as ongoing infection with the hepatitis B and C viruses and other environmental variables like alcohol consumption, cigarette smoking, aflatoxin exposure, and other environmental factors. On the other hand, cholangiocarcinoma is less common; in the US, it represents just 7.7% of malignant liver tumors. Over 60% of liver tumors in northeastern Thailand are caused by cholangiocarcinoma, which is more common in some regions of Southeast Asia [5].

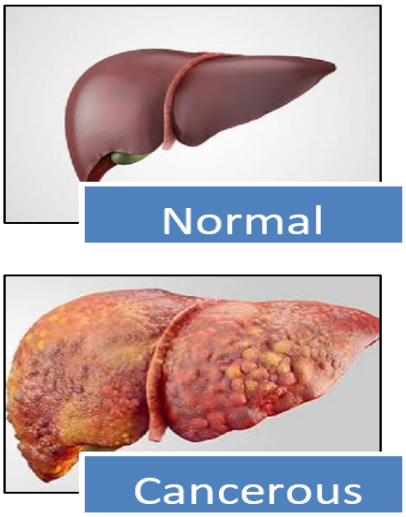


FIGURE 1.1: Healthy liver v/s Cancerous Liver [1, 2].

Over the past ten years, the incidence of liver cancer, one of the main causes of cancer-related fatalities worldwide, has steadily increased. The predicted global incidence of liver cancer in 2023 was 905,677 new cases; by 2024, the projected incidence is likely to reach 943,323 cases [6]. This concerning pattern emphasizes how urgently we need to learn more about the genetic variables that underlie the onset and spread of liver cancer in order to create more potent diagnostic and treatment approaches [6]. It is quite variable worldwide, with low frequency in Northern and Western Europe and the Americas and high rates in sub-Saharan Africa, eastern and southeast Asia, and Melanesia [5].

Compared to industrialized countries, prevalence rates in developing countries are

two to three times higher. Age-adjusted incidence rates (AAIRs) in Eastern Asia, Middle Africa, and some Central African countries range from 27.6 to 36.6 per 100,000 males, 20.8 to 38.1 per 100,000 men, and 30 to 48 per 100,000 men, respectively, making these regions the most vulnerable. With male AAIRs of less than 5.0 per 100,000, the locations with the lowest geographic risk of LC include Europe's northern regions, Australia, New Zealand, and the Caucasian communities in the Americas [7].

Environmental and genetic factors can affect the risk of liver cancer. Important environmental risk factors include cirrhosis, high alcohol use, obesity, exposure to aflatoxins, and chronic diseases such as hepatitis B or C. Additionally, vulnerability may be increased by genetic predispositions such as alpha-1 antitrypsin deficiency, hereditary hemochromatosis, and a family history of liver cancer. Although food and smoking are major lifestyle variables that increase the risk of liver cancer, certain people may be predisposed to the disease due to genetic abnormalities and specific hereditary diseases. As such, a combination of environmental factors and genetic features may lead to liver cancer [8]. TP53, CTNNB1 (beta-catenin), and TERT (telomerase reverse transcriptase) are common genes mutated in liver cancer, especially hepatocellular carcinoma (HCC). AXIN1 and ARID1A have further noteworthy alterations. These mutations frequently lead to errors in DNA repair, tumor suppression, and the activation of pathways such as Wnt/betacatenin, which in turn promotes the development of cancer [9].

HCC, or primary liver cancer, is still a challenging malignancy to treat. Geographic variations in the prevalence of viral hepatitis and liver cancer incidence are observed. Numerous staging systems have been created to account for regional preferences, the variability of primary liver cancer, and differences in respectability and transplant eligibility between regions. This diverse malignancy can be treated with a variety of modalities, and different specializations and geographical areas have different therapy guidelines for liver malignancies. As new therapeutic technologies have advanced, so too have novel treatment strategies [10]. Surgical resection, in which tumors are removed if found early, and liver transplantation, in situations of advanced liver cancer, are the available treatment options.

Transarterial chemoembolization (TACE), radiofrequency ablation, systemic medications like sorafenib, and immunotherapy drugs like nivoluma are examples of non-surgical treatments [11]. In the field of genetic research and disease analysis, bioinformatics—a multidisciplinary discipline that blends biology, computer science, and information technology—has emerged as a crucial instrument [10].

In the context of liver cancer, a devastating disease that continues to pose a significant global health challenge, the application of bioinformatics tools has proven to be invaluable in unraveling the complex genetic underpinnings of this malignancy [12]. Utilizing large-scale genomic and proteomic data, bioinformatics techniques are essential for studying genes linked to cancer. A few frequently used resources include Coremine Medical, which helps discover gene-disease connections by mining scientific literature, and Oncoming, cBioPortal, and GENT2 [13].

In order to help with structural bioinformatics, Protein Prompt Server suggests protein binding partner [14]. While DAVID makes gene functional annotation and pathway enrichment analysis easier [15]. STRING predicts protein-protein interactions that are essential for comprehending biological processes. Specific gene changes and interactions are identified with the use of these methods in liver cancer research [16]. For example, DAVID aids in mapping the biological processes important hub genes like CDC20 and CCNB2 [15].

1.1 Problem Statement

The genetic complexity of liver cancer complicates the disease. This study aims to fill the gap by using bioinformatics tools to analyze genes involved in the disease.

1.2 Aim and Objectives

The aim of this study is to predict and analyze the genes associated with liver cancer. Following are the objectives of the study:

1. To forecast genes associated with liver cancer using text mining.

2. To functionally annotate potential genes associated with liver cancer.

3. To gain insight into how expected genes function in liver cancer-related pathways

Chapter 2

Literature Review

Liver cancer, which is divided in to two categories: primary liver cancer and secondary liver cancer, is one of the most common and deadliest tumors worldwide, the third most common cause of cancer deaths among men and sixth most common among women. The most prevalent kind of primary liver cancer, hepatocellular carcinoma, as well as less frequent varieties including cholangiocarcinoma and angiosarcoma, start off in the liver cells. Cancer cells that have traveled to the liver from a different area of the body, such as the colon, breast, or lung, are referred to as metastatic or secondary liver cancer.

HCC includes a number of intriguing epidemiologic characteristics, such as dynamic temporal patterns, notable differences between men and women, racial and ethnic groupings, and geographic locations, and the existence of a number of well established environmental risk factors that may be avoidable. The molecular processes of hepatocarcinogenesis, which seldom ever happens in a healthy liver but substantially raises the risk of cancer in response to chronic liver damage at the cirrhosis stage, are now becoming well understood.

2.1 Causes and Risk Factors

Numerous risk factors contribute to its development, such as alcohol intake, non

-alcoholic fatty liver disease, and recurrent hepatitis infections (Fig 2.1). Since, liver cancer still has a high death rate despite advances in detection and therapy, it is essential to research the underlying molecular processes of the disease to develop more effective therapeutic approaches [17].

In Pakistan, chronic hepatitis B and hepatitis C virus (HCV) infections, particularly genotype 3, are the primary causes of hepatocellular carcinoma (HCC), accounting for 80% of cases. Contributing factors include unsafe medical practices like unsterilized injections and inadequate blood transfusion screening [18]. Additionally, aflatoxin exposure from contaminated food further raises HCC risk in Pakistan compared to global trends, where causes such as alcohol misuse and non-alcoholic fatty liver disease (NAFLD) are more common [19].

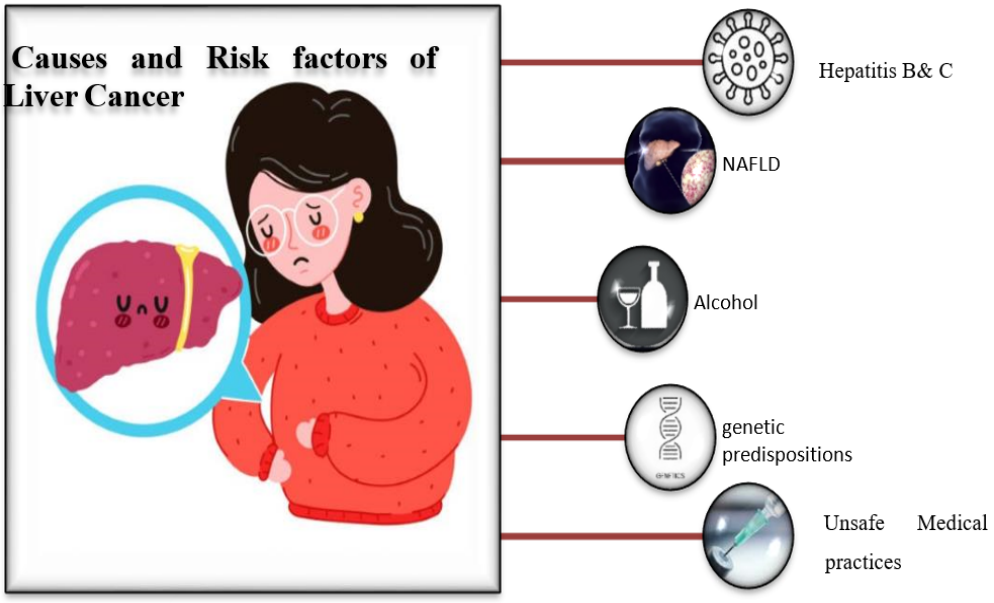


FIGURE 2.1: Etiological Factors of Liver Cancer [17, 20].

2.2 Prevalence

Primary liver cancer, mostly hepatocellular carcinoma, remains a difficult-to-treat cancer. Incidence of liver cancer varies geographically and parallels with the geographic prevalence of viral hepatitis. Globally and in Pakistan, there are important hereditary and environmental factors that contribute to liver cancer, especially

HCC. Hepatitis B (HBV) and HCV infections, which are common in Pakistan as a result of risky medical practices such reusing needles and receiving blood transfusions without screening, are the main environmental cause. These infections cause fibrosis, and chronic inflammation of the liver, which ultimately results in cancer. Pakistan has one of the highest worldwide rates of HCV-related HCC [20].

Furthermore, a major influence in the development of cancer is played by genetic factors, such as mutations that alter the liver fibrosis pathways. For example, excessive iron buildup, which adds to liver damage and carcinogenesis, puts people with certain genetic disorders, such as hemochromatosis, at higher risk of developing liver cancer [21]. The interaction between genes and the environment and the interplay of environmental factors, which include diet and other lifestyle parameters, illustrate the complexity underlying susceptibility. When genetic predispositions are made worse by toxic exposures or viral infections, gene-environment interactions play a critical role in the advancement of liver cancer in impacted populations [22].

2.3 Symptoms of Liver Cancer

Liver cancer symptoms can include vomiting, nausea, upper abdominal pain, unexplained weight loss, and loss of appetite. These symptoms are frequently more noticeable in the latter stages of the disease. Ascites, or swelling in the belly from a buildup of fluid, weariness, and jaundice— yellowing of the skin and eyes—are other typical symptoms. Due to their frequent overlap with symptoms of other liver disorders, early identification is difficult. The incidence of hepatocellular carcinoma varies greatly between nations, and this variation can be ascribed to variations in risk factors that are common in many nations. Among the risk factors for hepatocellular carcinoma include alcoholic liver disease, non-alcoholic fatty liver disease, and hepatitis B and C infections. Any underlying cause will eventually lead to liver cirrhosis and fibrosis, which will then develop into cancer. The high tumor recurrence rates and resistance to therapy make hepatocellular carcinoma difficult to treat and control [23].

2.4 Treatment

Surgery, including liver resection, is used to treat hepatocellular carcinoma in its early stages. Treatment options for advanced stages of hepatocellular carcinoma include immunotherapy, chemotherapy, and oncolytic virus usage. Novel treatment strategies have merged with the advance of new treatment modalities. Nanotechnology can be used in conjunction with conventional treatments to increase effectiveness and minimize adverse effects. Furthermore, to boost treatment effectiveness and get past resistance, immunotherapy and chemotherapy might be coupled. Not with-standing the existing treatment choices, the significant fatality rates show that the therapeutic objectives of advanced-stage hepatocellular carcinoma are not being met by the available therapy alternatives. There are several current clinical trials aimed at increasing treatment effectiveness, lowering recurrence rates, and eventually extending longevity [23]. Globally, HCC accounts for approximately 7% of all cancers, with around 750,000 new cases annually [24]. With rates of 1.6% for females and 13.1% for males, Pakistan has an exceptionally high incidence that is consistent with global trends I n the gender gap [25].

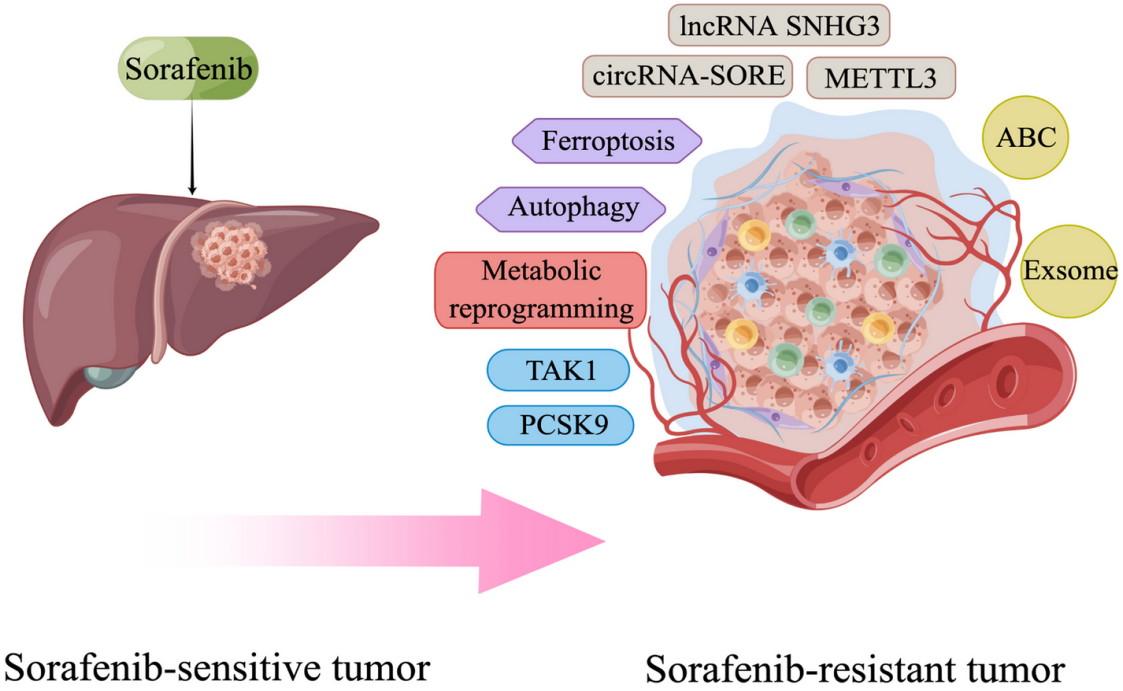


FIGURE 2.2: Mechanisms of sorafenib resistance. Signaling pathways in liver cancer: pathogenesis and targeted therapy [26].

2.4.1 Curative Treatment for Early-Stage HCC

In individuals diagnosed with small hepatocellular carcinoma (HCC) that is confined to one or a few specific areas of the liver, and who have well-preserved liver function and good overall health, curative treatment options are available. These patients often experience favorable long-term outcomes, with survival rates exceeding five years post-treatment.

2.4.2 Surgical Removal of the Tumor (Partial Liver Resection)

This treatment involves surgically removing the liver segment that contains the tumor. Where possible, surgeons use minimally invasive approaches, such as robotic or laparoscopic techniques, which can reduce recovery time. Even in cirrhotic livers, up to 60% can be safely removed, thanks to the liver's remarkable capacity to regenerate.

2.4.3 Liver Transplantation

A liver transplant may be an option when the cancer has not spread beyond the liver, there's no major blood vessel involvement, and the patient has up to three tumors, each around 5 cm or smaller. Allocation of donor livers is guided by the Model for End-Stage Liver Disease (MELD) score, with additional points granted to patients with HCC. These bonus points increase at regular intervals (usually every three months) while the patient remains on the transplant waiting list. Meanwhile, treatments such as embolization or ablation may be used to control tumor growth until a donor organ becomes available.

2.4.4 Ablation Therapy

Ablation is a minimally invasive procedure that uses needle-like probes to target

the tumor directly. These probes, inserted through the skin, apply extreme heat or cold to destroy cancer cells at the site.

2.4.5 Comparative Survival Outcomes

Liver resection, transplantation, and ablation offer similar five-year survival outcomes—commonly reported at around 70% across various treatment centers. However, liver transplantation tends to have a lower risk of cancer recurrence compared to surgical resection or ablation procedures.

2.4.6 Managing Advanced HCC

When HCC has progressed significantly—spreading to multiple parts of the liver, invading blood vessels, or metastasizing to other organs—curative treatment is no longer possible. At this stage, therapies are aimed at prolonging life and maintaining quality of life.

2.4.6.1 Embolization

This involves guiding a catheter through a groin artery into the liver to deliver drug-filled or radioactive beads directly into the tumor, effectively cutting off its blood supply and damaging cancer cells.

2.4.6.2 Chemotherapy

Among chemotherapeutic options, sorafenib remains the only drug proven to benefit advanced-stage HCC. It's taken orally once or twice a day. Ongoing clinical trials are actively testing newer agents that might improve treatment outcomes in the future.

Effective use of these therapies depends heavily on how well the liver is functioning. If liver function is too poor, these interventions may not be possible.

2.4.7 Supportive (Palliative) Care

Palliative care is a cornerstone of treatment for many HCC patients, particularly those with advanced disease or severe liver dysfunction. Its primary focus is on symptom relief—such as managing pain, fatigue, or nausea—and offering psychological and emotional support to help patients cope with the impact of a cancer diagnosis.

2.5 Liver Cirrhosis

Cirrhosis is a condition where the liver becomes damaged and scarred due to long-term injury. This scarring can interfere with the liver's ability to function properly. It's often caused by factors like excess use of alcohol, viral infections, fatty liver disease, or certain medications. Anticancer therapy for both liver and non-hepatic malignancies is limited by liver cirrhosis, the ultimate stage of all persistent liver disorders are major risk factor for the onset of hepatocellular carcinoma and Liver cirrhosis can lead to a greater risk for morbidity and death as it can impair surgical and interventional cancer treatment options, affect the pharmacokinetics of anticancer medications, exacerbate chemotherapy side effects, and make patients more vulnerable to hepatotoxicity [27].

Around the world, cirrhosis is rather frequent and can be caused by variety of illnesses, including autoimmune illnesses, cholestatic disorders, obesity, non-alcoholic fatty liver disease, excessive alcohol use, hepatitis B or C infection, and iron or copper overload.

After a prolonged inflammatory phase, healthy liver parenchyma is superceded by fibrotic tissue and growing nodules.causing portal hypertension and the development of cirrhosis. Hospitalization, a lower standard of living, and a high mortality rate are frequently outcomes of the disease's progression from an asymptomatic phase to a symptomatic phase. Disease outcomes are influenced by systemic inflammation, failure of the liver, and progressive portal hypertension [28].

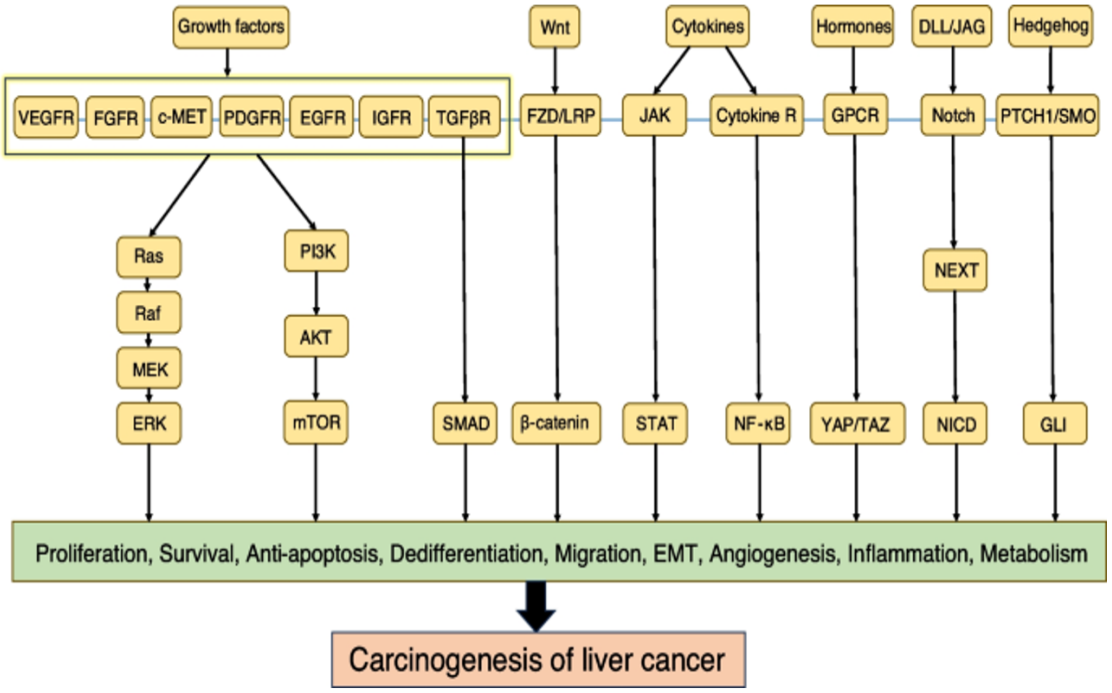


FIGURE 2.3: Signaling pathways in liver cancer [26].

2.6 Epigenetics of Hepatocellular Carcinoma

The second deadliest cancer globally is liver cancer [29]. Because of its high rate of aggressiveness and low survival rate, liver cancer is a major public health concern in many nations. Hepatocellular carcinoma (HCC) is the most prevalent kind of primary liver cancer, accounting for up to 90% of cases. Owing to the increased prevalence of related risk factors including diabetes and obesity, HCC incidence rates are rising in many nations [30].

In addition, smoking, exposure to aflatoxin in food, alcohol-induced cirrhosis, and persistent hepatitis B or C infection are known risk factors for HCC. Two of the main characteristics of cancer are abnormal molecular signaling and deregulation of gene expression, which provide cancer cells a survival advantage. It has recently been demonstrated that unfavorable modifications to epigenetic modifications might increase the cancer cells' selection advantage [31].

Given that the liver is one of the organs that continually adapts to wildly fluctu-

ating environmental circumstances, studying epigenetics in this organ is crucial. The liver is continuously undergoing repair and regeneration due to its ongoing adaptation to circadian rhythms, metabolic activities, alterations in the microbiota, and external influences including xenobiotics and viral infections [32].

The liver's Epigenome is hence very susceptible to its incredibly changing environment (Fig **??**). Therefore, hepatic Epigenome disruption is brought on by metabolic risk factors such obesity, excessive alcohol use, and viral hepatitis assaults. Epigenome modifications, including DNA methylation, chromatin modification, miRNAs, and lncRNAs, drive unchecked cell growth and proliferation, invasion, and metastasis. They also facilitate the progression of liver cancer from chronic inflammation and fibrosis to the accumulation of mutations and, ultimately, liver cancer [32].

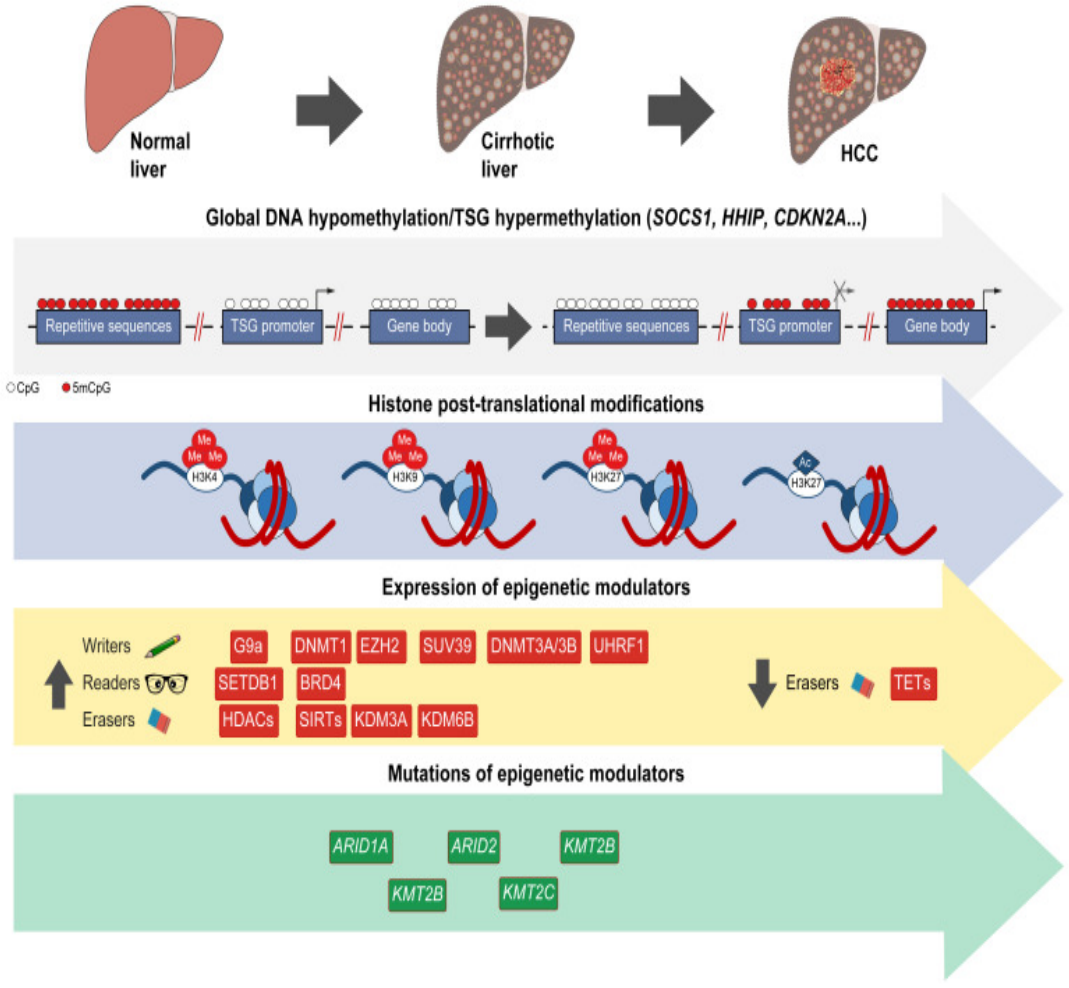


FIGURE 2.4: Epigenetic mechanisms during hepatocarcinogenesis [33].

2.7 DNA Methylation and HCC

DNA methylation is a process where in some DNA locations, cytosine bases are given methyl groups by DNA methyl transferases (DNMTs), usually at CpG sites. These CpG sites are commonly found in "CpG islands," which are areas rich in CpG sequences, often near gene promoters. Methylation in these areas can turn off gene activity, playing a key role in processes like X-chromosome inactivation [34, 35].

In stem cells, some methylation happens outside of CpG sites. Other regions near CpG islands, called CpG shores and shelves, also show different methylation patterns. DNA methylation is controlled by several DNMT enzymes, like DNMT1, which maintains methylation during DNA replication, and DNMT3A and DNMT3B, which add new methyl groups during development. Some DNMTs work together to methylate repetitive DNA regions [36]. Many malignancies, including HCC, frequently have dysregulated DNA methylation [37].

Research that correlated DNA methylation and gene expression data provided the first signs of a connection between epigenetics and cancer. Genome instability and tumor suppressor gene silence have been shown to be influenced by epigenetic modifications such as global hypomethylation and localized gene promoter hypermethylation [38]. One of the first actions in the pathophysiology of HCC is dysregulated DNA methylation, which is crucial in increasing chromosomal instability [39].

2.8 The Nature of Liver Cancer

Liver cancer, primarily hepatocellular carcinoma (HCC), is a malignancy that arises from the liver's hepatocytes—the main functional cells of the liver. This cancer is predominantly associated with chronic liver disease, including cirrhosis, which creates an environment conducive to carcinogenesis. The liver's regenerative capacity, while essential for recovery in cases of injury or disease, paradoxically

increases the risk of malignant transformation when combined with ongoing inflammation and cellular damage.

2.8.1 Cellular and Molecular Basis

Genetic and epigenetic changes in hepatocytes that accumulate over time are the hallmark of liver cancer. These changes impair important biological functions, such as the promotion of uncontrolled cell division by mutations in oncogenes (e.g., CTNNB1, implicated in the Wnt/β-catenin signaling pathway) [40].

Apoptosis can be avoided and unchecked cell proliferation can occur when TP53 or RB1, two essential tumor-suppressor genes, have loss-of-function mutations [41]. Rapid angiogenesis is a hallmark of liver cancer, as tumors create new blood vessels to support their development. In this process, the vascular endothelial growth factor (VEGF) pathway is crucial (Fig 2.5) [42].

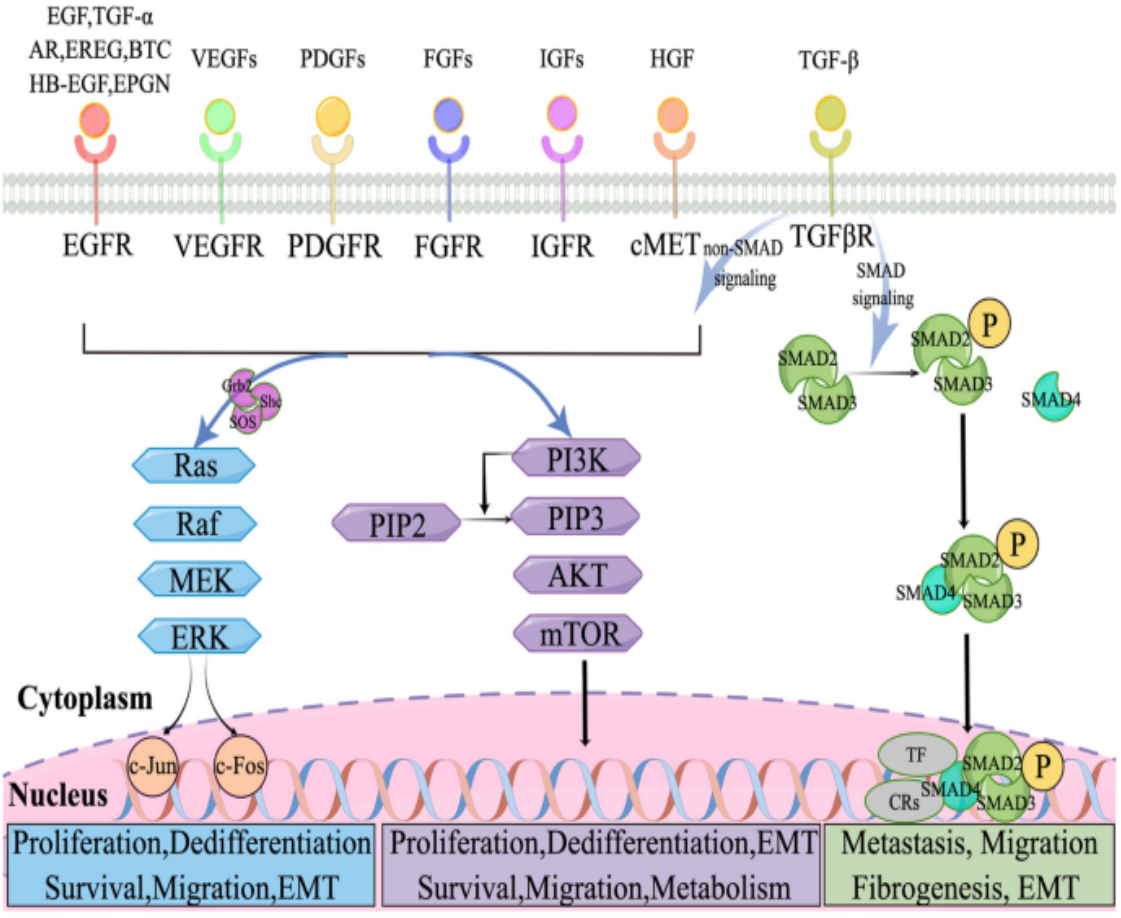


FIGURE 2.5: Growth factor receptor-related signaling pathways [26].

2.8.2 Disease Progression

Long-term liver inflammation, fibrosis, and ultimately cirrhosis are frequently seen before HCC occurs. The microenvironment of cirrhotic livers is characterized by oxidative stress, hypoxia, and inflammatory cytokines, all of which work together to promote the development of cancer from healthy liver tissue (Fig 2.6) [43].

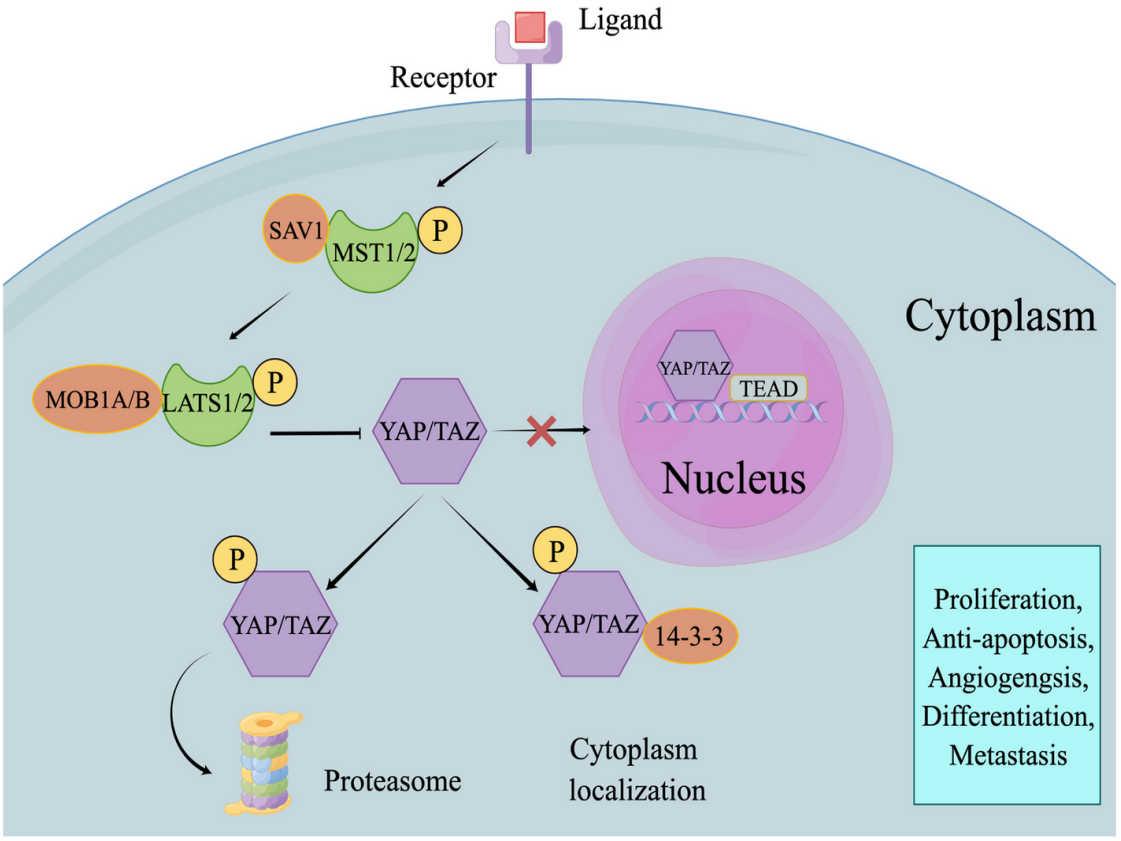


FIGURE 2.6: Hippo signaling pathway [26].

Treatment for liver cancer can become more difficult because, in contrast to many other tumors, it can spread quickly through the hepatic blood arteries. Particularly different from other cancer types, HCC tumors depend on aberrant blood artery networks to maintain their development.

A wide range of factors can cause liver cancer, such as metabolic disorders (obesity, diabetes), toxins (aflatoxins), and infections (hepatitis B and C) [44]. Tumor persistence is a result of the immune system's limited ability to detect and eradicate cancer cell due to the liver's immune-tolerant milieu [45].

2.8.3 Global Variability

Regional variations in risk variables are a major element in the geographical variation in liver cancer type. For example, the condition is frequently caused by hepatitis B virus infections that are contracted at birth or in early infancy in Asia and Africa. Non-alcoholic fatty liver disease (NAFLD) and alcohol addiction are more common causes in Western countries [46].

2.9 Genetics of Liver Cancer

Genetic and epigenetic changes interact intricately to promote the onset and spread of liver cancer, especially hepatocellular carcinoma (HCC). In the end, these alterations encourage cancer by affecting chromosomal stability, oncogenes, signaling pathways, and tumor suppressor genes. Some genetic mutations are explained in Table 2.1 [47–51].

TABLE 2.1: Genetic mutations that cause liver cancer

Genetic Mutation	**Frequency in HCC Cases**	**Impact on Cancer Progression**
TP53 Mutations	30–50% (higher in aflatoxin-exposed regions)	Loss of TP53 prevents apoptosis, allowing uncontrolled cell proliferation [47].
CTNNB1 Mutations (Wnt/β-catenin Pathway)	20–40%	Activating mutations enhance tumor growth and invasiveness [48].
AXIN1 Mutations	~10%	Disrupts Wnt/β-catenin signaling, contributing to HCC development [49].

Continue on next page

Table 2.1: Genetic mutations that cause liver cancer (Continued).

Genetic Mutation	**Frequency in HCC Cases**	**Impact on Cancer Progression**
TERT Promoter Mutations	∼60%	Enables cancer cells to maintain telomere length, leading to replicative immortality [50].
Epigenetic Alterations	Not specified	Abnormal DNA methylation and histone modifications silence tumor suppressor genes (e.g., RASSF1A, CDKN2A), promoting oncogenesis [51].

2.9.1 Chromosomal Instability and Structural Alterations

Heterozygosity loss is Inactivation of tumor suppressor genes such as TP53 and RB1 is linked to frequent LOH at chromosomal regions 1p, 8p, and 17p [52]. In some cases of hepatocellular carcinoma (HCC), oncogenes such as MYC and MET are found to be amplified. This amplification leads to the overexpression of these genes, which in turn promotes abnormal angiogenesis—the formation of new blood vessels that supply the tumor with nutrients and oxygen. Additionally, it contributes to dysregulation of the cell cycle, allowing cancer cells to proliferate uncontrollably.

2.9.2 Pathway Dysregulation

Multiple signaling pathways are altered in HCC. One such route, which promotes growth and survival, is often activated in HCC as a result of amplifications or mutations in genes such as PIK3CA [53]. Some instances of HCC are linked to

mutations in KRAS and associated components, however these are less common [54]. TGF-β signaling dysregulation has a role in metastasis and the epithelial-to-mesenchymal transition (EMT).

2.9.3 Genetic Heterogeneity and Subtypes

Molecular subtypes of HCC have been discovered by recent genomic research, and these correspond to different genetic changes: TERT promoter mutations, MYC amplification, and Wnt/β-catenin pathway activation are linked to the proliferation subtype. TP53 mutations and immune-related pathway activation are characteristics of the inflammatory subtype, which is frequently connected to hepatitis infections [55].

2.10 Bioinformatics in Liver Cancer Research

With the ability to examine massive amounts of transcriptomic and genomic data, The study of cancer has undergone a radical transformation because to bioinformatics. Bioinformatics techniques, important genes and pathways involved in the progression of liver cancer can be identified. These instruments can find possible biomarkers and treatment targets by integrating different datasets, such as methylation profiles, RNA sequencing, and DNA sequencing [56].

The Cancer Genome Atlas (TCGA), a database including comprehensive genomic data from liver cancer patients, is one of the most extensively utilized datasets in the field. In order to examine this data and find dysregulated genes and pathways unique to liver cancer, bioinformatics methods like gene set enrichment analysis (GSEA) and pathway analysis platforms like KEGG are frequently used Finding the genes that affect the development of liver cirrhosis into The type of liver cancer that is most common is HCC has been a major area of study. For example, it has been discovered that the gene CDKN3 is essential for the control of the cell cycle in HCC. Its potential as a therapeutic target is shown by the correlation between its high expression and a bad prognosis. Additional bioinformatics investigations

have revealed important pathways, including immunological responses, DNA replication, and cell cycle control, that are implicated in liver cancer [57].

2.11 Key Genes Involved in Liver Cancer

In the most frequent kind of liver cancer, hepatocellular carcinoma (HCC), recent bioinformatics driven research has found many important genes that are involved in the initiation and advancement of HCC. These include CTNNB1 and TP53, which are both often mutated in HCC. CTNNB1 mutations promote tumor development by activating the Wnt/β-catenin pathway, whereas TP53 mutations are important in cell cycle control and DNA repair. Both RB1 and CDKN2A are cell cycle regulators that are frequently inactive, which promotes uncontrolled cell division. Early-stage HCC is frequently associated with mutations in TERT, the gene that codes for telomerase reverse transcriptase, which extends the telomere and permits unrestricted cell division. Furthermore, AXIN1 and PIK3CA are connected to the PI3K/AKT and Wnt signaling pathways, respectively, which are essential for the survival and growth of tumors. By controlling the G1/S transition of the cell cycle, CDKN3 is another gene that has been found to accelerate the development of liver cancer. Patients with liver cancer who have elevated expression of CDKN3 are likely to have a bad prognosis. These genes present a prospective therapeutic target for liver cancer early detection and therapy [58]. Finding gene signatures that indicate a patient's likelihood of surviving liver cancer has been made possible in large part by bioinformatics methods. The relevance of bioinformatics in clinical applications has been demonstrated by studies that have found differentially expressed genes (DEGs) that associated with survival outcomes using transcriptome data from liver cancer patients [59].

2.12 Bioinformatics Tools for Gene Expression Analysis

In the analysis of gene expression data to find possible biomarkers and therapeutic

targets in liver cancer research, a number of bioinformatics techniques have proven helpful. For RNA-seq analysis, tools such as DESeq2 and EdgeR are widely employed. They improve statistical accuracy and normalize count data, which aids in the identification of differentially expressed genes (DEGs) in liver cancer. To show protein-protein interactions (PPI) and pinpoint important Genes linked to

the advancement of liver cancer, Cytoscape is used in conjunction with the STRING database. Utilizing data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases, this method aids in the construction of PPI networks and identifies hub genes that are essential for carcinogenesis, such as PLK1 and CDC20. Additionally, KEGG pathway analysis and GSEA (Gene Set Enrichment Analysis) offer insights into the biological processes these genes regulate, and tools like Oncomine enable researchers to confirm the mRNA expression levels of these hub genes across a variety of malignancies. A more thorough study of gene connections and pathways in liver hepatocellular carcinoma (HCC) is made possible by the helpful functional enrichment tool FunRich. These bioinformatics systems play a vital role in the integration of multiomics data and the discovery of molecular pathways that may direct therapeutic strategies for liver cancer [60]. Various proteins interact with one another in cancer cells, and this may be learned by building protein-protein interaction networks using tools such as Cytoscape. Moreover, relationships between proteins that are created by the relevant genes are predicted using the STRING database [16, 61].

2.13 Bioinformatics and Pathway Analysis in Liver Cancer

One of the most important methods in omics research is Pathway Analysis (PA), sometimes referred to as functional enrichment analysis. PA tools' primary function is to assess information obtained from large-scale devices and identify pertinent gene categories that are altered in case studies relative to a control. By doing this, Pathway analysis techniques aim to address the challenge of understanding the primary output of the majority of basic high-throughput data analysis-the

disproportionately huge lists of significant but separated genes devoid of biological context—as differential expression analysis. In order to facilitate interpretation and the subsequent creation of hypotheses, PA approaches provide experimental high-throughput biological data (HTBD) meaning. This has been accomplished by combining statistical tests, mathematical studies, and computational algorithms

with the biological knowledge that is now available in databases. In order to link genes to certain pathways, pathway analysis tools like KEGG and Reactome are essential in the study of liver cancer. This can aid in identifying important signaling pathways that are disrupted in liver cancer, notably the PI3K/AKT and Wnt/β-catenin pathways, which are frequently dysregulated as a result of mutations in genes like PTEN and CTNNB1 [62].

Researchers can now classify genes based on their molecular processes and biological functions using gene ontology (GO) enrichment analysis, which can provide more information about the biology of tumors and suggest possible therapeutic targets [63]. This is made possible by bioinformatics techniques.

Chapter 3

Material and Method

This study applies insilico tools to link predictions with potential clinical use. These tools facilitate the identification of critical genes, their interactions, and their biological roles in liver cancer progression (Fig 3.1).

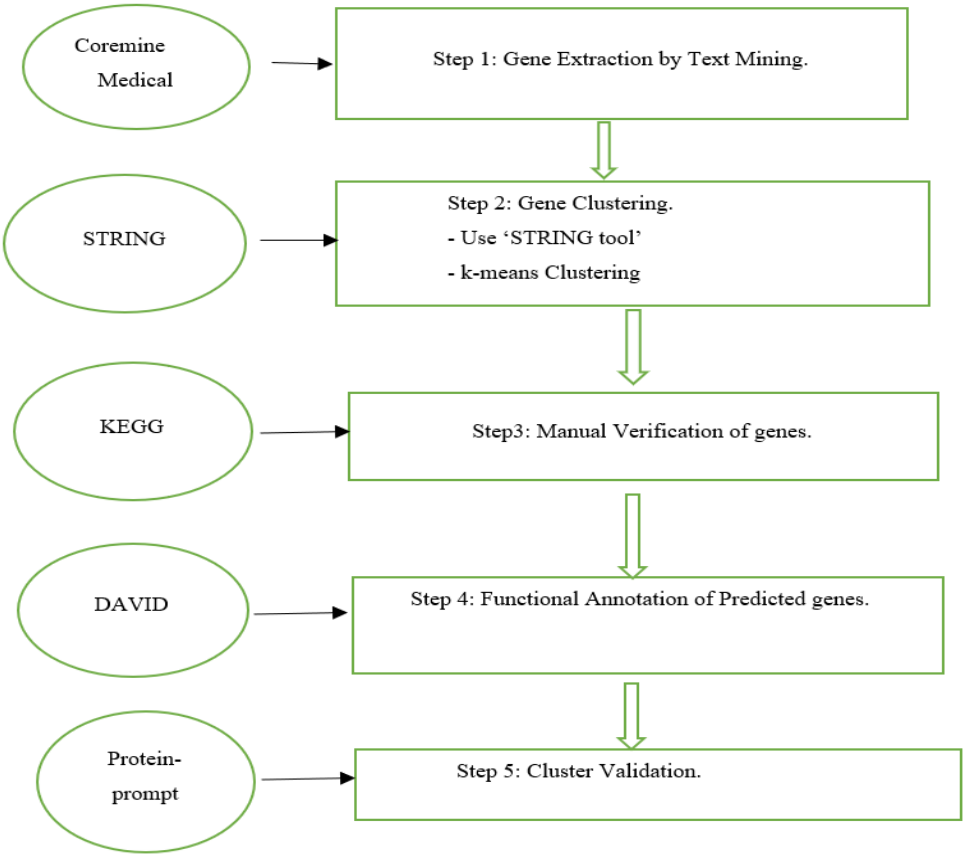


FIGURE 3.1: Mapping the path: Five bioinformatics tools unraveling liver cancer genes [63–75].

3.1 Functionality of COREMINE Medical

Text Mining Capabilities: COREMINE Medical links user queries with gene profiles using complex algorithms to provide an ordered list of the genes that are most pertinent to the input. keyword profiling improves the precision of gene identification based on user queries by assigning each gene a profile of keywords gathered from biomedical literature [63].

Provide Coremine with gene or biological data and utilize Text Mining to retrieve pertinent biological literature. Find and chart the relationships between genes and diseases. Show relationships and networks of genes and sort important genes accordings to the connections they make.

3.1.1 Investigating Links in Biomedical Literature

There are several technologies available for text analysis in the biomedical literature. A range of information was extracted from the published academic papers using a biomedical text mining tool called COREMINE medical. This data covered diseases, medications, processes, genes, proteins, and MeSH words. The findings that were found were downloaded as a text document [64].

3.1.2 Gene Occurrence via Text Mining

Text mining is the task of extracting meaningful information from text, by identifying connections between elements that don't exhibit obvious relationships [65, 66]. However, a significant portion of the text mining technique centers on the crucial pre-processing step of organizing the document collections since it concentrates on unstructured data. This is due to how crucial the pre-processing phase is. To achieve this, methods including information extraction (IE), word extraction, and text categorization are used. One of the procedures in the text mining process, which consists of many steps, is pre-processing the document collection [67, 68].

3.2 STRING

For the study of gene networks and interactions in biomedical research, the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database is an invaluable tool. Based on a range of evidence sources, such as experimental data, text mining, and co-expression patterns, it offers a thorough summary of anticipated protein-protein interactions (PPIs).

To use STRING, enter the names of liver cancer-related genes. STRING will generate a protein-protein interaction network, showing direct and indirect associations. You can adjust confidence levels and export interaction maps for further studies [69]. The objective of the STRING database is to offer an integrated and comprehensive evaluation of protein-protein interactions, including both functional and direct ties [70].

3.3 Clustering in Associated Genes of Liver Cancer Verified by KEGG

A bioinformatics tool called KEGG (Kyoto Encyclopedia of Genes and Genomes) may be used to comprehend the uses and capabilities of organisms and cells from both a high-level and genomic viewpoint. It is an independent, comprehensive resource that includes genetic, chemical, and network data along with cross-references to other external databases. It has all the building blocks (genes and chemicals) and interaction network wiring diagrams needed to model cellular operations. Place your gene list into programs like KEGG Mapper to map the genes into KEGG pathways. Insights into the functions of these genes in liver cancer will be provided by clustering them based on their participation in biological pathways [71].

Clustering is a standard technique with numerous algorithms (such as k-means, hierarchical methods and density based methods) data mining and machine lear-

ning approach that groups together comparable data points according to certain attributes. One of the goals of breaking a dataset into subsets or clusters is to create a situation where the data points within each cluster are more similar to one another than the data points within other clusters [72]. To automatically identify these groups, a variety of clustering algorithms are employed, including as K-means, hierarchical clustering, and DBSCAN [73].

3.4 Functional Annotation of Liver Cancer Genes by DAVID

The DAVID Resources now offer new tools and functions for any uploaded gene list, in addition to the standard gene-term enrichment analysis. Users may combine similar and redundant words together, look for interesting and relevant genes or terms, see genes dynamically from their lists on bio-pathways, and compress big gene lists into gene functional groupings with the use of these tools and functionalities.

The predicted genes are functionally annotated by DAVID. Utilize the DAVID gene list tool by uploading your gene list. To acquire functional annotation and uncover biological processes and pathways connected to the developing of the liver cancer, choose the pertinent species and run GO and KEGG enrichment analysis [74].

3.5 Cluster Validation through Protein-Protein Interaction

Physical linkages between two or more proteins are referred to as protein-protein interactions (PPIs), which play a critical role in facilitating various biological processes within cells, including signal transduction, immune responses, and metabolic pathways. Understanding PPIs is essential for deciphering the molecular mechan-

isms underlying both normal cellular functions and disease states.

Prompt server, a versatile online platform, allows researchers to seamlessly analyze and visualize protein sequences, structures, and interactions, providing invaluable insights into the structural and functional implications of genetic alterations associated with liver cancer [75].

Chapter 4

Results

4.1 Retrieval of Genes by COREMINE

COREMINE tool was used to extract the genes in the first phase. To do this, the COREMINE tool's search box was opened, and the disease—liver cancer—was entered. Next, the associations from 2020 to 2024 after selecting the extracted associations option in the COREMINE interface were extracted. A list of 2000 genes linked to liver cancer was produced by COREMINE using the query Liver Cancer (Genes/Proteins). The names of the related genes, illnesses, and descriptions were compiled into an Excel list and can be noted in Table 4.1.

TABLE 4.1: Genes associated with liver cancer and their significance values retrieved from the Coremine Medical database. Lower significance values indicate stronger associations between the gene and liver cancer.

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
AFP	1.33E-06	MIR7-3HG	2.62E-06	GPC3	1.27E-05
LAMTOR5	3.28E-05	TP53	4.36E-05	RSF1	4.55E-05
GOLM1	4.97E-05	TRIM26	6.48E-05	HCC	7.41E-05
CTNNB1	7.93E-05	TSC1	8.72E-05	RARB	9.2E-05
FAM126A	9.71E-05	CASP3	1.03E-04	MYC	0.000104
ALB	0.000106	EPCAM	0.000114	CDH1	0.000114

Table 4.1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
STAT3	0.000114	TAT	0.000115	CCND1	0.000121
KRT19	0.000122	PNPLA3	0.000133	CDKN1A	0.000133
PDCD1	0.000145	F2	0.000148	VEGFA	0.000149
TCEAL1	0.000157	H3F3AP6	0.000157	SHC3	0.000163
CYP1A1	0.000166	MET	0.000174	BIRC5	0.000181
MKLN1-AS	0.000186	RPL17	0.000188	PROM1	0.00019
MTOR	0.000194	AHR	0.000201	CD274	0.000212
TERT	0.000216	TM6SF2	0.000218	CASP9	0.000236
TM4SF5	0.000238	CCNB1	0.000247	HNF4A	0.000251
FGF19	0.000254	CDH2	0.000255	HIF1A	0.000258
PRRT2	0.000264	MMP9	0.00027	HGF	0.000272
MMP2	0.000293	GPC1	0.000302	VIM	0.000302
CDK1	0.000312	SLC10A1	0.000314	PTEN	0.000319
GGTLC1	0.000342	G6PC	0.000346	HNF1A	0.000349
GPM6A	0.000349	AKR1B10	0.000351	GNMT	0.000356
BCL2L1	0.000358	CASP8	0.000365	NRAV	0.000367
PARP1	0.000376	CDK2	0.000384	IFNL3	0.000401
JUN	0.000402	MLX	0.000403	TMCC1-AS1	0.000411
MIR122	0.000412	CDK4	0.000422	PSMD10	0.000429
COL11A2	0.000434	CCNA2	0.000435	SP1	0.000454
HEIH	0.00046	AXIN1	0.00046	PCNA	0.00046
BAX	0.000463	E2F1	0.000466	DLC1	0.000477
MAT2A	0.000479	CD81	0.000482	PSG1	0.000489
DAND5	0.000492	KDM4A-AS1	0.000494	CCL4	0.000495
SQSTM1	0.000508	ARNT	0.000512	HDGF	0.000517
CYP2E1	0.00053	EZH2	0.000534	IGF2	0.000543
ABCB1	0.000544	SREBF1	0.000548	KDR	0.000551
DLEC1	0.000565	TXK	0.000568	TNFSF10	0.000576
CXXC1	0.000583	PPARA	0.000599	SERPINA1	0.000605

Table 4.1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
SEC14L2	0.012374	CSTF2	0.012388	HN1	0.01239
HEG1	0.01239	GPX1	0.012416	LRP6	0.012425
DPYD	0.012456	TXNDC5	0.012474	ACAT1	0.012476
FGFR1	0.0125	NOP58	0.012523	TUBA1C	0.012523
DHX33	0.012539	MIR149	0.012539	ENGASE	0.012543
KLRC1	0.012545	HINT1	0.012576	IL18RAP	0.012596
HMGB2	0.012598	COMMD1	0.012605	LAYN	0.012625
TDGF1	0.012647	PES1	0.012648	UPF3B	0.012648
CRKL	0.012678	SLC41A3	0.012681	EFNA4	0.012693
HNF4G	0.012693	BMP6	0.012728	JUND	0.01273
CORO1A	0.012753	HSF1	0.012795	CDK5R2	0.012857
ACYP2	0.012861	TTN-AS1	0.012861	MIR137	0.012861
TRIM27	0.012861	ACSS1	0.012864	CENPQ	0.01288
NRSN2	0.01288	TRIM45	0.01288	SAAL1	0.01288
MIR424	0.01288	CERS6-AS1	0.01288	PDGFC	0.012882
VDAC1P5	0.012889	TRIM44	0.012889	CRYL1	0.012892
LINC00707	0.012892	CLEC4M	0.012904	COLEC10	0.01291
RECK	0.012946	NR2C2	0.01298	TRIM31	0.013011
CFL1	0.013021	EIF2S3	0.013044	PLTP	0.013047
APOM	0.013052	CISH	0.013082	CPT1A	0.013372
AKAP12	0.01338	UNC119	0.01338	DDX5	0.013381
SUMO1	0.013404	NAP1L1	0.013413	DR1	0.013422
ASF1B	0.013437	CAV1	0.013438	CCND2	0.013442
CTCFL	0.013445	HGFAC	0.013476	TMX2	0.013484
DDX58	0.011431	NCOA1	0.011469	SDHC	0.011479
RTKN2	0.011489	CCDC137	0.011489	ENG	0.011492
ADAM10	0.011496	ETV4	0.011546	QSOX1	0.011555
RFX1	0.011555	USF2	0.011567	MAPK9	0.011608
MFI2	0.01161	PCBP1	0.011612	IFNB1	0.011631

Table 4.1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
SLC4A4	0.011649	MCM10	0.011654	RBMY1A1	0.011656
EFNA3	0.011674	MIRLET7B	0.011676	NXF3	0.011684
PLGLA	0.011684	PCBP1-AS1	0.011684	LINC00426	0.011684
LIFR	0.0117	HNRNPK	0.011722	S100A4	0.01173
SULT1C2	0.011745	ACRBP	0.011745	SLC25A5	0.011745
ING3	0.011757	NDUFA13	0.011759	SLC17A2	0.011832
FBXW10	0.011838	TSPAN13	0.011838	HAGLROS	0.011838
SPOCK1	0.01184	VASH2	0.011847	ERRFI1	0.011851
NEK6	0.011859	WNT1	0.011873	SIRT5	0.011885
MAP3K7	0.011898	CCND3	0.011907	SLC50A1	0.01193
ING1	0.011954	RNGTT	0.011969	CUL4A	0.011973
VDAC3	0.012007	DAB2IP	0.012012	NFKBIA	0.012042
B3GNT3	0.012063	NCR3	0.012098	CD82	0.012098
CXCL8	0.01211	PDK2	0.012132	CBX6	0.012155
RHPN1	0.012155	ALDH6A1	0.012165	NPC1L1	0.012171
PPAP2C	0.012188	VDAC2	0.012215	MARCKSL1	0.012215
CXCL6	0.012217	SLC47A1	0.012218	CAT	0.01225
FLVCR1	0.012258	RMST	0.012258	FOXN3	0.012259
RPS15A	0.012259	EIF2S1	0.012272	MAGEB3	0.012279
OR1A2	0.012279	COMMD8	0.012279	DHRS4-AS1	0.012279
TATDN1	0.012279	LINC00239	0.012279	DEFB106B	0.012279
DLG1-AS1	0.012279	MIR4435-2	0.012279	BANCR	0.012279
ATXN7L3	0.01228	RNF187	0.01228	BOLA2	0.01228
TP53INP1	0.012283	CENPA	0.012319	SLC24A5	0.01232
WIF1	0.012339	ST6GALNAC4	0.012353	CLEC3B	0.012373
DDX58	0.011431	NCOA1	0.011469	SDHC	0.011479
RTKN2	0.011489	CCDC137	0.011489	ENG	0.011492
ADAM10	0.011496	ETV4	0.011546	QSOX1	0.011555

(Remaining genes can be noted in Appendix)

4.2 Network Creation

To complete this step, access to the STRING database was needed. A list of 2000 Liver Cancer genes are included here. Accordingly, humans were chosen as the organisms in this case. There were 1569 nodes in the network that was created after that (Fig 4.1).

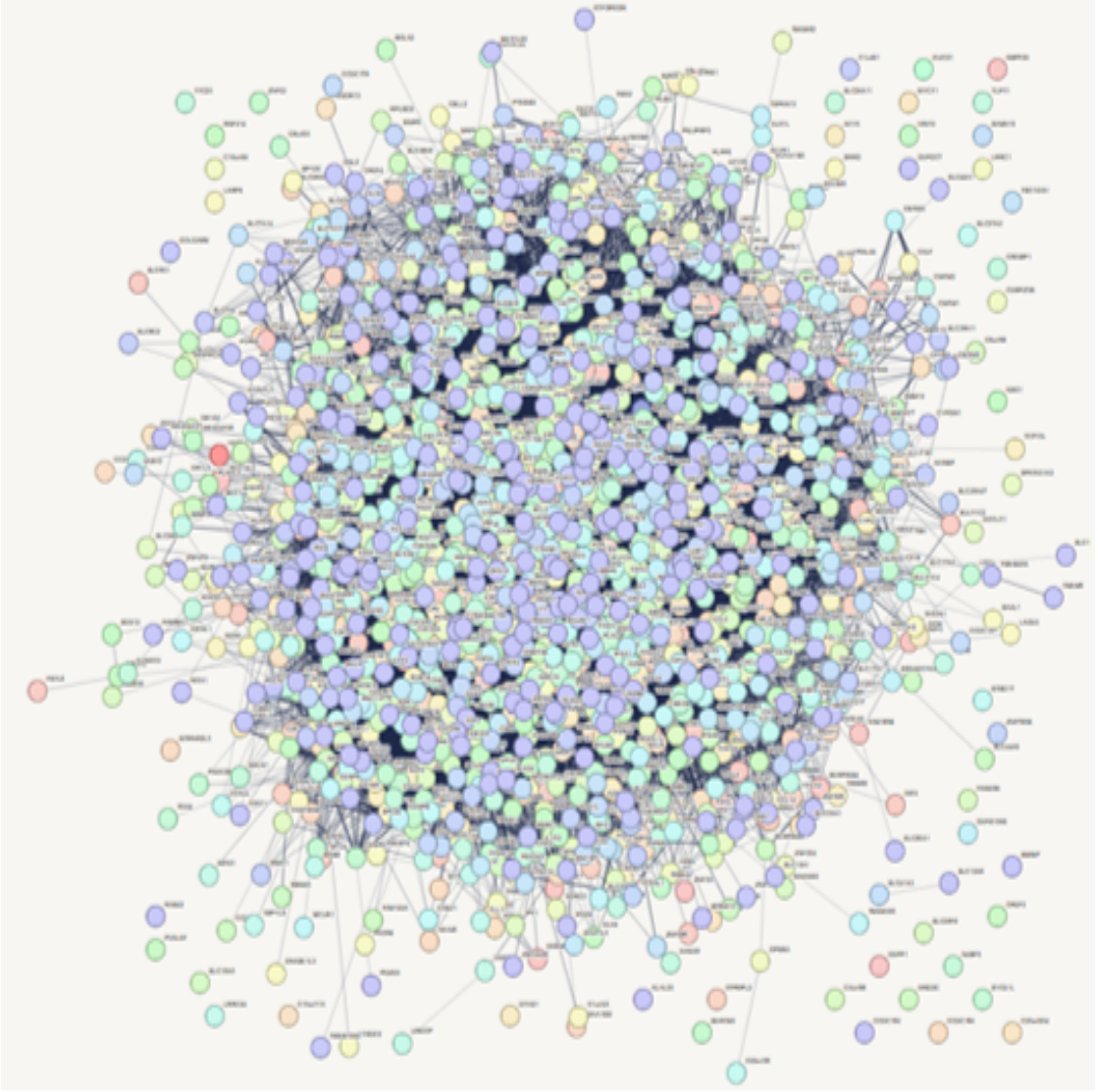


FIGURE 4.1: Network containing 2,000 genes.

4.3 Cluster Generation

The next step involved cluster generation. For this purpose, the 'Clusters' option

was selected, followed by the application of K-means clustering. The number of clusters was gradually adjusted, starting from three and increasing sequentially to four, five, six, and beyond, before applying the clustering algorithm (Fig 4.2).

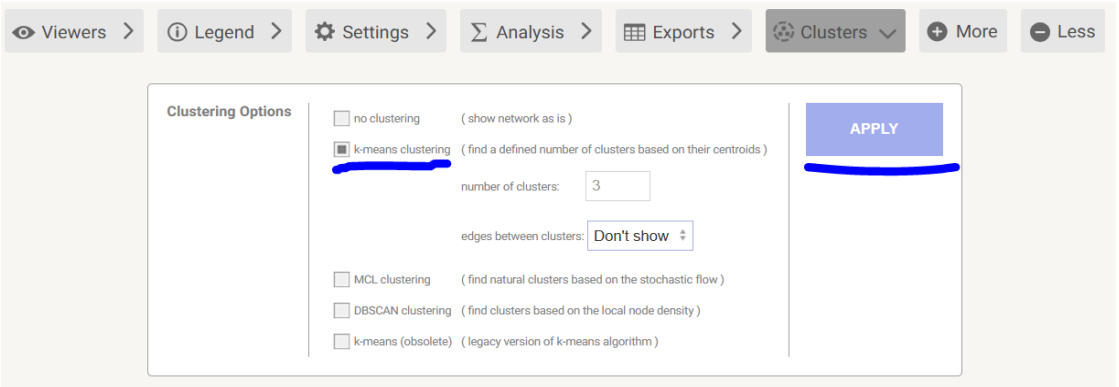


FIGURE 4.2: Executing k means clustering.

4.4 Verification by KEGG Database

- The KEGG Database contained 41 pathways of Liver Cancer. In all 41 pathways, clusters from STRING database ranging from 4 to 50 were observed. out of 169 genes, 37 matching genes were found in following pathways (Table 4.2):

TABLE 4.2: Observing matching genes in Liver Cancer pathways and Clusters.

Sr. No.	Pathway Name	Cluster No.	Genes in Cluster
1	Hepatocellular carcinoma		
2	Chemical carcinogenesis - DNA adducts		
3	Chemical carcinogenesis - receptor activation		
4	Other types of O-glycan biosynthesis		

Table 4.2: (Continued).

Sr. No.	Pathway Name	Cluster No.	Genes in Cluster
5	Metabolic pathways	Cluster no 4	PYCR2, ASS1, ARG1,OAT,OTC, ODC1, CPS1
		Cluster no 5	ALAS1, ATIC, ALDH1L1, BHMT, DMGDH, FTCD, GMPS, GNMT, MTR, MAT2B. MAT1A, MAT2A, PPAT, PDE10A
		Cluster no 7	B3GNT3, B4GALT1, GCNT2, FUT6, FUT8, MGAT3, MGAT5, MGAT4B, MGAT4A, ST6GAL1
		Cluster no 11	PNPLA3, PNLIPRP3
6	Endocrine resistance		
7	PPAR signaling pathway		
8	HIF-1 signaling pathway	Cluster no 6	TF TFRC
9	Protein processing in endoplasmic reticulum		
10	mTOR signaling pathway		
11	AMPK signaling pathway		
12	Wnt signaling pathway		
13	Notch signaling pathway		
14	Gap junction		

Table 4.2: (Continued).

Sr. No.	**Pathway Name**	**Cluster No.**	**Genes in Cluster**
15	Th1 and Th2 cell differentiation		
16	Thyroid hormone signaling pathway		
17	Regulation of lipolysis in adipocytes		
18	Relaxin signaling pathway		
19	Parathyroid hormone synthesis, secretion and action		
20	Insulin resistance		
21	Non-alcoholic fatty liver disease		
22	Growth hormone synthesis, secretion and action		
23	Alcoholic liver disease		
24	Alzheimer disease		
25	Spinocerebellar ataxia		
26	Prion disease		
27	Pathways of neurodegeneration - multiple diseases		
28	Shigellosis		
29	Malaria		
30	Amoebiasis		
31	Hepatitis C		

Table 4.2: (Continued).

Sr. No.	Pathway Name	Cluster No.	Genes in Cluster
32	Hepatitis B		
33	Human papillomavirus infection		
34	Epstein-Barr virus infection		
35	Pathways in cancer	Cluster no 6	RET
36	Viral carcinogenesis		
37	MicroRNAs in cancer	Cluster no 4	GLS
38	Breast cancer		
39	Gastric cancer		
40	Graft-versus-host disease		
41	Lipid and atherosclerosis		

- Pathway 5 naming (Metabolic Pathways) and cluster no 4 i.e. PYCR2, ASS1, ARG1, OAT, OTC, ODC1, CPS1 Cluster no. 5 i.e. ALAS1, ALDH1L1, BHMT, DMGDH, FTCD, GMPS, GNMT, MTR, MAT2B. MAT1A, MAT2A, PPAT, PDE10A, ATIC Cluster no.7 i.e. B3GNT3, B4GALT1, GCNT2, FUT6, FUT8, MGAT3, MGAT5, MGAT4B, MGAT4A, ST6GAL1 Cluster no. 11 i.e. PNPLA3, PNLIPRP3

- Pathway 8 naming (HIF 1 signaling pathway) and Cluster no 6 i.e. TF, TFRC

- Pathway 37 naming (MicroRNAs in cancer) and Cluster no 4 i.e. GLS

- Pathway 35 naming (Pathways in cancer) and Cluster no 6 i.e. RET

- Remaining 132 genes i.e. Were all predicted genes.

4.5 Functional Annotation by DAVID Tool

Functional annotation was carried out using the David tool. Clusters were the outcome of the functional annotation process. We were able to create twenty distinct clusters for the liver cancer genes based on the output of the David tool. The outcomes are displayed below:

4.5.1 Functional Annotation Clustering

4.5.1.1 Annotation Cluster 1

Functional annotation cluster 1 (Fig 4.3) shows 7.71 enrichment score. Many measures related to metal ion binding and homeostasis were significantly enriched in the study. Significant enrichment was seen for "metal ion homeostasis" (p-value: 1.48E-11) and "cellular response to toxic ions" (p-value: 7.41E-9). This suggests that dysregulation of metal ion transport and homeostasis mechanisms may play a critical role in liver cancer.

Annotation Cluster 1	Enrichment Score: 7.71	G		Count	P_Value	Benjamini
INTERPRO	Metalthion_vert_metal_BS	RT		7	6.4E-12	1.8E-9
GOTERM_BP_DIRECT	intracellular zinc ion homeostasis	RT		9	1.5E-11	8.7E-9
UP_KW_LIGAND	Cadmium	RT		7	2.1E-11	3.8E-10
UP_SEQ_FEATURE	REGION:Alpha	RT		7	2.7E-11	4.6E-9
UP_SEQ_FEATURE	REGION:Beta	RT		7	2.7E-11	4.6E-9
GOTERM_BP_DIRECT	cellular response to zinc ion	RT		8	4.4E-11	1.3E-8
INTERPRO	Metalthion_vert	RT		7	9.0E-11	6.1E-9
INTERPRO	Metalthion_dom_sf_vert	RT		7	9.0E-11	6.1E-9
INTERPRO	Metalthion_dom_sf	RT		7	9.0E-11	6.1E-9
UP_KW_LIGAND	Metal-thiolate cluster	RT		7	1.7E-10	1.5E-9
GOTERM_BP_DIRECT	detoxification of copper ion	RT		7	2.0E-10	3.9E-8
GOTERM_BP_DIRECT	negative regulation of growth	RT		7	4.6E-10	6.7E-8
KEGG_PATHWAY	Mineral absorption	RT		9	4.7E-10	1.6E-8
GOTERM_BP_DIRECT	cellular response to copper ion	RT		7	4.2E-9	5.0E-7
GOTERM_BP_DIRECT	cellular response to cadmium ion	RT		7	3.7E-8	3.1E-6
UP_KW_LIGAND	Copper	RT		7	1.2E-5	7.2E-5
GOTERM_MF_DIRECT	zinc ion binding	RT		9	1.6E-1	9.3E-1
UP_KW_LIGAND	Zinc	RT		17	8.1E-1	1.0E0
GOTERM_MF_DIRECT	metal ion binding	RT		13	9.2E-1	1.0E0
UP_KW_LIGAND	Metal-binding	RT		20	1.0E0	1.0E0

FIGURE 4.3: Functional annotation clustering of liver cancer genes.

4.5.1.2 Annotation Cluster 2

Functional annotation clustering study of liver cancer genes showed 7.54 enrichment in Annotation Cluster 2 (Fig 4.4). Many measures related to transmembrane regions and membrane components were significantly enriched in the study.

Significant enrichment was seen for "transmembrane helix" (p-value: 5.6E-13) and "membrane" (p-value: 1.2E-6). This suggests that changes in transmembrane structures and membrane-associated functions may contribute to liver cancer progression.

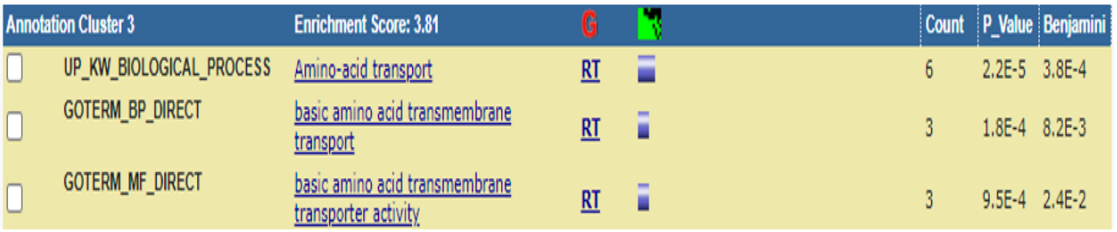
Annotation Cluster 2	Enrichment Score: 7.54	G		Count	P_Value	Benjamini
UP_KW_DOMAIN	Transmembrane helix	RT		72	5.6E-13	5.5E-12
UP_KW_DOMAIN	Transmembrane	RT		72	1.0E-12	5.5E-12
UP_SEQ_FEATURE	TRANSMEM:Helical	RT		68	1.9E-12	9.5E-10
GOTERM_CC_DIRECT	membrane	RT		57	2.0E-8	2.7E-6
UP_KW_CELLULAR_COMPONENT	Membrane	RT		77	1.2E-7	2.5E-6
UP_SEQ_FEATURE	TOPO_DOM:Extracellular	RT		39	7.3E-7	6.2E-5
UP_SEQ_FEATURE	CARBOHYD:N-linked (GlcNAc...) asparagine	RT		49	1.4E-6	1.0E-4
UP_SEQ_FEATURE	TOPO_DOM:Cytoplasmic	RT		45	1.8E-6	1.2E-4
GOTERM_CC_DIRECT	plasma membrane	RT		55	5.1E-6	3.5E-4
UP_KW_PTM	Glycoprotein	RT		50	1.9E-5	2.5E-4
UP_KW_CELLULAR_COMPONENT	Cell membrane	RT		44	2.4E-5	2.5E-4

FIGURE 4.4: Functional annotation clustering of liver cancer genes.

4.5.1.3 Annotation Cluster 3

Annotation Cluster 3 has an enrichment value of 3.81, indicating a significant increase in amino acid transport and transmembrane activity-related functional annotations in liver cancer (Fig 4.5). Important annotations include amino acid transport, basic amino acid transmembrane transport, and basic amino acid transport across membranes.

The words "amino acid transport" and "transmembrane activity" strongly suggest the involvement of cellular transport mechanisms. Most annotations have low P-values, indicating statistical significance; however, basic amino acid transmembrane transport has a slightly higher P-value (9.54E-4), indicating lower importance compared to other phrases in the cluster.



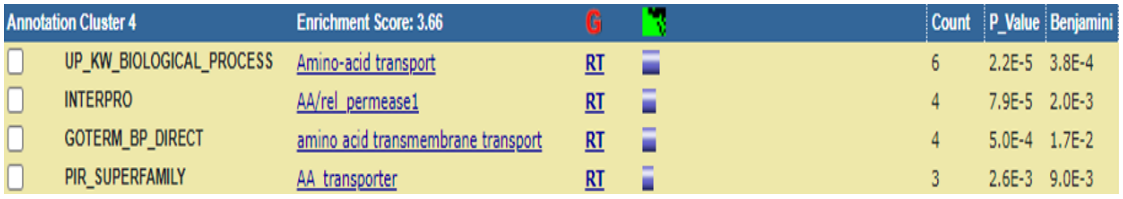
Annotation Cluster 3	Enrichment Score: 3.81			Count	P_Value	Benjamini
UP_KW_BIOLOGICAL_PROCESS	Amino-acid transport	RT		6	2.2E-5	3.8E-4
GOTERM_BP_DIRECT	basic amino acid transmembrane transport	RT		3	1.8E-4	8.2E-3
GOTERM_MF_DIRECT	basic amino acid transmembrane transporter activity	RT		3	9.5E-4	2.4E-2

FIGURE 4.5: Functional annotation clustering of liver cancer genes.

4.5.1.4 Annotation Cluster 4

Annotation Cluster 4 has an enrichment value of 3.66, indicating a significant increase in functional annotations related to amino acid transport and transmembrane transport in liver cancer (Fig 4.6). Important annotations include amino acid transport, amino acid transmembrane transport, and AA transport.

The words "amino acid transport" and "transmembrane transport" strongly suggest mechanisms of nutrient and ion movement across cellular membranes. Most annotations have low P-values, indicating high statistical significance; however, AA transport has a slightly lower P-value (2.64E-3), indicating it may be of slightly lower importance than other terms in the cluster.



Annotation Cluster 4	Enrichment Score: 3.66			Count	P_Value	Benjamini
UP_KW_BIOLOGICAL_PROCESS	Amino-acid transport	RT		6	2.2E-5	3.8E-4
INTERPRO	AA/rel permease1	RT		4	7.9E-5	2.0E-3
GOTERM_BP_DIRECT	amino acid transmembrane transport	RT		4	5.0E-4	1.7E-2
PIR_SUPERFAMILY	AA transporter	RT		3	2.6E-3	9.0E-3

FIGURE 4.6: Functional annotation clustering of liver cancer genes.

4.5.1.5 Annotation Cluster 5

Annotation Cluster 5 has an enrichment value of 3.1, indicating a significant increase in functional annotations related to transmembrane transporter activity and ABC transporters in liver cancer (Fig 4.7). The terms "transmembrane transport" and "ABC transporters" strongly suggest the involvement of transport mechanisms across membranes, particularly related to xenobiotics and metabolic resistance.

Most annotations have low P-values, indicating strong statistical significance, although terms such as lipid transport have slightly higher P-values (3.05E-1), indicating relatively lower importance in the cluster.

Annotation Cluster 5	Enrichment Score: 3.1	G		Count	P_Value	Benjami
GOTERM_MF_DIRECT	xenobiotic transmembrane transporter activity	RT		6	3.2E-7	3.6E-5
GOTERM_BP_DIRECT	transmembrane transport	RT		11	1.2E-6	7.7E-5
GOTERM_BP_DIRECT	xenobiotic transport across blood-brain barrier	RT		4	3.3E-6	1.9E-4
GOTERM_MF_DIRECT	organic anion transmembrane transporter activity	RT		5	9.4E-6	5.3E-4
GOTERM_CC_DIRECT	apical plasma membrane	RT		12	2.0E-5	8.8E-4
KEGG_PATHWAY	ABC transporters	RT		5	6.9E-5	1.2E-3
GOTERM_MF_DIRECT	ABC-type xenobiotic transporter activity	RT		4	7.0E-5	2.4E-3
GOTERM_MF_DIRECT	efflux transmembrane transporter activity	RT		4	7.0E-5	2.4E-3
GOTERM_MF_DIRECT	ABC-type transporter activity	RT		5	7.3E-5	2.4E-3
GOTERM_BP_DIRECT	transport across blood-brain barrier	RT		6	1.1E-4	5.7E-3
INTERPRO	ABC_transporter-like_ATP-bd	RT		5	2.1E-4	4.3E-3
GOTERM_BP_DIRECT	xenobiotic transport	RT		4	4.0E-4	1.5E-2
KEGG_PATHWAY	Antifolate resistance	RT		4	4.3E-4	3.6E-3
INTERPRO	ABC1_TM_dom	RT		4	6.0E-4	1.0E-2
INTERPRO	ABC1_TM_sf	RT		4	6.0E-4	1.0E-2
GOTERM_CC_DIRECT	basolateral plasma membrane	RT		8	8.1E-4	2.7E-2
GOTERM_BP_DIRECT	monoatomic anion transmembrane transport	RT		4	8.1E-4	2.6E-2
UP_KW_MOLECULAR_FUNCTION	Translocase	RT		5	9.7E-4	1.6E-2
GOTERM_BP_DIRECT	transepithelial transport	RT		3	1.3E-3	3.6E-2
GOTERM_BP_DIRECT	export across plasma membrane	RT		3	1.3E-3	3.6E-2
GOTERM_MF_DIRECT	ATPase-coupled transmembrane transporter activity	RT		4	1.5E-3	3.1E-2
SMART	AAA	RT		5	1.5E-3	1.0E-2
GOTERM_MF_DIRECT	ATPase-coupled inorganic anion transmembrane transporter activity	RT		3	1.5E-3	3.1E-2
INTERPRO	ABC_transporter-like_CS	RT		4	2.3E-3	2.8E-2
UP_SEQ_FEATURE	DOMAIN:ABC transporter	RT		4	2.7E-3	7.0E-2
UP_SEQ_FEATURE	DOMAIN:ABC transmembrane type-1 1	RT		3	3.7E-3	8.1E-2
UP_SEQ_FEATURE	DOMAIN:ABC transmembrane type-1 2	RT		3	3.7E-3	8.1E-2
GOTERM_BP_DIRECT	xenobiotic metabolic process	RT		5	4.7E-3	1.1E-1
INTERPRO	AAA+_ATPase	RT		5	8.6E-3	9.4E-2
UP_SEQ_FEATURE	DOMAIN:ABC transmembrane type-1	RT		3	1.1E-2	2.3E-1
UP_SEQ_FEATURE	DOMAIN:ABC transporter 1	RT		3	1.4E-2	2.7E-1
UP_SEQ_FEATURE	DOMAIN:ABC transporter 2	RT		3	1.4E-2	2.7E-1
KEGG_PATHWAY	Bile secretion	RT		3	7.4E-2	4.2E-1
INTERPRO	P-loop_NTPase	RT		9	1.5E-1	9.6E-1
GOTERM_MF_DIRECT	ATP hydrolysis activity	RT		5	2.5E-1	1.0E0
UP_KW_BIOLOGICAL_PROCESS	Lipid transport	RT		3	3.0E-1	1.0E0

FIGURE 4.7: Functional annotation clustering of liver cancer genes.

4.5.1.6 Annotation Cluster 6

Annotation Cluster 6 has an enrichment value of 3.08, indicating a significant increase in functional annotations related to extracellular functions and processes

in liver cancer (Fig 4.8). Important annotations include extracellular space, secreted proteins, and signal transduction.

The words "extracellular space" and "signal transduction" strongly suggest involvement in intercellular communication and extracellular processes. Most annotations have low P-values,indicating high statisticle significance, but terms like signal transduction have slightly higher P-values (7.0E-1), indicating relatively lower importance compared to other phrases in the cluster.

Annotation Cluster 6	Enrichment Score: 3.08	G		Count	P_Value	Benjamini
UP_SEQ_FEATURE	DOMAIN:MARVEL	RT		7	1.3E-8	1.3E-6
INTERPRO	Marvel	RT		7	1.6E-8	5.0E-7
UP_KW_MOLECULAR_FUNCTION	Cytokine	RT		6	1.3E-3	1.6E-2
UP_KW_BIOLOGICAL_PROCESS	Chemotaxis	RT		5	3.7E-3	3.2E-2
GOTERM_MF_DIRECT	cytokine activity	RT		5	2.9E-2	2.8E-1
GOTERM_BP_DIRECT	chemotaxis	RT		4	3.0E-2	4.5E-1
GOTERM_CC_DIRECT	extracellular space	RT		14	3.5E-1	1.0E0
GOTERM_BP_DIRECT	signal transduction	RT		7	7.0E-1	1.0E0

FIGURE 4.8: Functional annotation clustering of liver cancer genes.

4.5.1.7 Annotation Cluster 7

Annotation Cluster 7 has an enrichment value of 3.02, indicating a significant increase in functional annotations associated with transmembrane receptor activity and signaling pathways in liver cancer (Fig 4.9). Important annotations include transmembrane receptor activity, extracellular ligand-gated ion channel activity, and signal transduction pathways.

The terms "transmembrane receptor" and "extracellular signaling" strongly suggest the involvement of receptor-mediated signaling and intercellular communication mechanisms, which are vital for transmitting external signals into the cell to regulate various biological responses (such as cell growth, differentiation, immune responses, and cancer progression). Most annotations have low P-values, indicating high statistical significance, but terms such as plasma membrane receptor activity have slightly higher P-values (e.g., 2.06E-1), indicating lower importance in the cluster compared to other terms.

Annotation Cluster 7	Enrichment Score: 3.02	G		Count	P_Value	Benjamini
GOTERM_MF_DIRECT	transmembrane-ephrin receptor activity	RT		8	3.8E-12	8.7E-10
SMART	EPH_lbd	RT		6	8.4E-10	3.0E-8
UP_SEQ_FEATURE	DOMAIN:Eph LBD	RT		6	1.0E-8	1.3E-6
INTERPRO	Ephrin_rcpt_lig-bd_dom	RT		6	1.1E-8	3.7E-7
INTERPRO	Tyr_kinase_rcpt_V_CS	RT		6	1.1E-8	3.7E-7
INTERPRO	Tyr_kinase_ephrin_rcpt	RT		6	1.1E-8	3.7E-7
INTERPRO	Eph_TM	RT		6	1.1E-8	3.7E-7
GOTERM_BP_DIRECT	ephrin receptor signaling pathway	RT		8	1.3E-8	1.3E-6
PIR_SUPERFAMILY	TyrPK_ephrin_receptor	RT		6	1.7E-8	1.2E-7
SMART	Ephrin_rec_like	RT		6	2.2E-8	3.8E-7
GOTERM_MF_DIRECT	ephrin receptor activity	RT		4	6.9E-6	5.2E-4
SMART	SAM	RT		6	1.2E-5	1.2E-4
SMART	FN3	RT		7	1.3E-5	1.2E-4
INTERPRO	Tyr-kin_ephrin_A/B_rcpt-like	RT		4	6.4E-5	1.7E-3
INTERPRO	Galactose-bd-like_sf	RT		6	1.4E-4	3.1E-3
UP_SEQ_FEATURE	DOMAIN:Fibronectin type-III 1	RT		7	1.4E-4	5.3E-3
UP_SEQ_FEATURE	DOMAIN:Fibronectin type-III 2	RT		7	1.4E-4	5.3E-3
GOTERM_BP_DIRECT	axon guidance	RT		8	1.7E-4	8.1E-3
UP_SEQ_FEATURE	DOMAIN:SAM	RT		6	2.1E-4	6.7E-3
INTERPRO	SAM	RT		6	2.2E-4	4.3E-3
KEGG_PATHWAY	Axon guidance	RT		7	2.7E-4	3.1E-3
INTERPRO	SAM/pointed_sf	RT		6	8.0E-4	1.3E-2
INTERPRO	FN3_dom	RT		7	1.3E-3	1.9E-2
INTERPRO	FN3_sf	RT		7	1.4E-3	1.9E-2
INTERPRO	Growth_fac_rcpt_cys_sf	RT		6	1.4E-3	1.9E-2
UP_KW_MOLECULAR_FUNCTION	Tyrosine-protein kinase	RT		5	1.4E-3	1.6E-2
INTERPRO	Ser-Thr/Tyr_kinase_cat_dom	RT		6	1.8E-3	2.3E-2
UP_SEQ_FEATURE	MOTIF:PDZ-binding	RT		5	2.3E-3	6.2E-2
UP_SEQ_FEATURE	DOMAIN:Fibronectin type-III	RT		6	3.4E-3	8.1E-2
SMART	TyrKc	RT		4	4.1E-3	2.4E-2
GOTERM_CC_DIRECT	dendrite	RT		9	4.4E-3	1.2E-1
GOTERM_BP_DIRECT	multicellular organism development	RT		4	1.0E-2	1.9E-1
INTERPRO	Tyr_kinase_cat_dom	RT		4	1.5E-2	1.6E-1
INTERPRO	Tyr_kinase_AS	RT		4	1.9E-2	1.8E-1
GOTERM_MF_DIRECT	transmembrane receptor protein tyrosine kinase activity	RT		3	2.2E-2	2.5E-1
GOTERM_BP_DIRECT	positive regulation of kinase activity	RT		3	3.8E-2	5.0E-1
INTERPRO	Prot_kinase_dom	RT		7	6.7E-2	5.0E-1
UP_SEQ_FEATURE	DOMAIN:Protein kinase	RT		7	6.8E-2	9.4E-1
INTERPRO	Kinase-like_dom_sf	RT		7	9.3E-2	6.6E-1
UP_KW_MOLECULAR_FUNCTION	Kinase	RT		7	9.5E-2	5.8E-1
UP_KW_MOLECULAR_FUNCTION	Receptor	RT		12	1.5E-1	6.3E-1
INTERPRO	Protein_kinase_ATP_BS	RT		5	1.9E-1	1.0E0
GOTERM_MF_DIRECT	ATP binding	RT		13	2.0E-1	1.0E0
GOTERM_BP_DIRECT	protein autophosphorylation	RT		3	2.0E-1	1.0E0
UP_KW_BIOLOGICAL_PROCESS	Angiogenesis	RT		3	2.2E-1	1.0E0
UP_SEQ_FEATURE	ACT_SITE:Proton acceptor	RT		8	2.3E-1	1.0E0
UP_KW_MOLECULAR_FUNCTION	Transferase	RT		12	2.5E-1	9.3E-1
KEGG_PATHWAY	Rap1 signaling pathway	RT		3	2.8E-1	9.3E-1
INTERPRO	Ig-like_fold	RT		9	3.1E-1	1.0E0
KEGG_PATHWAY	Ras signaling pathway	RT		3	3.3E-1	9.3E-1
UP_KW_LIGAND	ATP-binding	RT		13	3.4E-1	1.0E0
GOTERM_BP_DIRECT	protein phosphorylation	RT		4	3.5E-1	1.0E0
GOTERM_CC_DIRECT	receptor complex	RT		3	3.6E-1	1.0E0
KEGG_PATHWAY	MAPK signaling pathway	RT		3	4.4E-1	9.3E-1
KEGG_PATHWAY	PI3K-Akt signaling pathway	RT		3	5.4E-1	1.0E0
GOTERM_BP_DIRECT	cell adhesion	RT		4	5.5E-1	1.0E0
GOTERM_BP_DIRECT	phosphorylation	RT		4	6.9E-1	1.0E0
UP_KW_LIGAND	Nucleotide-binding	RT		13	7.4E-1	1.0E0
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		3	8.2E-1	1.0E0

FIGURE 4.9: Functional annotation clustering of liver cancer genes.

4.5.1.8 Annotation Cluster 8

Annotation Cluster 8 has an enrichment score of 2.97, indicating moderate functional significance, with terms related to iron homeostasis, immune signaling (NOD1 / 2 pathway), and iron ion response (Fig 4.10). The P-values range from 0.0048 (highly significant) to 0.3845 (less significant), with gene counts between 3 and 5 for each term. Dysregulation of iron homeostasis, oxidative stress, and immune signaling are strongly implicated in liver cancer development, particularly due to their roles in chronic inflammation and DNA damage.

Low P-value annotations like intracellular iron ion homeostasis are more likely linked to cancer, while higher P-value terms might have weaker associations. Overall, these processes have a significant potential to contribute to liver cancer.

Annotation Cluster 8	Enrichment Score: 2.97	G		Count	P_Value	Benjamini
GOTERM_BP_DIRECT	intracellular iron ion homeostasis	RT		8	9.1E-8	6.6E-6
GOTERM_BP_DIRECT	multicellular organismal-level iron ion homeostasis	RT		4	4.0E-4	1.5E-2
GOTERM_BP_DIRECT	cellular response to iron ion	RT		3	8.4E-4	2.6E-2
GOTERM_BP_DIRECT	response to iron ion	RT		3	3.1E-3	7.5E-2
KEGG_PATHWAY	TGF-beta signaling pathway	RT		4	1.6E-2	1.1E-1
GOTERM_BP_DIRECT	positive regulation of transcription by RNA polymerase II	RT		5	9.3E-1	1.0E0

FIGURE 4.10: Functional annotation clustering of liver cancer genes.

4.5.1.9 Annotation Cluster 9

The DAVID annotation cluster 9, with an enrichment score of 2.91, highlights processes related to ion transport and sodium transport, which may have implications for cellular homeostasis and cancer development (Fig 4.11). Key enriched terms include ion transport (p-value: 3.94E-5) and sodium transport (p-value: 1.34E-4), both critical for maintaining ionic balance and cell signaling.

Other notable pathways include cation exchanger activity (p-value: 1.72E-3), which plays a role in calcium homeostasis, a key factor in cellular proliferation and apoptosis, processes often disrupted in liver cancer. The cluster also empha-

Annotation Cluster 9	Enrichment Score: 2.91	G		Count	P_Value	Benjamini
UP_KW_BIOLOGICAL_PROCESS	Antiport	RT		10	3.8E-9	9.9E-8
UP_KW_BIOLOGICAL_PROCESS	Ion transport	RT		15	3.9E-5	5.1E-4
UP_KW_BIOLOGICAL_PROCESS	Sodium transport	RT		7	1.3E-4	1.3E-3
UP_KW_LIGAND	Sodium	RT		7	5.2E-4	2.3E-3
INTERPRO	Cation/H_exchanger_CPA1	RT		3	1.7E-3	2.3E-2
GOTERM_MF_DIRECT	potassium:proton antiporter activity	RT		3	3.0E-3	5.5E-2
INTERPRO	Cation/H_exchanger	RT		3	3.3E-3	3.7E-2
GOTERM_MF_DIRECT	sodium:proton antiporter activity	RT		3	3.5E-3	5.6E-2
UP_KW_BIOLOGICAL_PROCESS	Symport	RT		5	6.3E-3	4.7E-2
GOTERM_BP_DIRECT	potassium ion transmembrane transport	RT		5	7.7E-3	1.6E-1
GOTERM_BP_DIRECT	proton transmembrane transport	RT		5	7.7E-3	1.6E-1
GOTERM_BP_DIRECT	sodium ion import across plasma membrane	RT		3	8.5E-3	1.7E-1
GOTERM_BP_DIRECT	regulation of intracellular pH	RT		3	1.1E-2	1.9E-1
GOTERM_BP_DIRECT	sodium ion transmembrane transport	RT		4	3.2E-2	4.7E-1
GOTERM_BP_DIRECT	monoatomic ion transport	RT		4	3.4E-2	4.7E-1

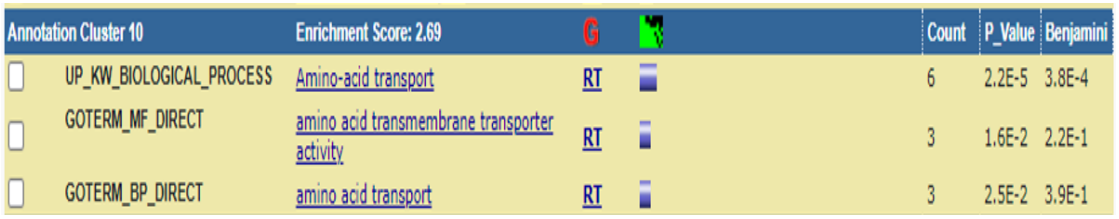
FIGURE 4.11: Functional annotation clustering of liver cancer genes.

sizes extracellular to intracellular ion transport (p-value: 3.9E-5) and sodium: proton antiporter activity (p-value: 3.8E-9), both of which are linked to pH regulation and cellular stress responses. These findings suggest that dysregulation of ion transport systems, particularly sodium and calcium transport, may contribute to liver cancer progression through disrupted signaling and homeostatic imbalances.

4.5.1.10 Annotation Cluster 10

The DAVID annotation cluster 10, with an enrichment score of 2.69, focuses on amino acid transport processes, which are critical for cellular metabolism and cancer progression (Fig 4.12). Key terms include amino acid transport (p-value: 2.28E-5) and amino acid transmembrane transporter activity (p-value: 1.42E-3), indicating the role of amino acid uptake in supporting cell growth and proliferation.

Dysregulated amino acid transport is a hallmark of cancer, as it fuels tumor cells' metabolic demands. Findings suggest that altered amino acid transport and transmembrane transporter activity may contribute to liver cancer by promoting nutrient uptake and metabolic reprogramming, essential for tumor growth and survival.

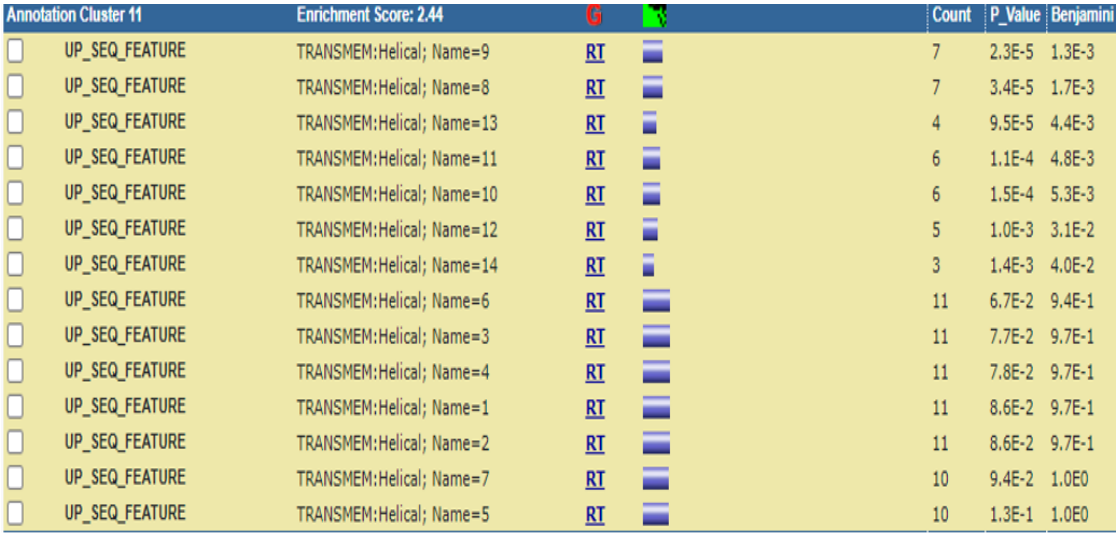


Annotation Cluster 10	Enrichment Score: 2.69			Count	P_Value	Benjamini
UP_KW_BIOLOGICAL_PROCESS	Amino-acid transport	RT		6	2.2E-5	3.8E-4
GOTERM_MF_DIRECT	amino acid transmembrane transporter activity	RT		3	1.6E-2	2.2E-1
GOTERM_BP_DIRECT	amino acid transport	RT		3	2.5E-2	3.9E-1

FIGURE 4.12: Functional annotation clustering of liver cancer genes.

4.5.1.11 Annotation Cluster 11

The DAVID annotation cluster 11, with an enrichment score of 2.44, highlights terms related to transmembrane helices, structural features of proteins involved in membrane transport and signaling (Fig 4.13). Key terms include TRANSMEM:Helix: Name-5 (p-value: 1.3E-1) and TRANSMEM:Helix: Name-3 (p-value: 7.7E-2), indicating a significant representation of transmembrane domains in the analyzed proteins.



Annotation Cluster 11	Enrichment Score: 2.44			Count	P_Value	Benjamini
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=9	RT		7	2.3E-5	1.3E-3
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=8	RT		7	3.4E-5	1.7E-3
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=13	RT		4	9.5E-5	4.4E-3
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=11	RT		6	1.1E-4	4.8E-3
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=10	RT		6	1.5E-4	5.3E-3
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=12	RT		5	1.0E-3	3.1E-2
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=14	RT		3	1.4E-3	4.0E-2
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=6	RT		11	6.7E-2	9.4E-1
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=3	RT		11	7.7E-2	9.7E-1
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=4	RT		11	7.8E-2	9.7E-1
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=1	RT		11	8.6E-2	9.7E-1
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=2	RT		11	8.6E-2	9.7E-1
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=7	RT		10	9.4E-2	1.0E0
UP_SEQ_FEATURE	TRANSMEM:Helical; Name=5	RT		10	1.3E-1	1.0E0

FIGURE 4.13: Functional annotation clustering of liver cancer genes.

These helices are typically associated with ion channels, receptors, and transporters, which are critical for cellular communication and nutrient uptake. Dysregulated transmembrane protein activity is often implicated in cancer, as it affects processes like signaling, growth, and metabolism. This cluster suggests that altered transmembrane helix-containing proteins may play a role in liver cancer progression by mediating abnormal transport and signaling pathways.

4.5.1.12 Annotation Cluster 12

The DAVID annotation cluster 12, with an enrichment score of 2.04, focuses on ion transmembrane transport, an essential process for cellular signaling and homeostasis (Fig 4.14). Key terms include ion transmembrane transport (p-value: 7.8E-3) and Zinc transport (p-value: 9.6E-3), highlighting the role of ion channels and transporters.

Dysregulated ion transport can disrupt cellular pH and energy metabolism, contributing to cancer progression by promoting tumor cell survival and proliferation. Abnormal proton gradients may foster an acidic tumor microenvironment, enabling invasion and metastasis. This underscores the potential role of ion transport dysregulation in cancer development.

Annotation Cluster 12	Enrichment Score: 2.04	G		Count	P_Value	Benjamini
GOTERM_BP_DIRECT	zinc ion transmembrane transport	RT		3	7.8E-3	1.6E-1
UP_KW_BIOLOGICAL_PROCESS	Zinc transport	RT		3	9.6E-3	6.2E-2
GOTERM_MF_DIRECT	zinc ion transmembrane transporter activity	RT		3	1.0E-2	1.5E-1

FIGURE 4.14: Functional annotation clustering of liver cancer genes.

4.5.1.13 Annotation Cluster 13

The DAVID annotation cluster 13, with an enrichment score of 1.35, focuses on transmembrane transport, which is essential for cellular function and signaling (Fig 4.15). The key terms include transmembrane transport (p-value: 3.1E-3), indicating the involvement of zinc ion transporters. Zinc is a cofactor for many enzymes and transcription factors, and its dysregulation is associated with cancer progression, including liver cancer.

Abnormal zinc transport can significantly impact critical cellular processes such as DNA repair, apoptosis, and the management of oxidative stress, all of which are essential for maintaining cellular integrity and function. Disruptions in these processes can lead to genomic instability and uncontrolled cell survival, potentially contributing to tumorigenesis. This suggests that disrupted zinc ion homeostasis

might play a meaningful role in the initiation and progression of liver cancer, making zinc-related pathways a potential area of interest for future research and therapeutic targeting.

Annotation Cluster 13	Enrichment Score: 1.35			Count	P_Value	Benjamini
GOTERM_MF_DIRECT	transmembrane transporter activity	RT		5	3.1E-3	5.5E-2
UP_SEQ_FEATURE	DOMAIN:Major facilitator superfamily (MFS) profile	RT		3	5.6E-2	8.7E-1
INTERPRO	MFS_dom	RT		3	1.1E-1	7.4E-1
INTERPRO	MFS_trans_sf	RT		3	2.1E-1	1.0E0

FIGURE 4.15: Functional annotation clustering of liver cancer genes.

4.5.1.14 Annotation Cluster 14

The DAVID annotation cluster 14, with an enrichment score of 4.18, emphasizes serine-type endopeptidase inhibitor action, which is essential for controlling protease activity and preserving proteostasis (Fig 4.16). Serine protease inhibitor (p-value: 5.42E-3), serine-type endopeptidase inhibitor activity (p-value: 1.3E-1), and protease inhibitor (p-value: 1.0E-1) are important phrases.

Excessive protease activity brought on by the dysregulation of these inhibitors can aid in the development of cancer by encouraging tissue invasion, metastasis, and immune evasion. This demonstrates the potential of serine protease inhibitors as therapeutic targets as well as their function in tumor suppression.

Annotation Cluster 14	Enrichment Score: 1.18			Count	P_Value	Benjamini
UP_SEQ_FEATURE	SITE:Reactive bond	RT		3	2.7E-2	4.8E-1
UP_KW_MOLECULAR_FUNCTION	Serine protease inhibitor	RT		3	5.4E-2	4.6E-1
UP_KW_MOLECULAR_FUNCTION	Protease inhibitor	RT		3	1.0E-1	5.8E-1
GOTERM_MF_DIRECT	serine-type endopeptidase inhibitor activity	RT		3	1.3E-1	8.2E-1

FIGURE 4.16: Functional annotation clustering of liver cancer genes.

4.5.1.15 Annotation Cluster 15

The DAVID annotation cluster 15 highlights the PDZ domain, a structural motif

essential for signaling cascades and protein-protein interactions, with an enrichment score of 0.56 (Fig 4.17). Important words include cell projection (p-value: 8.9E-1), domain-PDZ (p-value: 2.1E-1), and PDZ (p-value: 1.1E-1). Adhesion, intracellular communication, and cell polarity are all maintained by proteins with PDZ domains. Because dysregulated PDZ domain connections facilitate tumor invasion and promote aberrant cell signaling, they are linked to cancer. This highlights how crucial proteins containing PDZ are to the initiation and spread of cancer.

Annotation Cluster 15	Enrichment Score: 0.56	G		Count	P_Value	Benjamini
SMART	PDZ	RT		3	1.1E-1	4.3E-1
UP_SEQ_FEATURE	DOMAIN:PDZ	RT		3	2.1E-1	1.0E0
INTERPRO	PDZ	RT		3	2.3E-1	1.0E0
INTERPRO	PDZ_sf	RT		3	2.4E-1	1.0E0
UP_KW_CELLULAR_COMPONENT	Synapse	RT		5	4.0E-1	1.0E0
UP_KW_CELLULAR_COMPONENT	Cell projection	RT		6	8.9E-1	1.0E0

FIGURE 4.17: Functional annotation clustering of liver cancer genes.

4.5.1.16 Annotation Cluster 16

Annotation Cluster 16 exhibited 0.37 enrichment in the functional annotation clustering investigation of liver cancer genes. The study showed a high enrichment of many metrics pertaining to membrane components and mitochondrial areas (Fig 4.18).

Annotation Cluster 16	Enrichment Score: 0.37	G		Count	P_Value	Benjamini
UP_KW_CELLULAR_COMPONENT	Mitochondrion outer membrane	RT		3	2.5E-1	1.0E0
GOTERM_CC_DIRECT	mitochondrial outer membrane	RT		3	3.6E-1	1.0E0
UP_KW_CELLULAR_COMPONENT	Mitochondrion	RT		6	9.2E-1	1.0E0

FIGURE 4.18: Functional annotation clustering of liver cancer genes.

"Mitochondrion" (p-value: 9.2E-1), "mitochondrial outer membrane" (p-value: 2.5E-1), and "mitochondrion outer membrane" (p-value: 3.6E-1) all showed significant enrichment. This implies that modifications to mitochondrial architecture and membrane-related processes might accelerate the development of liver cancer.

4.5.1.17 Annotation Cluster 17

Annotation cluster 17 has an enrichment score of 0.34 and includes domains such as Ef hand one, Ef hand two, Ef hand dome pair, calcium ion binding, and calcium. The highest p-value is 8.6E-01 for calcium, followed by 6.5E-01 for calcium ion binding, and the lowest is 2.9E-01 for domain Ef hand one sequence feature. Annotation cluster 17, with its composition of Ef hand domains, calcium ion binding, and calcium, may potentially play a role in liver cancer. The presence of these elements suggests a possible involvement in cancer development or progression (Fig 4.19).

Annotation Cluster 17	Enrichment Score: 0.34	G		Count	P_Value	Benjamini
UP_SEQ_FEATURE	DOMAIN:EF-hand 1	RT		3	2.9E-1	1.0E0
UP_SEQ_FEATURE	DOMAIN:EF-hand 2	RT		3	3.0E-1	1.0E0
INTERPRO	EF_hand_dom	RT		3	3.9E-1	1.0E0
INTERPRO	EF-hand-dom_pair	RT		3	4.7E-1	1.0E0
GOTERM_MF_DIRECT	calcium ion binding	RT		5	6.5E-1	1.0E0
UP_KW_LIGAND	Calcium	RT		6	8.6E-1	1.0E0

FIGURE 4.19: Functional annotation clustering of liver cancer genes.

4.5.1.18 Annotation Cluster 18

Cluster 18 comprises internal membrane-bound organelles and protein transport, with an enrichment value of 0.24. 4.1E-01 to 7.7E-01 are the range of p-values for protein transport (Fig 4.20). Although it is not stated clearly what specific role this cluster plays in cancer, the presence of protein transport-related components raises the possibility of a meaningful connection to tumorigenesis, particularly in the context of liver cancer. Protein transport is essential for maintaining cellular organization and function, and its dysregulation could potentially contribute to altered signaling, immune evasion, or metabolic imbalance, causing cancer.

Annotation Cluster 18	Enrichment Score: 0.24	G		Count	P_Value	Benjamini
GOTERM_BP_DIRECT	protein transport	RT		4	4.1E-1	1.0E0
GOTERM_CC_DIRECT	intracellular membrane-bounded organelle	RT		6	5.9E-1	1.0E0
UP_KW_BIOLOGICAL_PROCESS	Protein transport	RT		4	7.7E-1	1.0E0

FIGURE 4.20: Functional annotation clustering of liver cancer genes.

4.5.1.19 Annotation Cluster 19

Differentiation and developmental proteins are the hallmarks of cluster number 19, which has an enrichment score of 0.09. P-values are 8.3E-01 for developmental proteins and 7.6E-01 for differentiation proteins (Fig 4.21). Although the precise effect of this cluster on liver cancer is not mentioned, the existence of developmental and differentiation proteins suggests a possible link.

Annotation Cluster 19	Enrichment Score: 0.09			Count	P_Value	Benjamini
UP_KW_BIOLOGICAL_PROCESS	Differentiation	RT		5	7.6E-1	1.0E0
UP_KW_MOLECULAR_FUNCTION	Developmental protein	RT		4	8.3E-1	1.0E0
UP_KW_MOLECULAR_FUNCTION	Developmental protein	RT		4	8.3E-1	1.0E0

FIGURE 4.21: Functional annotation clustering of liver cancer genes.

A comparatively lower likelihood is shown by the lower enrichment score of 0.09 when compared to clusters with higher values.

4.5.1.20 Annotation Cluster 20

With an enrichment value of 0.05, cluster number 20 includes components linked to transcriptional regulation, activators, DNA binding, and RNA polymerase II cis-regulatory region sequence-specific DNA binding (Fig 4.22). The p-values for activators and DNA binding vary from 6.0E-01 to 9.8E-01, whereas the p-values for transcriptional regulation and DNA binding unique to the RNA polymerase II cis-regulatory region sequence range from 1.0E+00.

Annotation Cluster 20	Enrichment Score: 0.05			Count	P_Value	Benjamini
UP_KW_MOLECULAR_FUNCTION	Activator	RT		4	6.0E-1	1.0E0
GOTERM_MF_DIRECT	DNA binding	RT		5	9.2E-1	1.0E0
UP_KW_MOLECULAR_FUNCTION	DNA-binding	RT		5	9.8E-1	1.0E0
GOTERM_MF_DIRECT	RNA polymerase II cis-regulatory region sequence-specific DNA binding	RT		3	1.0E0	1.0E0
UP_KW_BIOLOGICAL_PROCESS	Transcription regulation	RT		5	1.0E0	1.0E0
UP_KW_BIOLOGICAL_PROCESS	Transcription	RT		5	1.0E0	1.0E0

FIGURE 4.22: Functional annotation clustering of liver cancer genes.

Although the precise effect on cancer is not stated, the existence of these components raises the possibility of a link. A comparatively weaker relationship is shown by the lower enrichment score of 0.05 when compared to clusters with higher values.

4.6 Cluster Validation through Protein-Protein Interaction

We have validated some proteins from cluster no. 14 and cluster no. 25 through Protein-Protein interactions by Version 2.0 of the Protein Prompt server. The protein connections which were validated in these clusters are shown below:

Figure 4.23 contains a table with protein-related data, obtained from the Protein Prompt Server. The table includes three columns: Protein, Score (π = Binding), and UniProt ID. The identified protein is CSTF2, with a binding score of 0.6747, indicating the likelihood or strength of its interaction.

Protein	Score (1 = binding)	Uniprot ID
CSTF2	0.6747	P33240

FIGURE 4.23: Validated protein from cluster no 14.

The corresponding UniProt ID (P33240) links to the UniProt database, which provides detailed biological information about this protein. This data likely represents a prediction of protein-protein or protein-ligand interactions, where CSTF2 (Cleavage Stimulation Factor 2) is being assessed for its binding potential.

Figure 4.24 contains a table with protein-related data, likely obtained from the Protein Prompt Server. The table consists of three columns: Protein, Score (π = Binding), and UniProt ID. The identified protein is HOOK2, with a binding score of 0.6040, indicating the likelihood or strength of its interaction. The corresponding UniProt ID (Q96ED9) links to the UniProt database.

Protein	Score (1 = binding)	Uniprot ID
HOOK2	0.6040	Q96ED9

FIGURE 4.24: Validated protein from cluster no 25.

This provides detailed biological information about the HOOK2 protein, offering insights into its potential roles within the cellular environment. The data likely represents a prediction of protein-protein or protein-ligand interactions, aiming to assess HOOK2's binding potential and functional partnerships. Such predictive analysis can help infer the biological pathways HOOK2 may be involved in and its possible implications in liver cancer.

Chapter 5

Discussion

According to the available medical literature [1], Liver cancer is a severe and aggressive disease caused by the uncontrolled development of aberrant liver cells leading to high mortality rates worldwide. It is more common in older adults but can also affect younger individuals, depending on aspects linked to risk such as infections (e.g., hepatitis B and C), lifestyle, and genetic predisposition [1]. Access to cancer care remains a challenge in many regions, early detection is crucial for improving survival rates, with advancements in machine learning playing a key role in enhancing diagnostic accuracy. Despite medical progress, liver cancer continues to be a major health concern, requiring further research and innovation in both detection and treatment strategies [2].

Advanced biomedical text mining and bioinformatics techniques were employed to identify the mutations responsible for liver cancer, which was the primary focus of our study. To identify genes specifically associated with liver cancer, a comprehensive investigation utilizing the MeSH database was conducted. Information was gathered over a ten-year period, from 2020 to 2024, to ensure the inclusion of the most current and relevant academic sources. The COREMINE Medical tool facilitated the retrieval of several biological entities from these texts, including diseases, drugs, processes, genes, proteins, and MeSH terms.

Every piece of information that was collected was carefully arranged and examined. Identification of gene clusters ranging from four to fifty inside these pathways

using the STRING database was carried out. This demonstrates the existence of intricate networks of interactions that might contribute to the development of liver cancer. The identification of 41 pathways linked to liver cancer was made possible in large part by the KEGG pathway database.

There were four pathways that were particularly significant: Pathway 5, which is connected with cancer, and Pathway 8, which is associated with HIF 1 signalling pathway. Pathway 35 and pathway 38 also connected with cancer. Based on these findings, it is highly probable that cluster no 4 (PYCR2, ASS1, ARG1, OAT, OTC, ODC1, CPS1, GLS), cluster no. 5 (ALAS1, ATIC, ALDH1L1, BHMT, DMGDH, FTCD, GMPS, GNMT, MTR, MAT2B. MAT1A, MAT2A, PPAT, PDE10A), cluster no.7 (B3GNT3, B4GALT1, GCNT2, FUT6, FUT8, MGAT3, MGAT5, MGAT4B, MGAT4A, ST6GAL1), cluster no. 11 (PNPLA3, PNLIPRP3) and Cluster no. 6 (TF, TFRC, RET) are responsible for the development of Liver Cancer.

Further study was conducted using the DAVID program to broaden the list of anticipated genes, with particular attention given to those unique to humans. By restricting the experiment to a specific species, the findings were made directly relevant and applicable to liver cancer in humans. In the final phase of the study, specific proteins from Clusters 14 and 25 were identified by examining their interactions with other proteins. The Protein Prompt server, developed by the University of Leipzig (version 2.0), was utilized for this analysis. The validation method enhanced the credibility of the predictions by confirming the functional importance and interaction networks of the discovered proteins.

Finding and verifying unique genetic variants that could be involved in the development of Liver Cancer was the aim. This was accomplished by successfully combining a variety of bioinformatics methods and resources. Our knowledge of the molecular causes of Liver Cancer has grown as a result of these study findings, which might also direct future research and treatment strategies.

Chapter 6

Conclusion and Future Prospects

6.1 Conclusion

Liver cancer remains a major global health concern, necessitating advanced research for improved early detection and therapeutic intervention. This study employed a bioinformatics-driven approach to predict and analyze genes potentially associated with liver cancer. The methodology was grounded in the use of various computational tools, including text mining of scientific literature, functional annotation of gene lists, and pathway enrichment analysis to identify and characterize genetic factors involved in liver cancer progression.

Text mining, as an initial step in the research, was utilized to extract relevant gene information from an extensive corpus of biomedical literature. Through natural language processing algorithms, gene names, associated diseases, and functional terms were identified and compiled into a curated dataset. This method allowed the systematic retrieval of liver cancer-associated genes that have been previously mentioned or implicated in peer-reviewed studies. By parsing through thousands of abstracts and full-text articles, the approach ensured a broad and inclusive capture of candidate genes relevant to liver carcinogenesis.

Following gene identification, the study employed functional annotation to provide biological context to the curated gene set. Functional annotation involves mapping

genes to their known or predicted roles within biological systems, including cellular functions, molecular activities, and involvement in biological processes. This step was achieved using public databases such as Gene Ontology (GO), KEGG (Kyoto Encyclopedia of Genes and Genomes), and UniProt. The annotation process allowed the classification of genes based on their involvement in cancer-related pathways such as cell cycle regulation, apoptosis, immune responses, angiogenesis, and signal transduction.

A critical part of the analysis involved pathway enrichment, a statistical approach to determine whether a predefined set of genes is overrepresented in particular biological pathways. Enrichment analysis helped to highlight molecular pathways disproportionately associated with the genes identified through text mining. Tools like DAVID were employed for this purpose, revealing several key pathways that are frequently altered in liver cancer, including metabolic pathways, inflammatory signaling, and DNA repair mechanisms. This analysis supported the notion that liver cancer is a multifactorial disease involving numerous interrelated genetic and molecular events.

The integration of data from multiple bioinformatics platforms and tools added robustness to the findings. By cross-validating results obtained from different analytical methods, the study minimized biases and increased the confidence in the predicted gene-pathway associations. Furthermore, the study extended its analysis to the construction of protein-protein interaction (PPI) networks using databases such as STRING. This network visualized the physical and functional associations among the predicted genes at the protein level. The analysis uncovered several lesser-known or novel gene candidates that have not been extensively studied in the context of liver cancer. These findings highlight the potential of bioinformatics to reveal underexplored molecular targets.

It is important to note the limitations of a purely computational approach. While bioinformatics offers high-throughput and cost-effective analysis, experimental validation in cell lines, animal models, or clinical samples remains essential to confirm the functional role of the identified genes. Additionally, the dynamic and context-specific nature of gene expression implies that results may vary depending on tis-

sue type, disease stage, and individual patient variability. Nonetheless, the insights gained from bioinformatics analyses serve as a valuable starting point for experimental investigations. In conclusion, this study demonstrates the efficacy of a bioinformatics-driven methodology in elucidating key molecular players and pathways involved in liver cancer.

6.2 Future Recommendations

Future research should focus on validating these computational predictions through laboratory-based experiments and clinical studies to translate these findings into practical applications. By bridging computational biology with experimental research, this study contributes to the growing field of precision medicine, offering insights that may aid in the development of targeted diagnostic and therapeutic strategies for liver cancer.

6.2.1 Integration with Multi-Omics Data

To gain a more holistic understanding of liver cancer biology, such integration can demonstrate regulatory mechanisms, identify additional biomarkers, and uncover new therapeutic goals by providing a system-level view of tumor behavior. This approach can also help overcome support for one-dimensional data records and improve the accuracy and depth of cancer research.

6.2.2 Verification through Experimental and Clinical Research

Computational instruments provide valuable predictions, but experimental verification is important to confirm the biological significance of identified genes and pathways. Future studies should include in vitro and in vivo studies to assess gene expression, protein activity, and interactions in liver cancer cell lines and animal models. Additionally, large clinical studies are required to examine these results

in different patient populations and assess diagnosis, prognosis, or treatment relevance.

6.2.3 Development of Predictive Diagnostic Tools

The biomarkers identified by bioinformatics analysis are committed to developing non-invasive or minimal invasive diagnostic tools. Future research should focus on implementing these molecular markers in clinically available platforms such as gene expression collectors, blood-based assays, and AI-regulated diagnostic models. Such tools can significantly improve early detection of liver cancer, which can lead to improved survival and timely therapeutic interventions.

6.2.4 Targeted Drug Discovery and Reuse

With functional annotations of specific genes and pathways associated with liver cancer, future efforts can aim to develop or reuse drugs aimed at these molecular entities. Computer drug screening, virtual docking, and creation of pharmacological genomic profiles can accelerate the identification of promising candidates. Integrating these results into clinical research can lead to more effective and personalized liver cancer therapy, with less impact on Octat.

6.2.5 Personalized Medical Approach

In the context of liver cancer, which remains one of the leading causes of cancer-related mortality worldwide, the implementation of personalized medicine is particularly crucial. Hepatocellular carcinoma (HCC), the most common type of liver cancer, is known for its aggressive progression and resistance to standard therapies in many patients. The heterogeneity of HCC at the genetic, transcriptomic, and epigenetic levels contributes significantly to this therapeutic resistance. Therefore, future research must focus on identifying reliable biomarkers that reflect the individual molecular landscape of each tumor. This includes studying how epigenetic

modifications, such as histone acetylation, methylation, and non-coding RNAs, contribute to tumor development and progression.

By harnessing high-throughput technologies like next-generation sequencing and single-cell transcriptomics, it becomes possible to dissect the molecular intricacies of liver cancer with unparalleled precision. Such comprehensive molecular profiling can be used to stratify patients into subgroups that are more likely to respond to specific drugs or immunotherapies. Ultimately, this approach transforms liver cancer treatment from generalized protocols to dynamic, precision-guided interventions that evolve based on real-time molecular data. The path toward fully personalized liver cancer care requires interdisciplinary collaboration among bioinformaticians, oncologists, molecular biologists, and data scientists to translate research findings into actionable clinical strategies. As the field progresses, it is expected that personalized medicine will become the cornerstone of cancer therapy, offering hope for improved survival and quality of life in patients battling liver cancer.

Bibliography

[1] F. Sultana, O. A. Almassri, *et al.*, "Cancer opinion survey 2019: Cancer care in saudi arabia," *Saudi Journal of Oral and Dental Research*, vol. 5, no. 3, pp. 156–163, 2020.

[2] T. Saba, "Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges," *Journal of Infection and Public Health*, vol. 13, no. 9, pp. 1274–1289, 2020.

[3] S. Tu, W. Huang, C. Huang, Z. Luo, and X. Yan, "Contextual regulation of tgf- signaling in liver cancer," *Cells*, vol. 8, no. 10, p. 1235, 2019.

[4] S. Patel, P. Sadhu, M. Kumari, D. K. Dash, S. Jain, and A. K. Sen, "A review on anticancer agents: Conventional drugs and novel target specific inhibitors," *Oriental Journal of Chemistry*, vol. 39, no. 3, p. 657, 2023.

[5] P. Srivatanakul *et al.*, "Epidemiology of liver cancer: an overview," *Asian Pacific Journal of Cancer Prevention*, vol. 5, no. 2, pp. 118–125, 2004.

[6] L. Lin, L. Yan, Y. Liu, C. Qu, J. Ni, and H. Li, "The burden and trends of primary liver cancer caused by specific etiologies from 1990 to 2017 at the global, regional, national, age, and sex level results from the global burden of disease study 2017," *Liver Cancer*, vol. 9, no. 5, pp. 563–582, 2020.

[7] F. X. Bosch, J. Ribes, and J. Borràs, "Epidemiology of primary liver cancer," *Seminars in Liver Disease*, vol. 19, no. 3, pp. 271–285, 1999.

[8] S. Butt *et al.*, "Hepatitis c virus in pakistan: A systematic review of prevalence, genotype distribution and risk factors," *Liver International*, vol. 40, pp. 2535–2546, Oct. 2020.

[9] J. Calderaro, M. Ziol, V. Paradis, and J. Zucman-Rossi, "Molecular and histological correlations in liver cancer," *Journal of Hepatology*, vol. 71, no. 3, pp. 616–630, 2019.

[10] D. Anwanwan, S. K. Singh, S. Singh, V. Saikam, and R. Singh, "Challenges in liver cancer and possible treatment approaches," *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1873, no. 1, p. 188314, 2020.

[11] C. Y. Liu, K. F. Chen, and P. J. Chen, "Treatment of liver cancer," *Cold Spring Harbor Perspectives in Medicine*, vol. 5, no. 9, p. a021535, 2015.

[12] G. R. Oliver, S. N. Hart, and E. W. Klee, "Bioinformatics for clinical next generation sequencing," *Clinical Chemistry*, vol. 61, no. 1, pp. 124–135, 2015.

[13] S. ElShal, J. Davis, and Y. Moreau, "Biomedical text mining for disease-gene discovery," *EMBnet Journal 18*, vol. 18, no. B, pp. 108–109, 2013.

[14] T. Schmidt and D. Frishman, "Prompt: a protein mapping and comparison tool," *BMC Bioinformatics*, vol. 7, no. 331, pp. 1–15, 2006.

[15] D. W. Huang *et al.*, "David bioinformatics resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists," *Nucleic Acids Research*, vol. 35, no. 2, pp. W169–W175, 2007.

[16] S. Faro, T. Lecroq, S. Borzi, S. D. Mauro, and A. Maggio, "The string matching algorithms research tool," in *Stringology*, pp. 99–111, 2016.

[17] H. B. El-Serag and K. L. Rudolph, "Hepatocellular carcinoma: Epidemiology and molecular carcinogenesis," *Gastroenterology*, vol. 132, no. 7, pp. 2557–2576, 2007.

[18] S. Butt *et al.*, "Hepatitis c virus in pakistan: A systematic review of prevalence, genotype distribution and risk factors," *Liver International*, vol. 40, pp. 2535–2546, Oct. 2020.

[19] S. H. Hamid *et al.*, "Aflatoxin exposure and hepatocellular carcinoma in pakistan," *Journal of Hepatology*, vol. 58, no. 2, pp. 428–429, 2013.

[20] A. S. Butt, Z. Abbas, and W. Jafri, "Hepatocellular carcinoma in pakistan: Where do we stand?," *Hepatitis Monthly*, vol. 12, no. 10, p. e6023, 2012.

[21] A. Banerjee and P. Farci, "Fibrosis and hepatocarcinogenesis: Role of gene-environment interactions in liver disease progression," *International Journal of Molecular Sciences*, vol. 25, no. 16, p. 8641, 2024.

[22] S. Butt *et al.*, "Hepatocellular carcinoma in native south asian pakistani population; trends, clinico-pathological characteristics differences in viral marker negative viral-hepatocellular carcinoma," *BMC Research Notes*, vol. 6, pp. 1–13, 2013.

[23] G. Khemlina, S. Ikeda, and R. Kurzrock, "Hepatocellular carcinoma: A narrative review on current knowledge and future prospects," *Current Treatment Options in Oncology*, vol. 24, no. 2, pp. 173–186, 2023.

[24] S. M. Afzal, "Hepatocellular carcinoma in pakistan: An insight into future," *Journal of Cancer Science Therapy*, vol. 2, no. 12, pp. 746–747, 2016.

[25] A. Muhammad *et al.*, "Gender specific, worldwide picture of liver cancer: A special comparison with pakistan," *Pesquisa Agropecuaria Brasileira*, vol. 1, no. 9, pp. 2403–2414, 2020.

[26] Y. Xue, Y. Ruan, Y. Wang, P. Xiao, and J. Xu, "Signaling pathways in liver cancer: pathogenesis and targeted therapy," *Molecular Biomedicine*, vol. 5, no. 1, p. 20, 2024.

[27] P. Ginès, A. Krag, J. G. Abraldes, E. Solà, N. Fabrellas, and P. S. Kamath, "Liver cirrhosis," *The Lancet*, vol. 398, no. 10308, pp. 1359–1376, 2021.

[28] D. Schuppan and N. H. Afdhal, "Liver cirrhosis," *The Lancet*, vol. 371, no. 9615, pp. 838–851, 2008.

[29] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," *CA: A Cancer Journal for Clinicians*, vol. 65, no. 2, pp. 87–108, 2015.

[30] F. X. Bosch, J. Ribes, M. Díaz, and R. Cléries, "Primary liver cancer: Worldwide incidence and trends," *Gastroenterology*, vol. 127, no. 5, p. S5–S16, 2004.

[31] B. Sadikovic, K. Al-Romiah, J. A. Squire, and M. Zielenska, "Cause and consequences of genetic and epigenetic alterations in human cancer," *Current Genomics*, vol. 9, no. 6, pp. 394–408, 2008.

[32] C. L. Wilson, D. A. Mann, and L. A. Borthwick, "Epigenetic reprogramming in liver fibrosis and cancer," *Advanced Drug Delivery Reviews*, vol. 121, pp. 124–132, 2017.

[33] M. G. Fernández-Barrena, M. Arechederra, L. Colyn, C. Berasain, and M. A. Avila, "Epigenetics in hepatocellular carcinoma development and therapy: The tip of the iceberg," *JHEP Reports*, vol. 2, no. 6, p. 100167, 2020.

[34] T. H. Bestor, "The dna methyltransferases of mammals," *Human Molecular Genetics*, vol. 9, pp. 2395–2402, 2000.

[35] M. Okano, S. Xie, and E. Li, "Cloning and characterization of a family of novel mammalian dna (cytosine-5) methyltransferases," *Nature Genetics*, vol. 19, no. 3, pp. 219–220, 1998.

[36] J. Song, M. Teplova, S. Ishibe-Murakami, and D. J. Patel, "Structure-based mechanistic insights into dnmt1-mediated maintenance dna methylation," *Science*, vol. 335, no. 6069, p. 709–712, 2012.

[37] V. K. Rakyan, T. A. Down, D. J. Balding, and S. Beck, "Epigenome-wide association studies for common human diseases," *Nature Reviews Genetics*, vol. 12, no. 8, pp. 529–541, 2011.

[38] J.-P. Issa, "Cpg island methylator phenotype in cancer," *Nature Reviews Cancer*, vol. 4, no. 12, pp. 988–993, 2004.

[39] A. Eden, F. Gaudet, A. Waghmare, and R. Jaenisch, "Chromosomal instability and tumors promoted by dna hypomethylation," *Science*, vol. 300, no. 5168, p. 455, 2003.

[40] S. Villanueva, "Hepatocellular carcinoma," *The New England Journal of Medicine*, vol. 380, no. 15, pp. 1450–1462, 2019.

[41] M. E. Galle, C. Forner, and P. T. Llovet, "Molecular classification of hepatocellular carcinoma," *Journal of Hepatology*, vol. 75, no. 3, pp. 856–868, 2022.

[42] H. Sun, Z. Zhou, and S. Wang, "The role of vegf in hepatocellular carcinoma angiogenesis," *Cancer Biology Therapy*, vol. 19, no. 7, pp. 573–582, 2018.

[43] P. M. Sherman and M. Jonas, "Chronic liver inflammation and the progression to hepatocellular carcinoma," *Nature Reviews Gastroenterology Hepatology*, vol. 16, no. 5, pp. 256–270, 2019.

[44] I. A. for Research on Cancer, "Liver cancer: A review of etiological factors," *IARC Monographs*, vol. 101, pp. 37–44, 2021.

[45] L. Rimassa, A. E. Abou-Alfa, and T. Lamarca, "Immune evasion mechanisms in hepatocellular carcinoma," *The Lancet Oncology*, vol. 20, no. 6, pp. e263–e272, 2020.

[46] P. Konyn, A. Ahmed, and D. Kim, "The current trends in the health burden of primary liver cancer across the globe," *Clinical and Molecular Hepatology*, vol. 29, no. 2, pp. 358–362, 2023.

[47] Z. Liu *et al.*, "Genomic landscape and mutational analysis of hepatocellular carcinoma," *Nature Genetics*, vol. 50, no. 2, pp. 134–144, 2018.

[48] H. Schulze *et al.*, "Wnt/-catenin signaling in hepatocellular carcinoma," *Journal of Hepatology*, vol. 70, no. 1, pp. 176–185, 2019.

[49] A. Totoki *et al.*, "Axin1 mutations in hepatocellular carcinoma and their role in the wnt pathway," *Cancer Research*, vol. 78, no. 4, pp. 1128–1140, 2017.

[50] L. Nault and J. Zucman-Rossi, "Tert promoter mutations in hepatocellular carcinoma," *Hepatology*, vol. 67, no. 6, pp. 2323–2330, 2018.

[51] M. Kanwal *et al.*, "Epigenetic changes in liver cancer: Mechanisms and clinical implications," *Nature Reviews Gastroenterology Hepatology*, vol. 15, no. 9, pp. 536–553, 2020.

[52] S. Villanueva, "Hepatocellular carcinoma: Novel molecular approaches," *The New England Journal of Medicine*, vol. 380, no. 15, pp. 1450–1462, 2019.

[53] C. Llovet *et al.*, "Genomics and precision medicine in hepatocellular carcinoma," *Nature Reviews Clinical Oncology*, vol. 18, no. 5, pp. 293–312, 2021.

[54] E. J. Park *et al.*, "Role of kras mutations in liver cancer," *Journal of Clinical Oncology*, vol. 36, no. 8, pp. 719–728, 2018.

[55] F. Hoshida *et al.*, "Gene expression in molecular subtypes of liver cancer," *Cancer Cell*, vol. 29, no. 1, pp. 102–115, 2016.

[56] X. Zhao, Z. Yang, and Y. Bao, "Application of bioinformatics in liver cancer research," *Biomedicine Pharmacotherapy*, vol. 121, p. 109634, 2020.

[57] M. Yu *et al.*, "Bioinformatics-based screening of key genes for the transformation of liver cirrhosis to hepatocellular carcinoma," *Journal of Translational Medicine*, vol. 21, no. 1, p. 387, 2023.

[58] C. Yang *et al.*, "Identifying prognostic gene expression signatures in liver cancer using bioinformatics methods," *Frontiers in Oncology*, vol. 8, p. 618, 2018.

[59] A. Raj, R. C. Petreaca, and G. Mirzaei, "Multi-omics integration for liver cancer using regression analysis," *Current Issues in Molecular Biology*, vol. 46, no. 4, pp. 3551–3562, 2024.

[60] D. Szklarczyk *et al.*, "String v10: Protein-protein interaction networks, integrated over the tree of life," *Nucleic Acids Research*, vol. 45, no. D1, pp. D362–D368, 2017.

[61] J. J. Harding, S. Nandakumar, and J. Armenia, "Prospective genomic profiling of hepatocellular carcinoma: Clinical implications," *JCO Precision Oncology*, vol. 3, no. 7, pp. 1–15, 2019.

[62] M. Ashburner *et al.*, "Gene ontology: Tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[63] L. Rokach and O. Maimon, "Clustering methods," in *Data Mining and Knowledge Discovery Handbook*, p. 321–352, 2005.

[64] S. ElShal, J. Davis, and Y. Moreau, "Biomedical text mining for disease gene discovery," *EMBnet Journal*, vol. 18, pp. 108–109, Nov. 2012.

[65] S. V. Gaikwad, A. Chaugule, and P. Patil, "Text mining methods and techniques," *International Journal of Computer Applications*, vol. 85, no. 17, pp. 42–45, 2014.

[66] M. Hearst, "What is text mining," *SIMS, UC Berkeley*, vol. 5, 2003.

[67] C. C. Aggarwal and C. Zhai, "An introduction to text mining," in *Mining Text Data*, pp. 1–10, Springer US, 2012.

[68] K. G. . Jain and A. Kumar, "Text mining and network analysis for discovery of novel genes associated with staphylococcus biofilms," *MOJ Proteomics Bioinform*, vol. 4, no. 5, p. 00135, 2016.

[69] D. Szklarczyk *et al.*, "String v11: Protein-protein association networks in the age of genomics," *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–D613, 2019.

[70] H. V. Cook, N. T. Doncheva, D. Szklarczyk, C. V. Mering, and L. J. Jensen, "String: A virus-host protein-protein interaction database," *Viruses*, vol. 10, no. 10, p. 519, 2018.

[71] K. F. Aoki and M. Kanehisa, "Using the kegg database resource," *Current Protocols in Bioinformatics*, vol. 11, no. 1, pp. 1–12, 2005.

[72] A. N. Srivastava and M. Sahami, *Text Mining: Classification, Clustering, and Applications*. Boca Raton, FL, USA: CRC Press, 2009.

[73] D. W. Huang *et al.*, "David bioinformatics resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists," *Nucleic Acids Research*, vol. 35, no. 2, pp. W169–W175, 2007.

[74] G. Dennis, B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, and R. A. Lempicki, "David: Database for annotation, visualization, and integrated discovery," *Genome Biology*, vol. 4, pp. 1–11, 2003.

[75] V. S. Rao, K. Srinivas, G. N. Sujini, and G. S. Kumar, "Protein-protein interaction detection: Methods and analysis," *International Journal of Proteomics*, vol. 2014, no. 1, p. 147648, 2014.

Appendix

In the initial phase, genes were extracted using the COREMINE tool. This involved accessing the COREMINE search interface and entering the keyword "liver cancer" into the search bar. By selecting the "extracted associations" option and filtering results from the years 2020 to 2024, relevant data was retrieved. COREMINE generated a list of 2000 genes associated with liver cancer based on the query "Liver Cancer (Genes/Proteins)." The resulting data, including gene names and significance value can be noted in Table 1.

TABLE 1: Genes associated with liver cancer and their significance values retrieved from the Coremine Medical database. Lower significance values indicate stronger associations between the gene and liver cancer.

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
NFE2L2	0.000607	MCL1	0.000616	MTDH	0.000623
ZFPM2-AS1	0.000624	MRPS11	0.000624	INSR	0.00063
RASSF1	0.000636	FGFR4	0.000649	NCAPG	0.000651
DDX11-AS1	0.000654	LAPTM4B	0.000656	PKM	0.00066
TNS1	0.000663	FOXM1	0.000669	CCNB2	0.000678
BSG	0.00068	MAT1A	0.000684	NUP62	0.000688
FUNDC2	0.000692	KHDRBS1	0.000698	GABPA	0.000698
HK2	0.000702	SLC45A2	0.000705	CDC20	0.000718
CCNE1	0.000731	RGN	0.000736	LINC01134	0.00074
CDKN2A	0.000746	CDCA8	0.000756	YAP1	0.000766
YY1AP1	0.000767	TBC1D9	0.000768	SPP1	0.000771
GSTP1	0.000786	IFNA1	0.000791	SLC7A11	0.000799

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
ONECUT1	0.000803	KIF20A	0.000805	LINC00205	0.000809
CDKN1B	0.00081	SALL4	0.000811	KPNA2	0.000819
TCF4	0.000821	NHCP1	0.000871	HSD17B13	0.000874
DLGAP5	0.000876	LINC01093	0.000881	CYP3A4	0.000886
CENPK	0.0009	NQO1	0.000902	SMYD3	0.000906
LINC01224	0.000908	MAP1LC3A	0.000908	ZEB1	0.000917
MIR21	0.000921	MTTP	0.000923	FASN	0.000927
MGAT5	0.000929	MIA3	0.00093	TNFRSF10B	0.000932
HTATIP2	0.000936	CD44	0.00094	AKT1	0.000941
HAVCR2	0.000951	SNHG3	0.000951	METTL3	0.000952
SETD2	0.000954	CYP1A2	0.000956	CEBPB	0.000956
PRKAA2	0.000978	IL6	0.001	HCRP1	0.001001
DDIT3	0.001001	ARID1A	0.001007	PRKAB1	0.001007
KRAS	0.001007	NR1I3	0.00101	HFE	0.001019
CHD1L	0.001021	RBM39	0.001021	PRKAA1	0.001025
CEBPA	0.001027	MDK	0.001029	PCK2	0.001045
SOCS1	0.001054	MDM2	0.001061	PTK2B	0.001061
FOXA1	0.001061	ASPM	0.001067	NR1I2	0.001067
TPX2	0.001075	NUSAP1	0.001111	KRT7	0.001114
MYLK-AS1	0.001117	CYCS	0.001118	LINC01419	0.001121
HNF1B	0.001129	SNORD126	0.001137	SARNP	0.001145
HSPA5	0.001149	PRKACA	0.00115	SNAI1	0.001156
PRR34-AS1	0.001166	CYP7A1	0.001169	TM4SF4	0.001176
LINC01138	0.001181	HHCM	0.001182	FOXO1	0.001186
APOA1	0.001196	DNMT1	0.001198	SERPINB3	0.001202
HMGCR	0.001207	ABCB11	0.001221	ANGPT2	0.001224
TF	0.001231	UCK2	0.001233	NR1H4	0.001233
SULT1C4	0.001236	SREBF2	0.001236	NME1	0.00124
CTLA4	0.001248	BECN1	0.00125	CCL15	0.001251

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
SNHG1	0.001259	MT1G	0.001262	FOXA2	0.001264
MGAT3	0.001268	CIB1	0.001286	SMAD3	0.001291
KRT18	0.001293	ZWINT	0.001314	ARID2	0.001315
LRPPRC	0.001323	HAMP	0.001327	ADH4	0.001328
TMEM220-AS1	0.001329	CCL14	0.001331	THY1	0.001335
FASLG	0.001336	NFKB1	0.001356	MBOAT7	0.001358
SNHG6	0.001359	ABCC2	0.00136	APOB	0.001361
XIAP	0.001363	IL6ST	0.001363	MAT2B	0.001364
IQGAP2	0.001369	STC2	0.001371	STAT1	0.001375
ABCB4	0.001382	CD34	0.00139	WNK1	0.001392
RELA	0.001393	FGL1	0.001407	TGFB1	0.001407
HOTAIR	0.001409	NOTCH1	0.001412	HDAC1	0.001419
STAM	0.00142	RRM2	0.001426	SLC26A5	0.001429
YBX1	0.001433	SLC1A5	0.001434	NEK2	0.00144
DKK1	0.001444	LINC00501	0.001456	GCNT2	0.001457
GLUL	0.00149	ZC3H13	0.001494	NR0B2	0.001505
GPX4	0.001457	C1QTNF1-AS1	0.001481	CDK6	0.001486
TMEM106C	0.001521	MIR4435-2HG	0.001524	TSPAN1	0.001526
NANOGP8	0.001526	RAF1	0.001533	NANOG	0.001538
SYT1	0.001539	MALAT1	0.001542	KEAP1	0.001542
KLRK1	0.001543	ST8SIA6-AS1	0.001545	IRS1	0.001553
CASZ1	0.001554	YTHDF1	0.001557	LHPP	0.001561
HRAS	0.001564	MIR210HG	0.001571	SYCE1L	0.001574
GUCD1	0.001576	MIR1269B	0.001576	LOC101927789	0.001576
COL18A1	0.001589	DNMT3B	0.00159	MTUS2	0.001597

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
ZEB2	0.001602	KLHL1	0.001603	ATP7B	0.001604
YTHDF2	0.001627	ZHX2	0.001636	LGALS3BP	0.001639
MAPKAPK5-AS1	0.001667	JAK1	0.001686	CDC25C	0.001697
PLK1	0.001699	DUXAP8	0.001702	LDLR	0.001714
LINC00161	0.001715	HCCAT5	0.001715	LDHA	0.001717
ARHGEF39	0.001726	DNASE1L3	0.001732	SOCS3	0.001733
ATG5	0.001736	MT1M	0.001752	SLC2A1	0.001752
MAGEA1	0.001763	HOOK2	0.001769	OSM	0.001771
DNAJB1	0.001774	ABCC3	0.001779	PRR34	0.001781
HTR2A-AS1	0.001781	PTK2	0.001785	BMI1	0.001817
SLC5A5	0.00182	SMAD4	0.001822	ALDH2	0.001827
PEG10	0.001828	KIAA1429	0.00184	CRNDE	0.00184
TGFA	0.00185	ABCC1	0.001852	URGCP	0.001852
POLD1	0.001857	DYNLL1	0.001869	SMAD2	0.00187
ARG1	0.001882	MTUS1	0.001887	TIGIT	0.001893
C10orf91	0.001909	CDC25A	0.001912	PRC1	0.001913
MMP7	0.001929	PTBP1	0.001937	MAGEC2	0.001941
MSC-AS1	0.001946	RB1	0.00196	SOX4	0.001967
CDKN3	0.001971	MCM7	0.001972	UBD	0.001987
GNA14	0.001992	GALNT14	0.001992	ANXA2	0.001992
KLF6	0.001992	ODC1	0.001993	IGF2BP3	0.001995
LAG3	0.001998	CCNE2	0.002	SKP2	0.002007
FAS	0.002009	LINC-ROR	0.002014	SLCO1B3	0.002025
CBX2	0.002026	TXNRD1	0.002028	COMMD7	0.00204
BUB1B	0.002043	LETMD1	0.002046	MIR192	0.002072
TOP2A	0.002072	ZSCAN16-AS1	0.002083	FBP1	0.002084
FZD7	0.002091	SLC12A9	0.002095	SCAMP3	0.002098

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
DEPDC1	0.002101	PRRT3-AS1	0.002104	NCOA5	0.002105
GLI1	0.002106	TSPEAR-AS1	0.002108	FAM99B	0.002108
UBE2C	0.002111	SOCS2	0.002111	TCF7L2	0.002153
RBMS3	0.002163	PBLD	0.002166	LINC00942	0.002166
NDRG1	0.002167	WNK2	0.002172	SULF1	0.002172
NELFE	0.002173	CENPF	0.002189	CD24	0.002196
PFKFB4	0.002205	MRPL28	0.002205	GJB1	0.002213
FAH	0.002214	FOXO3	0.002215	RDH16	0.002224
VWCE	0.002224	AXL	0.002228	DNASE1	0.002235
USP22	0.002236	SULF2	0.002238	CPEB3	0.002248
ACSL4	0.002251	IQGAP3	0.00226	SPATS2	0.002272
POU5F1	0.002296	WAC-AS1	0.0023	PECAM1	0.0023
RIMS2	0.002308	EGFR	0.002321	SLC22A1	0.002326
RHOC	0.002326	SPATA2	0.002333	SUB1	0.002347
RHPN1-AS1	0.00235	EPHA2	0.002351	VPS37A	0.002353
SMAD7	0.002358	VPS53	0.00236	SLC9C1	0.002361
MAPK1	0.002364	E2F3P1	0.002364	LINC00869	0.002364
SMIM7	0.002364	INTS6P1	0.002364	AKR1B10P1	0.002364
LINC01136	0.002364	PLCE1-AS1	0.002364	WARS2-IT1	0.002364
ENPP7P13	0.002364	AL359878.1	0.002364	RP11-40C6.2	0.002364
AC020915.2	0.002364	ZBTB7A	0.002366	MACC1	0.002368
LPCAT1	0.002382	HEPN1	0.002386	HPSE	0.002389
CYP2B6	0.002395	GADD45A	0.002409	CDCA5	0.002409
LGALS3	0.002422	DLX6-AS1	0.002422	LIN28B	0.002434
PRDM2	0.002436	KRT88P	0.002443	RUNX3	0.002450
FOXA3	0.002456	MCM3	0.002457	ATAT1	0.002492
LINC01063	0.002496	UBE2T	0.002509	LECT2	0.002513
FTCD	0.002516	NDC80	0.002523	INS	0.002527

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
HPX	0.002543	IGFBP1	0.002544	DLGAP1-AS1	0.002547
PPM1G	0.002553	FUT8	0.002559	PTTG1	0.002569
SOX12	0.002586	NDRG2	0.002591	DERL2	0.002602
SIRT1	0.002611	GFER	0.002641	ATG7	0.002649
HOTTIP	0.002652	TP53I11	0.002677	WSPAR	0.002685
PPAPDC1B	0.002691	ZNF689	0.002691	ISG20L2	0.002693
NRAS	0.002694	PAK3	0.002711	SNHG16	0.002712
GLYAT	0.002729	FHIT	0.002729	MGAT4A	0.002736
NSG1	0.00274	RACGAP1	0.002747	COPS5	0.00275
NET1	0.00276	MIR22HG	0.00276	CHEK1	0.002761
MIR145	0.002761	CD5L	0.002767	CNBP	0.00277
NEAT1	0.002779	BBS9	0.002783	MDM4	0.002787
LMOD1	0.002788	LINC00665	0.002792	OGDHL	0.002792
MIRLET7C	0.002809	CLIC4	0.002809	ATF4	0.002829
ELAVL1	0.00283	HCCS	0.002851	HINT2	0.002856
PSMB4	0.002873	HDAC9	0.00288	MS	0.002882
PRKCA	0.002884	CFLAR	0.002899	MAD2L1	0.002914
TP53BP2	0.00292	HNRNPC	0.002926	YY1	0.002932
ASS1	0.002941	CEP55	0.00298	FABP1	0.002986
SRPRB	0.002986	BRI3BP	0.002989	METTL1	0.00299
MAPK8	0.002996	LINC00462	0.002997	LINC01287	0.002997
BID	0.003	PTPRE	0.003001	RNANC	0.003013
SPC25	0.003015	BCL9	0.003015	IGF2BP1	0.003043
JAK2	0.003059	ACOX1	0.003063	AURKB	0.003078
EPS8L3	0.00309	FLVCR1-AS1	0.00309	HHLA2	0.003091
SAC3D1	0.003103	ANXA7	0.003117	HMGCS2	0.003125
GZMB	0.003126	MSH3	0.00313	SNHG20	0.003133
ZFAS1	0.003141	WDR4	0.003145	FCN3	0.003151

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
CAP2	0.003423	SCARB1	0.003456	KIFC1	0.003558
LUCAT1	0.003425	PLAU	0.003457	HP	0.00356
ECT2	0.003443	PTPN6	0.003462	NUF2	0.003562
CFHR3	0.00345	SERPINE1	0.003463	LINC01225	0.003562
SPHK1	0.003454	MAP1LC3B	0.00348	GTF2IRD2B	0.003562
ANXA10	0.003511	BUB1	0.003516	LINC00238	0.003562
KIF11	0.003522	KIAA1522	0.003524	MIG7	0.003562
DUXAP10	0.003524	GSTA2	0.003529	FAM182B	0.003562
BBC3	0.003531	CCNG1	0.003531	LINC01549	0.003562
SLC6A1-AS1	0.003562	SEMA6A-AS1	0.003562	LINC01517	0.003562
TAS2R13	0.003569	CMTM6	0.003601	CYP2A	0.003602
BMND7	0.003603	BMND8	0.003603	SLC44A5	0.003626
CDCA3	0.003638	MCM6	0.003642	ADH7	0.003653
SLC7A1	0.003657	DNLZ	0.003666	TNF	0.00367
ACYP1	0.00369	ETS1	0.003691	SNHG4	0.003697
LATS1	0.003697	TXN	0.003716	DIO3OS	0.003719
APP	0.003721	PCNAP1	0.003724	TMEM220	0.003724
MT1F	0.003753	EP300	0.003753	PDCD4	0.003755
JAG1	0.003784	SAA1	0.003791	CDC123	0.003796
HDAC2	0.003804	YEATS2	0.003808	MIR146A	0.003812
SLC25A11	0.003829	SPTBN1	0.003831	SLC7A14	0.003854
HMOX1	0.003859	TYMS	0.003861	E2F8	0.003871
XPO5	0.003874	RXRA	0.003882	ASPH	0.003882
HMGA1	0.003892	ABCG2	0.003899	TCHP	0.003913
E2F3	0.003931	EIF5A2	0.003942	FZR1	0.003943
UGT1A7	0.003943	ARID4B	0.00395	E2F2	0.003952
APOA2	0.003955	PFDN5	0.003961	IGF2BP2	0.003979
PDCD1LG2	0.003987	PARK2	0.003996	EGR1	0.003999

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
EIF2AK2	0.004001	MCM2	0.004012	SLC38A11	0.004023
UBL4A	0.004034	TWIST1	0.004047	MYCN	0.004058
MTRNR2L3	0.004063	RFC4	0.004068	SLC7A5	0.004072
ENO1	0.004079	MAGEA6	0.004086	SLC16A11	0.004089
CES1	0.004094	CENPL	0.004094	GCLC	0.004095
ANLN	0.004108	TSPYL5	0.004132	TRIM35	0.004144
PLXNB2	0.004147	TFRC	0.00416	CLDN1	0.004175
PPARGC1A	0.004185	URI1	0.004208	EVA1A	0.004215
HCG18	0.004227	MARK2	0.004227	FUBP1	0.00423
MRPL9	0.004236	NOL10	0.004236	GLYATL1	0.004236
LINC00844	0.004236	PCBP2-OT1	0.004236	CYP1B1	0.004238
MTA1	0.004238	AKR7A3	0.004251	TRMT6	0.004251
NM	0.004251	CRAT	0.00426	TMCO3	0.004269
KLHL23	0.004269	LINC00176	0.004269	FAM99A	0.004269
EIF3J-AS1	0.004269	SNORA47	0.004269	PXN-AS1	0.004269
HMMR-AS1	0.004269	IGF1R	0.004279	ANXA5	0.004281
FOXK2	0.004293	CTAG1B	0.004295	TTC36	0.004297
OIT3	0.004297	TEAD2	0.004305	HOXA11-AS	0.004339
MAPK3	0.004345	ZNF148	0.004359	WNT3A	0.004371
SULT1A3	0.004381	SLC27A5	0.004405	ATF3	0.004407
HAND2-AS1	0.004409	TRIP13	0.004411	GCLM	0.004418
EIF4EBP1	0.00442	MST1	0.004421	IGF2R	0.004429
MEG3	0.00443	ST6GAL1	0.004436	NR5A2	0.004444
SRSF3	0.004464	APOBEC1	0.004464	CCL20	0.004466
DNMT3A	0.004485	MIR96	0.004488	S100A11	0.004492
SLC9C2	0.004494	MTFR2	0.004501	TMEM9	0.004501
STK4	0.004505	MTSS1	0.004506	FGF21	0.004507
RNF38	0.004524	CLEC4G	0.00454	SLC25A13	0.004541
SUGP1	0.004548	LINC00467	0.004549	FBXL6	0.00455

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
KIF4A	0.004557	PRDX5	0.004579	SLC25A25	0.00459
TCF19	0.004596	RAC1	0.004611	USF1	0.004615
CDC25B	0.004618	XRCC1	0.00462	GSK3B	0.00462
G6PD	0.004628	SBF2-AS1	0.004633	PCK1	0.004634
SLC26A11	0.004642	NR3C1	0.004655	SLC7A13	0.00466
ACP1	0.004661	XPO4	0.004669	EPX	0.004669
CMTM4	0.004681	FOXP3	0.004691	MTF1	0.004696
SLC7A7	0.004704	TMEM74	0.004706	MIR22	0.004714
VLDLR	0.004725	FRA5C	0.004727	DKFZP434L187	0.004727
OR4F4	0.004727	OR4F3	0.004727	SNORD52	0.004727
IGKV1D-33	0.004727	OR4F17	0.004727	C16orf46	0.004727
C22orf39	0.004727	C5orf58	0.004727	ZNF738	0.004727
LINC00479	0.004727	FAM43B	0.004727	SNRPFP1	0.004727
CDRT4	0.004727	RPL7AP6	0.004727	HIST2H2BC	0.004727
EHHADH-AS1	0.004727	LOC341056	0.004727	ZNF233	0.004727
C10orf113	0.004727	AATK-AS1	0.004727	LINC00272	0.004727
LOC390638	0.004727	SRRD	0.004727	HCG20	0.004727
WHAMMP2	0.004727	DNM1P41	0.004727	LRRC24	0.004727
BPY2C	0.004727	JST	0.004727	ADAMTSL4-AS1	0.004727
MIR520F	0.004727	MIR519C	0.004727	MIR518D	0.004727
UBE2MP1	0.004727	LOC643387	0.004727	PPIAP22	0.004727
PDPK2P	0.004727	TMEM78	0.004727	MIR550A1	0.004727
MIR645	0.004727	SPATA31A3	0.004727	LOC728290	0.004727
LINC01183	0.004727	GAGE12H	0.004727	TBC1D3H	0.004727
LINC01348	0.004727	SNORD113-1	0.004727	PDZK1P1	0.004727
MIR541	0.004727	RPL7P24	0.004727	FUNDC2P4	0.004727
LOC100128398	0.004727	LOC100133920	0.004727	LINC01056	0.004727

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
SH3RF3-AS1	0.004727	HSPA8P4	0.004727	LINC00674	0.004727
MIR3660	0.004727	MIR3653	0.004727	SNHG21	0.004727
ALKBH3-AS1	0.004727	LINC01604	0.004727	BLOC1S5-TXNDC5	0.004727
RNASEK-C17orf49	0.004727	MIR4800	0.004727	MIR4784	0.004727
MIR4644	0.004727	MIR5003	0.004727	ELMO1-AS1	0.004727
LINC00556	0.004727	MTND4P20	0.004727	RNA5SP38	0.004727
SNRK-AS1	0.004727	ZNF385D-AS2	0.004727	COL18A1-AS2	0.004727
ARHGEF7-AS2	0.004727	SLC25A30-AS1	0.004727	EHMT2-AS1	0.004727
EGLN3-AS1	0.004727	MIR6883	0.004727	MIR7706	0.004727
MIR6089	0.004727	LINC01374	0.004727	DLX2-AS1	0.004727
ZNF529-AS1	0.004727	LINC01607	0.004727	LINC01608	0.004727
LOC101926913	0.004727	LOC101927051	0.004727	LOC101928775	0.004727
CYP3A	0.004787	MAFG-AS1	0.00479	ATF6	0.004793
LOC101929241	0.004727	AP000769.1	0.004727	RP11-863K10.7	0.004727
RP5-1014O16.1	0.004727	RP11-640M9.2	0.004727	AC108463.2	0.004727
RP11-495P10.1	0.004727	AC007389.1	0.004727	RP11-241J12.3	0.004727
CTD-2313N18.8	0.004727	RP11-564D11.3	0.004727	RP11-424C20.2	0.004727
RP11-467L13.4	0.004727	RP11-187E13.1	0.004727	RP11-343N15.5	0.004727
CCNJL	0.004772	LINC00635	0.004772	RNF214	0.004772
ARHGEF37	0.004772	BAIAP2-AS1	0.004772	RNF216P1	0.004772

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
SNORD76	0.004772	SENP3-EIF4A1	0.004772	CD276	0.004784
RET	0.004805	LAGE3	0.004814	CTGF	0.004818
MRPL54	0.004818	SNAI3-AS1	0.004818	LINC00200	0.004818
LASP1	0.004829	HLA-DRB1	0.004834	ABCC5	0.004844
SOAT1	0.004844	RBM12	0.004864	METTL6	0.004864
TRERNA1	0.004864	CENPM	0.004887	SLCO3A1	0.004892
CBX4	0.004892	TUFT1	0.004895	MMP14	0.004901
GSTM1	0.004901	AURKA	0.00492	LIN28A	0.004921
ZEB1-AS1	0.004923	SERPINB8	0.004928	PVT1	0.004947
SLC1A1	0.004948	IFNL4	0.004955	PTTG3P	0.004959
ISX	0.004959	KIAA1524	0.004967	MVP	0.00497
MIR203A	0.004978	CASC2	0.004996	SLC45A1	0.005004
MPZL1	0.005008	CRYZ	0.005011	REG3A	0.005027
EPHA1	0.005036	MCM3AP-AS1	0.005038	TP53I3	0.005066
HES1	0.005071	SLC39A11	0.005079	NEIL3	0.005101
KIF18B	0.005106	NKD1	0.005113	ALDOA	0.005113
NR2F1-AS1	0.005117	OAT	0.005123	EPHB6	0.005136
EPHB4	0.005140	MBD2	0.005142	RHOA	0.005145
CKLF	0.005179	MT1B	0.005180	PNPLA7	0.005181
SLC7A11-AS1	0.005181	BAK1	0.005183	POU2F1	0.005191
SLC7A6	0.005196	SLC9A9	0.005205	MIR183	0.005212
DUSP1	0.005216	BAD	0.005217	LRRC1	0.005226
SLC2A2	0.005247	ADH1B	0.005247	STK25	0.005271
SAV1	0.005273	HJURP	0.005278	IL24	0.005291
ADAM9	0.005302	JUNB	0.005304	TNFRSF6B	0.005310
RNASEH2A	0.005311	ENPP2	0.005315	PAH	0.005317

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
NUPR1	0.005324	CDC37L1	0.005325	FSTL5	0.005325
PAQR4	0.005325	SSRP1	0.005326	WWOX	0.005329
SS2R	0.005336	SPRTN	0.005336	ROCK2	0.005352
MIR93	0.005354	H2AFX	0.005357	ILF3	0.005360
SLC7A10	0.005360	EIF3H	0.005366	C1orf61	0.005371
IGKV2D-24	0.005371	TTC23	0.005371	LIX1L	0.005371
A1BG-AS1	0.005371	CECR7	0.005371	EHMT2	0.005381
FBXW7	0.005390	CYP4F3	0.005391	ANPEP	0.005415
SNHG7	0.005420	AXIN2	0.005437	BACE1-AS	0.005447
CD68	0.005467	FABP5P1	0.005479	LINC00473	0.005485
CYP2C9	0.005492	NAA40	0.005499	RUSC1-AS1	0.005499
SULT1A1	0.0055	SFRP1	0.005533	SLC16A1-AS1	0.005533
FADD	0.005543	BTBD7	0.005546	PEMT	0.005548
EGFL7	0.005555	UBE2S	0.005557	ARHGAP1	0.005559
UBE2B	0.005559	PDGFRB	0.005566	KLF4	0.005581
HSP90B1	0.005585	TTR	0.005628	MAP2K1	0.005632
PDK1	0.005632	FN1	0.00565	APAF1	0.005653
SNRPD2	0.00566	FBXO43	0.00566	MYBL2	0.005664
SOX2	0.005672	PIN1	0.005682	CCT3	0.005687
LRP1B	0.005689	HSPB1	0.005692	LINC01296	0.005694
MAGI1	0.005703	CKS2	0.005707	SHH	0.005712
ACTB	0.005719	MKI67	0.005721	MT-TP	0.005734
IGDCC3	0.005741	NT5DC2	0.005741	IRF1	0.005755
CLDN7	0.005761	SNAI2	0.005764	PCLO	0.005787
PIK3R3	0.005813	CDCA2	0.005827	IL2	0.005829
SPC24	0.005832	TAB3	0.005836	CMTM5	0.005842
EHHADH	0.005854	SKA1	0.00587	PLG	0.005873
BHMT	0.005874	CXCR4	0.005889	GAGE1	0.00589

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
ROCK1	0.005906	HDAC3	0.005908	CCDC58	0.005927
GRPEL2	0.005927	WEE2-AS1	0.005927	MIR9-2	0.005927
HNF4A-AS1	0.005927	FOS	0.005927	DTYMK	0.005943
BAIAP2L2	0.005951	DKK3	0.005988	TMEM147	0.005991
PHC3	0.005991	FAM189B	0.005994	LINC00652	0.005994
C8orf33	0.005994	ZNF479	0.005994	SOX9-AS1	0.005994
POTEKP	0.005994	DNAJC25	0.005994	MIR3677	0.005994
LINC00441	0.005994	UPK1A-AS1	0.005994	TRPC7-AS1	0.005994
PLBD1-AS1	0.005994	KIF1B	0.005999	LOXL2	0.006000
ALDH1A1	0.006003	TEAD4	0.006037	DEGS1	0.006044
ARCN1	0.006047	COPD	0.006055	MICA	0.006058
AKT2	0.006076	BCLAF1	0.006081	GSDMC	0.006082
MIR185	0.006083	CHD5	0.006094	TROAP	0.006114
MT1X	0.006131	MIR224	0.006139	CASC11	0.006140
RASSF5	0.006145	HNRNPDL	0.006152	GAPDH	0.006158
TIMP3	0.006168	STMN1	0.006185	POLR2E	0.006195
SLC44A1	0.006196	ATIC	0.006198	MLXIPL	0.006201
PDZD2	0.006204	MIR26B	0.006256	SLC28A1	0.006262
SLCO1B1	0.006265	ACADL	0.006279	SLC38A1	0.006297
RNF2	0.006312	TIPARP	0.006334	DDX3X	0.006342
HAGLR	0.006366	SLC7A4	0.006372	ACY3	0.006372
PKLR	0.006373	STEAP3	0.006379	UCA1	0.006380
EPHA3	0.006382	FOXQ1	0.006386	SERPINA7	0.006396
CDKN1C	0.006400	MIR25	0.006410	METTL14	0.006410
MELK	0.006415	NRF1	0.006438	MIR490	0.006440
DGCR5	0.006452	NDUFA4L2	0.006452	PREX2	0.006455
TUG1	0.006460	GBAP1	0.006462	ASB16-AS1	0.006462
HNRNPA2B1	0.006473	POGK	0.006486	MRPL21	0.006486
ZNF674-AS1	0.006486	MIR326	0.006486	GAS5	0.006488

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
FNDC3B	0.006490	LINC00511	0.006490	CPS1	0.006490
LEF1	0.006493	GPNMB	0.006495	CTHRC1	0.006497
ALAS1	0.006518	AKT3	0.006525	LINC01116	0.006535
KLF8	0.006549	GSTM2	0.006559	UGT1A1	0.006567
GOT2	0.006571	GSTA1	0.006576	RPS6KB1	0.006580
BAP1	0.006584	BZW2	0.006584	PAFAH1B3	0.006591
NCAM1	0.006596	SAA2	0.006601	NUAK1	0.006604
HMGCS1	0.006616	TIMP2	0.006622	AHSG	0.006640
S100A9	0.006665	YES1	0.006666	SULT1B1	0.006668
SLC27A6	0.006668	PDPK1	0.006682	FLT4	0.006704
WWTR1	0.006706	MT1H	0.006713	TLL1	0.006713
OCM	0.006718	LILRB1	0.006727	GMPS	0.006727
HDLBP	0.006747	APOC3	0.006762	ULBP1	0.006778
SUMO1P3	0.006786	LIPT1	0.006796	SNHG10	0.006797
GLS	0.006842	PVR	0.006855	KMT2B	0.006894
YTHDC1	0.006907	DKK4	0.006907	CEACAM5	0.006912
CLU	0.006945	LINC00673	0.006951	SFN	0.006972
KIF18A	0.006974	STAT2	0.007001	MIR107	0.007009
WDR76	0.007011	CASC19	0.007011	DVL2	0.007011
SLC51B	0.007022	SLC25A47	0.007022	SF3B4	0.007034
ASMTL-AS1	0.007049	MITD1	0.007049	LINC00221	0.007049
OTC	0.007049	EGF	0.007061	DUSP12	0.007066
DBP	0.007082	TCF3	0.007090	CD80	0.007103
CDK16	0.007109	MRPS23	0.007112	SPANXC	0.007112
TRIM52	0.007112	RAET1G	0.007112	TPTEP1	0.007112
CREB1	0.007122	MKLN1	0.007136	TUBG1	0.007136
SGOL2	0.007136	HSPA9	0.007151	SCD	0.007169
NFE2L1	0.007170	MGAT4B	0.007199	SLU7	0.007199

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
CDKN2B-AS1	0.007201	HOXD9	0.007209	ZSCAN20	0.007227
N4BP3	0.007227	POLR2J4	0.007227	DHRS13	0.007227
PRR7-AS1	0.007227	TLR8-AS1	0.007227	ACF	0.007227
C3P1	0.007227	LINC00998	0.007227	DUXAP9	0.007227
LINC00624	0.007227	PCNA-AS1	0.007227	LINC01508	0.007227
MAFA-AS1	0.007227	DIABLO	0.007244	RNASE1	0.00725
PCAT6	0.007252	BIRC3	0.007264	DLK1	0.007288
PPP1R1A	0.007292	PDK4	0.007299	AKR1A1	0.007312
LGALS9	0.007313	CEACAM1	0.007317	AKR1C3	0.007338
BIRC2	0.00735	GSTZ1	0.007391	GCK	0.007396
SNRPE	0.0074	TFAP2A	0.007425	SKA3	0.007433
RUSC1	0.007439	MIR499B	0.007439	CSPG5	0.00744
TMEM45A	0.00744	SLC10A5	0.007452	GINS1	0.007453
TRIM16	0.007453	PRR11	0.007483	CD53	0.007493
SLC27A2	0.007497	MAP3K5	0.007497	STK11	0.007503
EEF1E1	0.007505	METTL5	0.007505	CCAT2	0.007514
HFM1	0.007549	CYP2A6	0.007568	ATAD2	0.007569
CLEC1B	0.007578	XBP1	0.007579	KIF23	0.00759
PTP4A1	0.007596	MEX3A	0.007599	CCDC88A	0.007614
ZC3H3	0.007616	TFAM	0.007617	STAT5A	0.007621
ACSL3	0.007622	ANXA3	0.007623	PMAIP1	0.007639
CXCR6	0.007646	TGFBR2	0.007706	HSPB2	0.00772
SOD2	0.007731	HNRNPM	0.007749	CTAG2	0.007749
PCED1B-AS1	0.007767	SLC22A11	0.007768	PTPN1	0.00777
NOTCH2	0.007787	TP73	0.007811	BEX2	0.007821
BAMBI	0.007824	PRDX1	0.007827	MAP1S	0.007832

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
MIR15B	0.007841	FLT1	0.007842	ACLY	0.007848
SOX9	0.007865	SRFBP1	0.007883	AMFR	0.007886
PSRC1	0.007891	SLC46A1	0.007901	SLC48A1	0.007917
TGFBR1	0.00792	C10orf54	0.007949	GOLPH3	0.007956
CYP4A11	0.007967	SRSF10	0.007974	SLC16A3	0.007976
CD151	0.007981	BIRC7	0.007981	BP8	0.007983
FGD5-AS1	0.007983	ALKBH5	0.008009	ADH1C	0.008041
PTP4A3	0.008051	IRS2	0.008054	FOXK1	0.008071
MCM4	0.008074	STAT4	0.008078	APC	0.008111
CXCL14	0.008113	KDM5B	0.008159	HIST1H2AH	0.008168
STAT5B	0.008173	MIR186	0.008175	CCL16	0.008198
STK3	0.008217	CA1	0.008232	UGT1A6	0.008249
AZIN1	0.00826	ROBO1	0.008283	OGG1	0.00829
MIR98	0.008294	TRIM28	0.008307	PRIM1	0.00832
SNRPB	0.008353	AKR1D1	0.008353	RAD51AP1	0.008353
CREB3L3	0.008353	TRIM24	0.008356	MUC1	0.008401
KDM8	0.008421	VPS72	0.008435	OGT	0.008437
ZNFX1	0.008438	MIR198	0.008447	XBP1P1	0.008455
LRRC41	0.008472	FAM160B2	0.008472	DEPDC7	0.008472
RBM43	0.008472	HCG15	0.008472	SAMD12-AS1	0.008472
NEDE	0.008472	SNORA23	0.008472	MIR596	0.008472
LINC00106	0.008472	MSTO2P	0.008472	LMCD1-AS1	0.008472
MIR1269A	0.008472	LINC00882	0.008472	PTOV1-AS1	0.008472
LINC01564	0.008472	SLC39A5	0.008477	DDR1	0.008496
CPD	0.008497	FUT6	0.008509	CXCL9	0.008525
PON1	0.008528	UBAP2	0.008528	MIR499A	0.008540
GTSE1	0.008540	EIF2S2	0.008559	GADD45B	0.008570
SERPINB4	0.008578	CSMD1	0.008578	CXCL10	0.008580

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
CHEK2	0.008594	ATG3	0.008596	E2F4	0.008603
HSP90AB1	0.008608	GDF15	0.008614	CERS2	0.008617
XAF1	0.008617	CYP2C8	0.008618	ESR1	0.008622
IFI27	0.008642	SEPP1	0.008665	FGFR2	0.008683
MX1	0.008697	LPAR6	0.008699	LAMP1	0.008705
CCR6	0.008705	IL6R	0.008711	C3	0.008725
CEBPD	0.008733	NS2	0.008748	CDYL	0.008749
DHX37	0.008749	DANCR	0.008753	FAM122A	0.008760
RABL3	0.008760	LINC00052	0.008763	FBXL19-AS1	0.008763
CDK5RAP3	0.008778	SLC23A1	0.008782	VEGFC	0.008784
CYP8B1	0.008803	TCP1	0.008821	ID1	0.008825
LARP6	0.008876	OXCT1	0.008882	IGFBP3	0.008883
H2AFY	0.008890	ABCB6	0.008909	HOXD3	0.008916
GADD45G	0.008932	KIT	0.008941	PGAM5	0.008942
KIF15	0.008954	C10orf10	0.008970	CBX8	0.008983
RNMT	0.008993	SLC6A6	0.009003	OIP5	0.009011
MIR221	0.009011	SIX1	0.009013	COL15A1	0.009018
PRL	0.009019	CYP2B7P	0.009037	FAM83D	0.009038
CKS1B	0.009062	ECH1	0.009069	SPINK1	0.009081
TAGLN2	0.009092	PXN	0.009101	LINC01554	0.009113
TRADD	0.009116	ANGPTL4	0.009138	APOBEC3B	0.009153
TET1	0.009158	PINX1	0.009190	HPN	0.009190
ATF2	0.009223	CLIC1	0.009232	HNRNPU	0.009250
CIDEB	0.009259	APOA4	0.009261	TLN1	0.009267
IFNAR2	0.009267	KIF14	0.009276	NCL	0.009282
CD86	0.009289	TALDO1	0.009293	SLC38A5	0.009309
AIFM1	0.009320	ZNF32	0.009337	COMMD3	0.009337
C1orf43	0.009337	RAB42	0.009337	RAB42P1	0.009337

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
PITPNA-AS1	0.009337	LOC100419081	0.009337	ESM1	0.009348
CSRP1	0.009370	DAP3	0.009392	NUAK2	0.009392
SPRY4-IT1	0.009392	LCAT	0.009415	POLA2	0.009435
MIR16-1	0.009435	TJP2	0.009471	SLC38A6	0.009488
REN	0.009501	MTCH1	0.009501	LGR5	0.009510
GLI2	0.009511	BDNF	0.009515	IRF2	0.009523
PGK1	0.009532	YTHDF3	0.009533	OR1L3	0.009545
RPUSD2	0.009545	IGHV5-10-1	0.009545	FAM86C1	0.009545
ZNF529	0.009545	METTL25	0.009545	TMEM191A	0.009545
LOC90784	0.009545	PNLIPRP3	0.009545	SAMD13	0.009545
ITPRIPL2	0.009545	ZNF780B	0.009545	DSTNP2	0.009545
NBPF11	0.009545	ALG1L	0.009545	CXorf58	0.009545
ZDHHC24	0.009545	FLJ36000	0.009545	LGALS9B	0.009545
INTS4P1	0.009545	RFESD	0.009545	HIST2H2BD	0.009545
ZNF774	0.009545	FBLL1	0.009545	BTBD17	0.009545
TMEM82	0.009545	C6orf222	0.009545	LINC00999	0.009545
SNORA52	0.009545	CCT6P1	0.009545	PIK3CD-AS1	0.009545
NAV2-AS4	0.009545	FLJ33360	0.009545	LINC00890	0.009545
MIR299	0.009545	LINC00595	0.009545	DUX4L10	0.009545
BW34	0.009545	SERTAD4-AS1	0.009545	MIR522	0.009545
GOLGA8M	0.009545	SCARNA16	0.009545	SNORA11	0.009545
SNORA55	0.009545	SNORD5	0.009545	SNORD88B	0.009545
SNORD105	0.009545	MIR627	0.009545	MIR638	0.009545
KRT16P1	0.009545	PPIAL4E	0.009545	RNA5-8S5	0.009545
DNM1P35	0.009545	TRAPPC3L	0.009545	MRVI1-AS1	0.009545
CSAG4	0.009545	TEKT4P2	0.009545	PPIAP1	0.009545
LOC100294145	0.009545	DUH2	0.009545	MIR548F1	0.009545

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
FTO-IT1	0.009545	CEBPZOS	0.009545	LINC01511	0.009545
PHOSPHO2-KLHL23	0.009545	LINC00506	0.009545	MIR548AQ	0.009545
RNF185-AS1	0.009545	NBPF26	0.009545	SNORD124	0.009545
UCHL1-AS1	0.009545	LINC01149	0.009545	LINC00682	0.009545
CAPN10-AS1	0.009545	IGHV1-69D	0.009545	LINC01273	0.009545
LINC01358	0.009545	LINC01359	0.009545	LINC01391	0.009545
LINC01447	0.009545	LINC01449	0.009545	RP1-13D10.2	0.009545
CTB-63M22.1	0.009545	RP11-87C12.2	0.009545	RP11-1094M14.8	0.009545
SIRT3	0.009555	AKR1C2	0.009563	NUS1	0.009584
SLC22A7	0.009600	SIRT6	0.009611	PPAT	0.009633
ZNF263	0.009638	EPO	0.009642	INSIG1	0.009654
IMMP1L	0.009658	MBTPS1	0.009658	HK1	0.009658
HLA-DRB4	0.009663	UGT1A	0.009672	SCFV	0.009676
IGF1	0.009680	MYL4	0.009707	FEZF1-AS1	0.009709
DCAF4L2	0.009729	ARID3C	0.009729	NBPF12	0.009729
CCDC178	0.009729	DTHD1	0.009729	LINC00210	0.009729
LINC01352	0.009729	CD47	0.009734	SENP1	0.00976
ATP6V0B	0.009773	CDHR2	0.009773	CCDC25	0.009773
RND3	0.009786	MUC13	0.009798	PTH	0.009811
BEX1	0.009815	TYK2	0.009853	LINC00346	0.009862
CLDN10	0.009864	SLC25A1	0.009867	B4GALT1	0.009886
METTL16	0.009894	PPIAP10	0.009908	RBM28	0.009918
CRAMP1L	0.009918	CMTM1	0.009918	MIR9-3	0.009918
PBK	0.009923	CTF1	0.00993	BCAT1	0.009937
EIF2AK3	0.009937	SOAT2	0.00994	KMT2D	0.00994
PSMD14	0.009945	SETDB1	0.009965	BACH1	0.009993
PLAGL2	0.009996	PAK4	0.010009	AQP9	0.010016

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
FAM46C	0.010032	CYP2W1	0.010032	ERVW-1	0.010045
TERC	0.010083	TXNIP	0.010088	TNFSF9	0.010111
BOP1	0.010113	ZNF154	0.010113	CMTM2	0.010113
NSUN4	0.010113	CCAT1	0.010149	VTCN1	0.010188
MT1E	0.010211	PRF1	0.010231	MAGEA4	0.010238
PIGU	0.01026	ATG10	0.010261	NAT10	0.010263
CDKN2C	0.010277	ABCA1	0.010296	SLC27A1	0.010313
SLC30A1	0.010313	LOC100507703	0.010317	CLDN6	0.010319
RPS6	0.010349	MOB2	0.010357	CASC15	0.01036
PART1	0.010361	ESRRB	0.01037	SPINT2	0.01037
ITGA5	0.010393	CDO1	0.010395	PPP2CA	0.010395
SPRY2	0.010401	ING4	0.010433	ERN1	0.010482
HILPDA	0.010484	FTH1	0.010491	PPP1R15A	0.010497
HMGN4	0.010503	GAS2L3	0.010503	C8B	0.010521
NDRG3	0.010521	SNHG17	0.010521	IQGAP1	0.010525
TM4SF1	0.010531	EPHA10	0.010541	GHR	0.010547
ULK1	0.010572	PRMT5	0.010642	CYR61	0.010653
PPIB	0.010662	ERP29	0.010663	REG1A	0.010666
RITA1	0.010686	TRIM71	0.010692	IFNG	0.010707
SRXN1	0.010719	HIPK3	0.010733	IFITM3	0.010739
SIRT4	0.010746	ZNF331	0.010764	MYCT1	0.010769
PDRG1	0.010769	MED8	0.010769	SLC17A1	0.010776
SEA	0.010798	TRIM66	0.010798	DCAF13	0.010798
FIGNL1	0.010798	DDIAS	0.010798	FOXP4-AS1	0.010798
YTHDC2	0.010798	MYT1	0.010827	PRKDC	0.010828
SLC35F5	0.010849	SLC6A11	0.010849	SLC13A1	0.010851
KMT2C	0.010851	TRIM14	0.010859	ANXA6	0.01087
UGT1A9	0.010872	CDK7	0.010917	IL13	0.010921
MIR17HG	0.01093	USP10	0.010933	RORA	0.010939

Table 1: Genes associated with liver cancer (Continued).

Gene	**Sig.**	**Gene**	**Sig.**	**Gene**	**Sig.**
SLC15A5	0.010953	ALDH1L1	0.010983	TP73-AS1	0.010984
ASAP1-IT1	0.010998	ZNF334	0.010998	DDX55	0.010998
CDK15	0.010998	MRPL53	0.010998	LINC01551	0.010998
NME1-NME2	0.010998	LINC00630	0.010998	LINC01572	0.010998
HSPA4L	0.011004	PLIN2	0.011012	KLF9	0.011014
MICB	0.011028	MIR100	0.01103	AQP7	0.011063
CDR1-AS	0.01107	GOT1	0.011075	JOSD2	0.011092
THUMPD3-AS1	0.011092	MACC1-AS1	0.011092	UHRF2	0.011097
AGO2	0.011106	CTSD	0.011118	PLIN3	0.01112
TLR3	0.011126	SLC37A2	0.011139	ROMO1	0.011153
CXCL12	0.011167	CDH17	0.011178	PTDSS2	0.011179
POMC	0.0112	DLX6	0.011207	GPX2	0.011215
TARBP1	0.011228	1-Mar	0.011228	AFAP1L2	0.011228
CMTM3	0.011228	BTG2	0.011269	RUVBL2	0.011298
ADAM17	0.011312	SOCS6	0.011337	IL17D	0.01134
EPT1	0.011344	DMGDH	0.011345	MAGI2-AS3	0.011345
CXCL5	0.011356	CP	0.011386	FGL2	0.011388
CD163	0.011404	FCN2	0.01141	FGG	0.011418