Aus dem Zentrum für Operative Medizin der Universität zu Köln
Klinik und Poliklinik für Allgemein-, Viszeral-, Tumorchirurgie und Transplantationschirurgie
Direktorin: Universitätsprofessorin Dr. med. C. Bruns

The landscape of mutation in plasma circulating tumor DNA sequencing as potential predictive biomarkers in hepatocellular carcinoma

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von
Xiaolin Wu
aus Yunnan China

promoviert am 27.Juni 2024

Gedruckt mit Genehmigung der Medizinischen Fakultät der Universität zu Köln

2024

Dekan: Universitätsprofessor Dr. med. G. R. Fink

1. Gutachter: Professor Dr. med. R. Wahba
2. Gutachterin: Professorin Dr. med. U. Drebber

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich keine Unterstützungsleistungen von folgenden Personen erhalten:

Professor Dr. med. Roger Wahba

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen. Dissertationsschrift

Die wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die Versuchsplanung, das Screening von mutierten Genen für die Aufnahme in das Sequenzpanel, die Auswahl der Patienten und die Kombination von klinischen und Sequenzdaten für die Analyse wurden von mir selbst durchgeführt. Margaux Bamberger half bei den klinischen Daten und der Nachbeobachtung der Patienten. Die Extraktion von zellfreier DNA, Keimbahn-DNA und Tumor-DNA wurde gemeinsam mit Anke Wienand-Dorweiler und Michaela Heitmann durchgeführt. Die Entnahme von FFPE-Proben wurde von Prof. Uta Drebber und Susanne Neiss unterstützt. Dr. Janine Altmüller und Dr. Kerstin Becker assistierten bei der Sequenzierung. Die Primäranalyse der Sequenzier-Rohdaten wurde von PD Dr. Sven Borchmann und Dr. Jan-Michel Heger durchgeführt.

Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Köln, den 23.01.2024

Unterschrift: .....Xiaolin Wu.....

Acknowledgment

As time flies, my medical doctoral research is coming to an end. Looking back on my busy and precious study experience, I feel so joyful and satisfied. Although COVID-19 has brought huge challenges to my research, I have successfully made it to this point with the support of professors, teachers, friends, and families. At this time of farewell, a simple "thanks" is not enough. I hope to write down my gratitude to all those who helped me during the writing of this dissertation.

Firstly, I would like to express my deepest appreciation to Prof. Dr. Roger Wahba, my research supervisor, for his unwavering support throughout my M.D. studies. He has been an exceptional mentor, and I am truly thankful for his guidance and patience. I am also grateful to Dr. Yue Zhao, who gave me full guidance, endless encouragement, and gentle companionship, especially during my life's sad and depressing times. Meanwhile, I would like to express gratitude to Prof. Dr. Christiane J. Bruns, for her strong support and feedback on my project and publication. They allowed me to do medical research in Germany and helped me complete my studies. I will always remember this kindness in my heart.

I extend my sincere gratitude to Dr. Jiahui Li, for his valuable insightful feedback, and support throughout my research. I also feel grateful to Dr. Jiangang Zhao, Dr. Zhefang Wang, Dr. Chenghui Zhou, and Qiye Sun for their invaluable expertise in my academic and personal development. I would like to thank Dr. Dai Li, Dr. Jie Qin, Ningbo Fan, and Feng Ju for providing a nurturing academic environment that has enriched my learning experience.

Then, I want to appreciate: Prof. Dr. Hakan Alakus and Dr. Asmae Gassa for their ctDNA-related experience, and powerful support of the preliminary experiment and funding application; Prof. Dr. Uta Drebber and Dr. Heihe Löser for providing HCC paraffin samples; Dr. Janine Altmüller and Elisabeth Kirst for next-generation sequencing; Dr. Sven Borchmann and Dr. Jan-Michel Heger for data analysis; Dr. Markus Ball for the technical guidance of Cell-free DNA ScreenTape; Prof. Dr. Margarete Odenthal for the helpful advice on the ctDNA manuscripts.

Moreover, I would like to sincerely thank Susanne Neiß, Anke Wienand-Dorweiler, Michaela Heitman, and Lisa Raatz, my laboratory colleagues, for their full support and care in my work and life here. And I also want to thank Margaux Bamberger for the help with HCC survival information.

In addition, I hope to acknowledge my family for their unconditional love and encouragement. I also want to thank my friends who are in China: Mengdie Ou, Lizhu Bao, and Danyu Lin. Your

love and support for me remain steadfast during my M.D. journey, even with a six-hour time difference between us.

Thank you all for your invaluable contribution to my M.D. journey. No matter what I may experience in the future, I will always miss you all and the joyful time we spent together!

Dedication: this dissertation is dedicated to all the people who never give up in adversity.

Content

ABBREVIATIONS 8

1. SUMMARY 15

2. INTRODUCTION 17

2.1. Hepatocellular carcinoma 17
2.1.1. Introduction 17
2.1.2. Diagnosis of HCC 17
2.1.3. Biomarkers of HCC 18

2.2. Liquid biopsy 19
2.2.1. Introduction 19
2.2.2. ctDNA/cfDNA 20
2.2.3. Typical biomarkers of ctDNA 21

2.3. Genetic mutation 22
2.3.1. Introduction 22
2.3.2. The mutation of ctDNA 23

2.4. Research aim 25

3. MATERIALS AND METHODS 27

3.1. Materials 27
3.1.1. Human samples 27
3.1.2. Materials of ctDNA/cfDNA extraction 27
3.1.3. Materials of ctDNA/cfDNA quantification 27
3.1.4. Materials for library preparation and NGS 28
3.1.5. Laboratory equipment 28
3.1.6. Consumable materials 29
3.1.7. Software 30

3.2. Methods 30
3.2.1. Patients 30
3.2.2. CfDNA extraction 31
3.2.3. CfDNA quantification and quality appraisal 31
3.2.4. Germline-DNA extraction 31
3.2.5. Tumor-DNA extraction 31
3.2.6. HCC sequencing panel design 32

3.2.7. Library preparation and NGS 35
3.2.8. Basic data processing, variant calling and filtering 35
3.2.9. Gene Ontology (GO) enrichment 36
3.2.10. Statistical analyses 36

4. RESULTS 37

4.1. Clinical characteristics of enrolled HCC patients 37

4.2. The quality test for cfDNA of HCC 41

4.3. Genetic mutation of ctDNA showed an effect in distinguishing ctDNA and cfDNA 43

4.4. Landscapes of mutations of ctDNA and associated signaling pathways 45
4.4.1. Genetic landscape of ctDNA 48
4.4.2. Mutant genes in ctDNA and tDNA 49
4.4.3. Mutation genes and signaling pathways 50

4.5. Concordant mutations and correlation with the clinical factors 52
4.5.1. Mutation concordance between plasma ctDNA and matched HCC tDNA 52
4.5.2. Clinical variable and concordant mutations 55

4.6. CtDNA and HCC diagnosis and prognosis 57
4.6.1. Combination of ctDNA and AFP for HCC diagnosis 57
4.6.2. The specific mutation set in ctDNA and prediction of HCC survival 59

5. DISCUSSION 64

6. REFERENCE 69

ABBREVIATIONS

ACVR2A: activin receptor type-2A

AFF3: ALF transcription elongation factor 3

AFP: alpha-fetoprotein

ALB: albumin

ALK: ALK receptor tyrosine kinase

APC: adenomatous polyposis coli protein

ARID1A: AT-rich interactive domain-containing protein 1A

ARID1B: AT-rich interactive domain-containing protein 1B

ARID2: AT-rich interactive domain 2

ASXL1: additional sex combs like transcriptional regulator 1

ATM: ATM serine/threonine kinase

ATR: ATR serine/threonine kinase

*AXIN*1: axis inhibition protein 1

AUC: area under the curve

BAP1: BRCA1 associated protein 1

BCLC: Barcelona clinic liver cancer

BCORL1: BCL6 corepressor like 1

BIOMASOTA: Biologische Material Sammlung zur Optimierung Therapeutischer Ansätze

BLM: BLM RecQ like helicase

BRCA1: breast cancer gene 1

BRCA2: breast cancer gene 2

CARD11: caspase recruitment domain family member 11

CDKN1A: cyclin dependent kinase inhibitor 1A

CDKN2A: cyclin dependent kinase inhibitor 2A

CES: comparative error suppression

CHD2: chromodomain helicase DNA binding protein 2

CI: confidence interval

CREBBP: CREB binding protein

CT: computed tomography

cfDNA: cell-free DNA

ctDNA: circulating tumor DNA

CTCs: circulating tumor cells

CTP: Child-Turcotte-Pugh

CTNNB1: catenin beta 1

CNVs: copy number variations

DCP: Des-Gamma-Carboxy prothrombin

ddPCR: digital droplet PCR

DKK1: Dikkopf-1

EGFR: epidermal growth factor receptor

EP300: E1A binding protein P300

EPHA3: EPH receptor A3

ERBB2: Erb-B2 receptor tyrosine kinase 2

ERBB4: Erb-B2 receptor tyrosine kinase 4

ERCC5: ERCC excision repair 5, endonuclease

FAT4: FAT atypical cadherin 4

FFPE: formalin-fixed paraffin-embedded tissue

FGFR2: fibroblast growth factor receptor 2

FGFR3: fibroblast growth factor receptor 3

FGFR4: fibroblast growth factor receptor 4

FLT3: Fms related receptor tyrosine kinase 4

FLT4: Fms related receptor tyrosine kinase 4

GO: gene ontology

GP73: Golgi Protein-73

GPC3: Glypican 3

HCC: hepatocellular carcinoma

HBV: hepatitis B virus

HCV: hepatitis C virus

hGE: haploid genome equivalents

HGF: hepatocyte growth factor

HNF1A: HNF1 homeobox A

IDH2: isocitrate dehydrogenase (NADP(+)) 2

IGF1R: insulin like growth factor 1 receptor

IL6ST: interleukin 6 cytokine family signal transducer

IQR: interquartile range

JAK1: janus kinase 1

KEAP1: Kelch like ECH associated protein 1

KDM6A: lysine demethylase 6A

KIT: KIT proto-oncogene, receptor tyrosine kinase

KMT2A: lysine methyltransferase 2A

KMT2B: lysine methyltransferase 2B

KMT2C: lysine methyltransferase 2C

KMT2D: lysine methyltransferase 2D

KRAS: KRAS proto-oncogene, GTPase

LRP1B: LDL receptor related protein 1B

LRRK2: leucine rich repeat kinase 2

mAF: mean mutated allele frequency

MAPK: mitogen-activated protein kinase 1

MDK: Midikine

MECOM: MDS1 and EVI1 complex locus

MED12: mediator complex subunit 12

MET: MET proto-oncogene, receptor tyrosine kinase

MGA: max dimerization protein MGA

MGAM: maltase-glucoamylase

MKI67: marker of proliferation Ki-67

MRI: magnetic resonance imaging

MSH6: MutS homolog 6

MTOR: mechanistic target of rapamycin kinase

MYC: MYC proto-oncogene, BHLH transcription factor

MYO18A: myosin XVIIIA

NAFLD: nonalcoholic fatty liver disease

NCOR1: nuclear receptor corepressor 1

NCOR2: nuclear receptor corepressor 2

NF1: Neurofibromin 1

NFE2L2: NFE2 like BZIP transcription factor 2

NGS: next-generation sequencing

NOTCH1: neurogenic locus Notch homolog protein 1

NOTCH3: neurogenic locus Notch homolog protein 3

NOTCH4: neurogenic locus Notch homolog protein 4

NRAS: NRAS proto-oncogene, GTPase

NTRK2: neurotrophic receptor tyrosine kinase 2

NTRK3: neurotrophic receptor tyrosine kinase 3

NUP214: nucleoporin 214

OPN: Osteopontin

OS: overall survival

PAK5: P21 (RAC1) activated kinase 5

PBRM1: polybromo 1

PCLO: Piccolo presynaptic cytomatrix protein

PCR: polymerase-chain-reaction

PDE4DIP: phosphodiesterase 4D interacting protein

PDGFRA: platelet derived growth factor receptor alpha

PFS: progression-free survival

PIK3CA: phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha

PIK3CG: phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma

POLQ: DNA polymerase Theta

PREX2: phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 2

PRKDC: protein kinase, DNA-activated, catalytic subunit

pTNM: pathology TNM stage

PTEN: phosphatase and tensin homolog

PTPN13: protein tyrosine phosphatase non-receptor type 13

PTPRB: protein tyrosine phosphatase receptor type B

PTPRT: protein tyrosine phosphatase receptor type T

RAD50: RAD50 double strand break repair protein

RAS: RAS GTPase

RANBP2: RAN binding protein 2

RB1: retinoblastoma protein 1

RELN: reelin

RNF213: ring finger protein 213

ROBO1: roundabout guidance receptor 1

ROC: receiver operating characteristic

SETD2: SET domain containing 2, histone lysine methyltransferase

SF3B1: splicing factor 3b subunit 1

SMARCA4: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4

SNV: single-nucleotide polymorphism

SPEN: Spen family transcriptional repressor

SWI/SNF: Switch/Sucrose Non-Fermentable

TCGA: the cancer genome atlas

TAF1: TATA-box binding protein associated factor 1

TERT: telomerase reverse transcriptase

TNM: tumor nodes metastasis

TP53: tumor protein p53

TRRAP: transformation/transcription domain associated protein

TSC1: TSC complex subunit 1

TSC2: TSC complex subunit 1

WGS: whole genome sequencing

WRN: WRN RecQ like helicase

ZFHX3: zinc finger homeobox 3

ZNF521: zinc finger protein 521

1. SUMMARY

Recently, the advancements in liquid biopsy have facilitated the techniques for clinical diagnosis and treatment monitoring of hepatocellular carcinoma. Genetic mutations are practical markers of distinguishing ctDNA, eliminating interference from cfDNA. Our study conducts a sequencing panel for ctDNA/cfDNA in patients with HCC and patients with benign liver disease, using NGS technology to identify mutant targets. After excluding germline and silent mutations, we obtain the mutation profiles of ctDNA in HCC, comprising solely functional mutations.

Genetic mutations were found in free nucleic acid from 66.7% of HCC patients, while no mutant gene was in the control group. In our HCC cohort, ctDNA analysis was constituted of 49 genes and 91 exon mutations, with 15 genes (*NCOR2, HGF, MECOM, ROBO1, MKI67, PEPN13, RANBP2, RELN, ALB, FAT4, KMT2B, MGAM, PAK5, PTPRB, ZFHX3*) being identified for the first time in the ctDNA of HCC. *NOCR2* and *CTNNB1* were the highest frequent mutant genes in ctDNA, reaching 13.3%. The majority of these mutant genes were distributed in the classical molecular pathways of HCC, and the gene-enriched pathways showed a strong consistency between ctDNA and tDNA. A total of nineteen concordant mutations were detected in both ctDNA and matched tDNA, with 23 exons. We also found that the ratio of concordant mutation was highly correlated to tumor burden, especially vascular invasion. No mutations were found in the cfDNA of the control group, suggesting that mutant genes in ctDNA exhibit the potential to differentiate between benign and malignant liver diseases. Consequently, we further explored the diagnostic capabilities of ctDNA and discovered a great improvement in diagnostic accuracy of a combination of ctDNA mutation and AFP level over either one alone. Additionally, our research found the specific mutation-based gene set from ctDNA could contribute to predicting the prognosis of HCC patients. The mutation set screened according to TNM stages 2-4 consisted of twelve genes: *NCOR2, ARID2, ERBB4, ERCC5, KMT2A, MSH6, PIK3CA, PIK3CG, POLQ, PEPRB, TERT*, and *TSC1*. This analysis showed that the mutation of the *NCOR2* gene was detected particularly frequently in HCC ctDNA and could therefore indicate a high potential for the prognosis of HCC patients.

In total, these findings have demonstrated the potential of ctDNA mutation as a specific biomarker for liquid biopsy technique and deepened our understanding of the mutation profile in HCC. Our research supports the implications of mutations in ctDNA for precision medicine and illustrates the clinical prospect of ctDNA in the future.

Zusammenfassung

In der heutigen Zeit haben die Fortschritte bei der Liquid-Biopsy die Techniken für die klinische Diagnose und die Überwachung der Behandlung des Leberzellkarzinoms erleichtert. Genetische Mutationen sind Merkmale zur Identifizierung ctDNA im Blut und dienen zusätzlich zur Unterscheidung dieser von cfDNA. In unserer Studie wird ein Sequenzierungspanel für ctDNA/cfDNA bei Patienten mit HCC und Patienten mit gutartigen Lebererkrankungen durchgeführt, wobei die NGS-Technologie zur Identifizierung von Mutationen eingesetzt wird. Nach Ausschluss von Keimbahn- und stillen Mutationen erhalten wir die Mutationsprofile der ctDNA bei HCC, die ausschließlich funktionelle Mutationen umfassen. Genetische Mutationen wurden in der ctDNA von 66,7 % der HCC-Patienten gefunden, während in der Kontrollgruppe kein mutiertes Gen zu finden war. In unserer HCC-Kohorte umfasste die ctDNA-Analyse 49 Gene und 91 Exon-Mutationen, wobei 15 Gene (*NCOR2, HGF, MECOM, ROBO1, MKI67, PEPN13, RANBP2, RELN, ALB, FAT4, KMT2B, MGAM, PAK5, PTPRB, ZFHX3*) zum ersten Mal in der ctDNA von HCC Patienten identifiziert wurden. *NCOR2* und *CTNNB1* waren mit 13,3 % die am häufigsten mutierten Gene in der ctDNA. Die meisten dieser mutierten Gene waren in den für HCC klassischen molekularen Pathways verteilt, wobei diese eine starke Übereinstimmung zwischen ctDNA und tDNA aufzeigten. Insgesamt wurden 19 übereinstimmende Mutationen mit 23 Exons sowohl in der ctDNA als auch in der korrespondierenden tDNA nachgewiesen. Zudem fanden wir heraus, dass das Verhältnis der übereinstimmenden Mutationen sowohl mit der Tumorlast als auch der makrovaskulären und der mikrovaskulären vaskulären Invasion korreliert. In der cfDNA der Kontrollgruppe wurden keine Mutationen gefunden, was darauf schließen lässt, dass mutierte Gene in der ctDNA das Potenzial haben, zwischen gutartigen und bösartigen Lebererkrankungen zu unterscheiden. Folglich untersuchten wir die diagnostischen Möglichkeiten der ctDNA weiter und entdeckten, dass die Kombination von ctDNA-Mutation und AFP-Spiegel die diagnostische Genauigkeit gegenüber einem der beiden Werte allein erheblich verbessert. Darüber hinaus fanden wir heraus, dass der spezifische, auf Mutationen basierende Gensatz aus ctDNA zur Vorhersage der Prognose von HCC-Patienten beitragen kann. Der nach den TNM-Stadien 2-4 untersuchte Mutationssatz bestand aus zwölf Genen. Bei dieser Analyse zeigte sich, dass die Mutation des Gens *NCOR2* besonders häufig in der HCC ctDNA detektiert wurde und somit auf ein hohes Potenzial zur Prognose von HCC-Patienten hindeuten könnte.

Insgesamt haben diese Ergebnisse das Potenzial von ctDNA-Mutationen als spezifischer Biomarker für die Liquid-Biopsy gezeigt und unser Verständnis des Mutationsprofils beim HCC vertieft. Unsere Forschung hebt die Bedeutung von Mutationen in der ctDNA von HCC-Patienten für die Präzisionsmedizin hervor und veranschaulicht die klinischen Aussichten dieser Methode in der Zukunft.

2. INTRODUCTION

2.1. Hepatocellular carcinoma

2.1.1. Introduction

Primary liver cancer is one of the most globally prevalent malignant tumors, presenting a grave prognosis with a 5-year survival rate of only 18% [1]. It is the 6th most frequently diagnosed cancer and the third leading cause of cancer-related fatalities [2,3]. More than 900000 people were diagnosed with liver cancer globally in 2020. Still, the incidence of liver cancer continues to surge: projections indicate an increase of 55.0% between 2020 and 2040, which suggests nearly 1.4 million individuals may be affected by liver cancer, and 1.3 million fatalities resulting from the disease in 2040 [1]. Nowadays, primary liver cancer has become a major health problem worldwide. Hepatocellular carcinoma (HCC) is a significant type of primary liver carcinoma, comprising nearly 90% of all cases [4].

More than 90% of HCC cases develop due to chronic liver disease, and the risk factors may cause the development of liver cirrhosis and further lead to malignant tumors [5]. Removing risk factors by treating the cause of liver cancer is the only method to reduce the incidence of HCC. The significant risk factors of HCC are virus infection (hepatitis virus B & C) and alcohol, accounting for 84% of HCC deaths [6]. Hepatitis B (HBV) can induce chronic inflammatory disease and promote mutations in liver cells, leading to HCC; thus, HBV carriers undertake the lifetime risk of HCC ranging from 10-25% [7]. Nearly 20% of chronic HCV patients will develop cirrhosis in 20-30 years, significantly increasing the risk for HCC [8]. Alcohol-associated cirrhosis brings a high cumulative incidence of HCC at 10-year follow-up, reaching 9% [9]. Besides virus and alcohol, nonalcoholic fatty liver disease (NAFLD) has become a significant etiology of HCC, with the annual incidence of HCC 2.4–12.8% [10]. Aflatoxin exposure is another risk factor that promotes the onset of HCC, especially for patients with the hepatitis virus [11].

Additionally, age and gender have been associated with HCC as well. The incidence of HCC increases with age, with the highest period of initial diagnosis being 70-79 years [12]. Epidemiological studies have displayed a higher prevalence of HCC among males than females due to many factors, including decreased adiponectin levels in men [13].

2.1.2. Diagnosis of HCC

HCC patients demonstrate a vast disparity in prognosis between the early and late stages. Patients in the early stage experience a significantly improved 5-year survival rate exceeding 70%, while those in the late stage merely show a 5-year survival rate of less than 16% [14].

Despite the systemic therapies, patients with HCC in advanced-stage typically have a median survival of 1-1.5 years [2]. Therefore, early diagnosis plays a crucial role in HCC as it facilitates the availability of multiple curative therapy options. Regrettably, HCC generally remains asymptomatic and missing inflammation during the early stages, leading to the challenge of detecting underlying cirrhosis with progression to carcinoma. The diagnosis of HCC heavily relies on modern medical technology till now.

Currently, the standard clinical method for early diagnosis of HCC involves a combination of imaging techniques and alpha-fetoprotein (AFP) measurement. The primary noninvasive imaging methods employed for HCC diagnosis are computed tomography (CT) and magnetic resonance imaging (MRI). Many studies have shown that MRI carries a higher sensitivity than CT, with a specificity of 85%-100% [15]. Contrast-enhanced ultrasound is not typically considered the primary choice for HCC because of the risk of diagnosis, making it a secondary option following CT and MRI [5]. Regardless of the imaging technique, only show favorable performance in detecting HCCs larger than 2cm in diameter. However, their effectiveness notably declines when detecting HCCs smaller than 1cm [15]. Although AFP is widely utilized as a tumor biomarker for HCC diagnosis and monitoring, its cut-off value of 20 ng/ml only exhibits a sensitivity of 62.4% and specificity of 89.4%, indicating that it lacks sufficient accuracy [16].

Histopathology is the gold standard for HCC diagnosis, with the characteristics of increased cell density, wide trabeculae (> three cells), obvious acinar pattern, mitotic activity, absence of Kuffer cells, and vascular invasion [17]. Undoubtedly, the pathological biopsy is a highly accurate diagnostic method for defining HCC. Nevertheless, due to its invasive status, it is limited in frequency of use and lacks the capacity to detect HCC in real time.

2.1.3. Biomarkers of HCC

Due to the unsatisfactory current diagnostic methods for HCC, scientists have shifted their research focus to blood-based biomarkers, which are recognized for significant potent in enhancing tumor detection. Several biomarkers in early-stage validation are shown in Table 1, including AFP-L3, des-gamma carboxyprothrombin (DCP), dikkop-1 (DKK1), golgi protein-73 (GP73), osteopontin (OPN), midikine (MDK) and Glypican 3 (GPC3). However, a comprehensive evaluation of the sensitivity and specificity indicates that their diagnostic validity is not significantly superior to AFP[18].

Consequently, the early diagnosis of HCC remains a significant challenge, necessitating the development of more precise technologies. The ideal tools for diagnosis and surveillance must be highly accurate, reproducible, and not dependent on clinical settings. In this regard, liquid biopsy is a promising solution that fulfills all these requirements.

Table 1. Blood-based biomarkers for diagnosis of HCC

Biomarker	Sensitivity	Specificity	Reference
AFP L3	42%	97%	[19]
DCP	74%	70%	[19]
DKK1	69.1%	90.6%	[20]
GP73	69%	75%	[21]
OPN	74%	66%	[22]
MDK	86.9%	83.9%	[23]
GPC3	60%	52%	[24]

2.2. Liquid biopsy

2.2.1. Introduction

Liquid biopsy offers a promising solution to the challenges in HCC. With the advancements in the integration of 'omics information,' cancer research could be more comprehensive in clinical applications, including diagnosis, prognostic predicting, monitoring, and therapy [25]. This noninvasive approach utilizes blood and other bodily fluids sampling, facilitating a deeper understanding of the disease [26].

Liquid biopsy could involve different components: circulating tumor DNA (ctDNA), cell-free DNA (cfDNA), circulating tumor cells (CTCs), microRNA, and extracellular vesicles [27]. Among them, ctDNA/cfDNA and CTCs are both the vital cornerstones of liquid biopsy (Fig.1). CTCs are the tumor cells that have sloughed off the primary tumor and released into and circulate in the blood from the primary cancer [28]. Extracellular vesicles are nanoscale capsules released from cells, possess a lipid bilayer membrane and contain protein, DNA and RNA for cell communication [29]. Circulating miRNA is secreted from apoptosis, inflammation, and necrosis cells, potentially becoming cancer biomarkers [30].

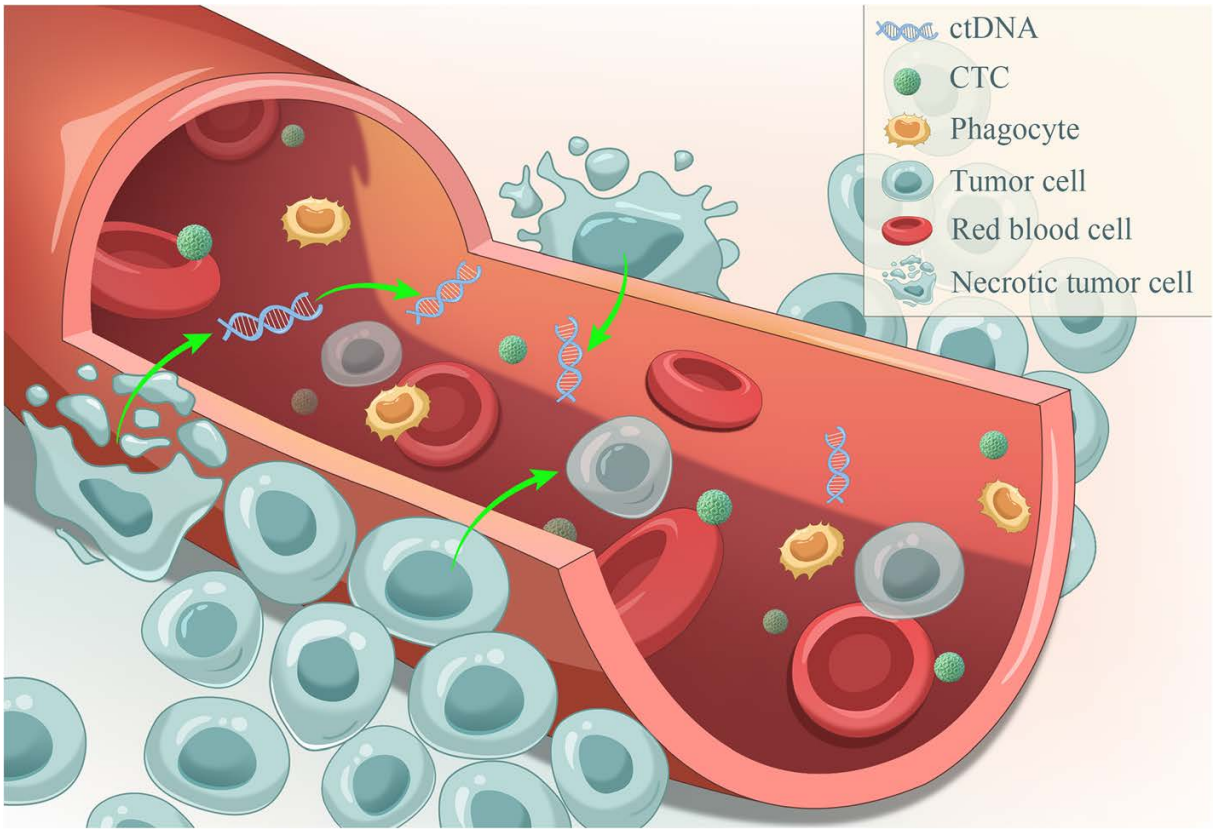


Figure 1. Liquid biopsy: ctDNA and CTCs are easily accessible in peripheral blood. The ctDNA is free nucleic acid fragments released from tumor cells undergoing apoptosis or necrosis, and CTCs are cancer cells naturally shed from the primary or metastatic tumors. They are both precise tumor markers that perform a large amount of tumor information [27].

2.2.2. ctDNA/cfDNA

As a crucial element of liquid biopsy, ctDNA/cfDNA consists of double-stranded DNA fragments measuring approximately 150-200 base pairs in length [31]. Free nucleic acid fragments in human peripheral blood were first reported by Mandel and Metais in 1948 [32]. Leon and colleagues demonstrated the cfDNA level increased significantly in more than half of cancer patients than in normal control individuals in 1977 [33]. The ctDNA/cfDNA population generally peaks at 166 bp, but in cancer patients, ct/cfDNA exhibits more fragments, spread between 40 and 150 bp [34]. The definitions of cfDNA and ctDNA have slight distinctions: cfDNA is derived from both healthy and malignant cells undergoing necrosis or apoptosis and is released into the circulatory system. In contrast, ctDNA is released explicitly from tumor cells [35]. In healthy individuals, cfDNA levels are typically low, with average concentrations ranging from 10 to 15 ng/ml, but they can increase in response to tumor, inflammation, or tissue damage [35]. The high quantity of ctDNA is one of the essential reasons for the elevation in

cfDNA. Since ctDNA is derived from tumor cells, it carries specific tumor-related information and has emerged as a potential alternative source for molecular profiling in cancer patients [36]. With a half-life of less than one hour, ctDNA provides real-time insights into the dynamic progression of the carcinoma [35].

These constitute the key advantages of ctDNA in clinical applications, mainly its potential for early cancer detection, continuous monitoring and prognostication [37]. Moreover, ctDNA has been shown to have the ability to make molecular genotypes and detect acquired chemoresistance [38].

With the advances in molecular and computational biology in recent years, sequencing approaches for ctDNA have been significantly improved, including next-generation sequencing (NGS), droplet digital PCR (ddPCR), whole genome sequencing (WGS) and other analysis methods based on fragment omics of epigenetic feature [39]. DNA sequencing technology has been developed for over 40 years [40]. Sanger sequencing is the primary and standard sequencing method, with a capacity shortage and high consumable cost [41]. NGS is an optimizing DNA sequencing technology over traditional Sanger sequencing, with higher parallelism than before [42]. NGS could produce an enormous number of DNA sequencing data and shows the advantage of lower cost [43]. The ddPCR is the third-generation sequencing technology that uses water-in-oil droplets for high-throughput technology for PCR [44]. This high-throughput DNA sequencing technology displays excellent sensitivity and specificity in cancer diagnosis [45]. NGS and ddPCR are currently the dominant analysis methods for ctDNA research. NGS holds a broad detection range and encompasses the whole genome or hundreds to the whole exome, whereas ddPCR has a superior sensitivity of 0.1%–0.001% [46]. NGS is generally used for comprehensive gene screening, whereas ddPCR is well-suited for the targeted detection of a few known gene mutations.

However, the field of ctDNA research and clinical utility encounters a primary challenge: how to recognize the ctDNA from normal cfDNA precisely?

2.2.3. Typical biomarkers of ctDNA

CtDNA displays a variable proportion within cfDNA, ranging from less than 0.1% to over 90% [35]. Therefore, it is hard to discriminate ctDNA in the background of normal cfDNA accurately. Identifying specific biomarkers in ctDNA may be an effective solution to this problem. The molecular biomarkers of ctDNA should be detectable in both tumor cells and plasma DNA within the same individual while remaining absent in the cfDNA.

Since free nucleic acid fragments are shed from tumor cells, ctDNA carries a vast amount of genetic information about the tumor, which could be classic biomarkers (Fig. 2). The majority of studies published recently have concentrated on investigating genetic variations, such as copy number variations (CNVs), gene integrity, genetic mutations, and methylation, as we previously reviewed [47-51].

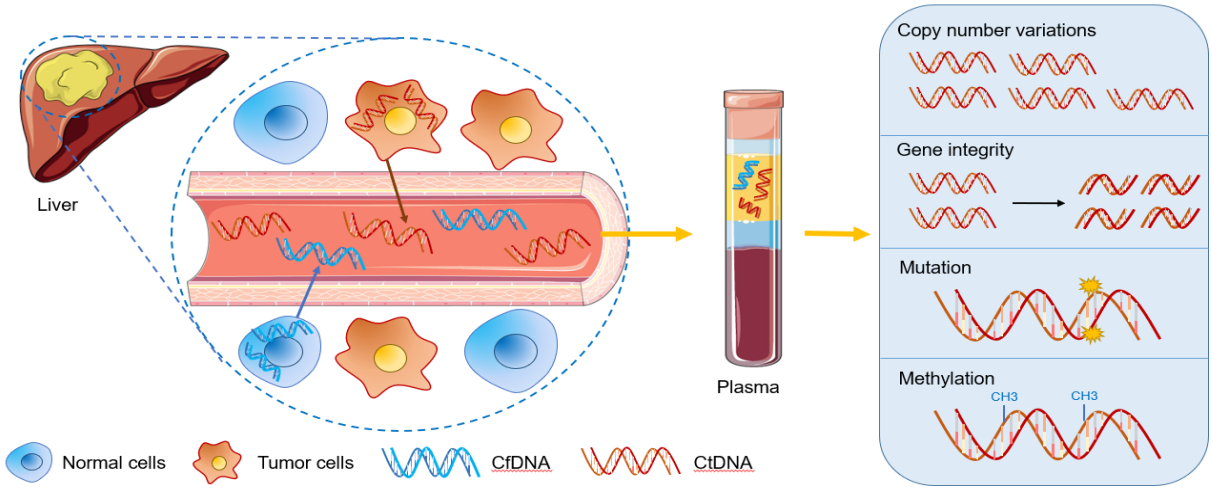


Figure 2. As presented in the figure, both ctDNA and cfDNA can be identified in the peripheral blood of individuals with HCC. CfDNA is released from normal cells across the body, while ctDNA solely originates from tumor cells. Identifying ctDNA within cfDNA necessitates the recognition of specific ctDNA biomarkers that rely on genetic changes in ctDNA, including copy number variations, gene integrity, mutations, and methylation [51].

2.3. Genetic mutation

2.3.1. Introduction

Almost all cancers of human have a common characteristic: genomic instability [52]. Genetic mutations accumulated gradually in somatic cells, with the majority being harmless; however, some mutations serve as key contributors to tumor [53]. In malignant hepatocytes, the accumulation of mutations in DNA shows the mechanisms of HCC development [54]. All tumors (including HCC) are essentially caused by somatic mutation, which could be the molecular fingerprint carried in tumor genome [55]. The causes of somatic mutations fall into two broad categories: endogenous (such as exposition to hepatitis virus, aflatoxin, etc) and exogenous (such as age, DNA repair mechanisms defection, etc) [54]. Mutation-specific research helps scientists explore the malignant transformation of hepatocytes. The common genes in HCC nodules with high frequency were: *TERT* promoter (44%), *TP53* (31%), *CTNNB1* (27%), *AXIN1* (8%), *ARID1A* (7%) [56].

2.3.2. The mutation of ctDNA

In recent years, more studies have explored the potential of mutations as specific biomarkers in ctDNA [57] [58]. Mutant genes significantly affect clinical application, as evidenced by numerous studies [59]. The mutations in ctDNA provide the tumor information from the primary malignant nodules and relate to tumor burden. In HCC cohorts of Europe, at least one mutation was detectable in 86% of HCC cases with a tumor diameter over 5cm or metastasis, displaying the ability of ctDNA to capture genetic information that corresponds to the condition of cancer [60]. Several common mutations identified in tDNA from HCC tissues could also be detected in ctDNA from blood, such as ARID1A, *CTNNB1* and *TP53* [61]. The frequency of mutations in ctDNA is random and may be influenced by many clinical factors, such as tumor stage, viral infection, etc. In advanced HCC, Johann von Felden et al. showed the ctDNA frequency of *ARID1A*, *CTNNB1* and *TP53* mutations were 6%, 17% and 32%, respectively, in addition to the other ctDNA mutations: *TERT* promoter (51%), *Axin1* (6%) [62]. Among the HCC cohort in Chinese, 22/66 patients carried mutated genes: *TP53* exhibited the highest mutation rate at 60.0%, followed by *CTNNB1* at 15.7%, *Axin1* and *ARID1A* at 14.3% [63].

Despite the high-frequency mutations in HCC, some low-frequent mutations could be identified in ctDNA. In the study conducted by Lim HY and colleagues, they investigated the *RAS* (*KRAS* and *NRAS*) gene in a small cohort of HCC patients (27 cases) and found RAS mutational status could be confirmed in 44% of HCC patients using NGS, whereas the frequency of *RAS* mutations in tumor tissue was only 1-2% [64] [56]. These discoveries suggested the mutation of ctDNA may be a promising approach for exploring novel diagnosis, monitoring, and prognosis evaluation methods.

The mutant genes commonly confirmed in the plasma of HCC patients contain *TERT, CTNNB1, TP53, Axin1, ARID1A, KRAS* and *NRAS*. These genes could be classified into oncogenes (*TERT, CTNNB1, KRAS, NRAS*) and tumor suppressor genes (*TP53, Axin1, ARID1A*) [65] [66] [67] [68].

Moreover, the typical mutant ctDNA genes are integral to crucial molecular signaling pathways, which greatly contribute to the development and advancement of HCC, as depicted in our review paper (Fig. 3). The pathways implicated are the *RAS/MAPK* pathway, Telomere maintenance mechanism (TMM) pathway, p53 signaling pathway, Wnt-β catenin pathway, and *SWI/SNF* complex-related signaling pathway [51]. These pathways are associated with various vital functions, including proliferation, immortalization, cell differentiation, genomic stability and prognosis.

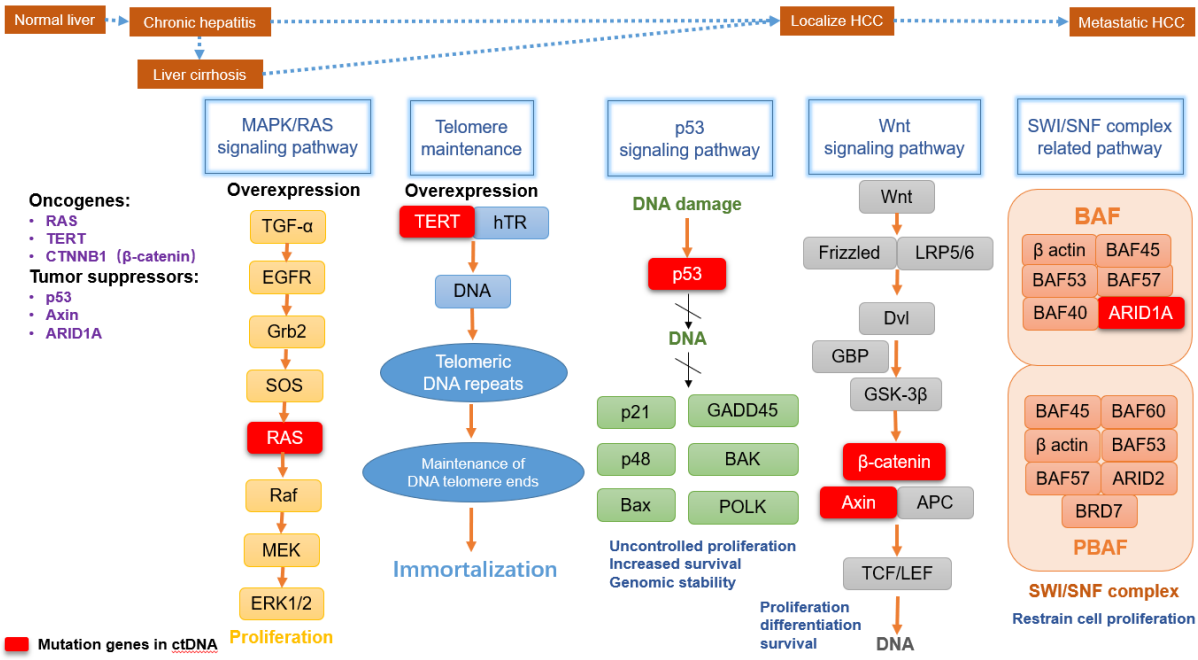


Figure 3. Major mutations in ctDNA and signaling pathways of HCC. *RAS/MAPK* pathway (yellow boxes), *TERT* mutation (blue boxes), p53 signaling pathway (green boxes), Wnt-β catenin pathway (gray boxes), and *SWI/SNF* complex-related pathway (light red boxes) are the core HCC signaling pathways. The general mutation genes from ctDNA show significant roles in pathways (red boxes) related to tumorigenesis and progression of hepatocellular carcinoma, involving common oncogenes and tumor suppressors [51].

The *RAS* gene is a group of oncogenes, coding proteins of the *RAS* family (*KRAS*, *NRAS* and *HRAS*), and function as small molecular-weight GTP-binding proteins for regulating cell growth [69]. The *RAS/MAPK* pathway is a general and conserved signaling pathway present in mammalian cells, and its kinases hold great potential as targets for the discovery of novel therapies [70]. Though *RAS* mutations are not commonly observed in HCC, the *MAPK/RAS* pathway is activated in nearly all advanced-stage HCCs and almost half of early-stage HCCs [71 72].

As the major subunit of the telomerase complex, the telomerase reverse transcriptase (*TERT*, or *hTERT*) maintains the telomere length, relating to the activity of telomerase [73]. Usually, the telomerase is kept inactive in mammalian somatic cells, but it will be activated in cancer or proliferating cells. In HCC patients, *TERT* has a high risk of being upregulated, especially in HCV infection [74]. The *TERT* promoter mutation is identified in dysplastic nodules in liver cirrhosis, suggesting that it is the key factor in the progression from chronic hepatitis to liver cancer [75].

The tumor protein *p53* (*TP53*) is a recognized tumor suppressor gene, acting as an adaptor in DNA repair proteins and facilitating the repair of DNA damage, arresting the cell cycle at checkpoints [76]. Mutation of *TP53* could inhibit this function and lead to "gain of function" effects in hepatoma cells, including disordered proliferation, cancer cell migration, and therapy resistance. Several etiological factors for HCC, such as chronic inflammation, infection of hepatitis viruses (HBV and HCV), and exposure in chemistry (aflatoxin B), can contribute to *TP53* mutations [77]. *TP53* mutation exhibits a high detecting frequency in ctDNA of HCC, but tissue-specific evidence of *TP53* is scarce due to its prevalence across multiple malignant tumors [51].

As a significant cascade comprising a series of factors for signal delivery, the *Wnt* signaling pathway impacts liver homeostasis, developmental regulation, and tumorigenesis [78]. The *Wnt* signaling pathway is one of the most commonly activated signaling pathways in HCC [79]. The β-catenin (*CTNNB1*) and *Axin* exert essential but opposite roles in the *Wnt* pathway: *CTNNB1* acts as an oncogene, while *Axin* is a tumor suppressor. The mutation of *CTNNB1* prevents the phosphorylation and degradation of the β-catenin protein and leads to anti-apoptosis, cell proliferation, and angiogenesis [80]. *Axin* is widely recognized as exerting a negative function in the *Wnt* signaling pathway. However, its expression can be reduced through tumor-specific promoter methylation or histone deacetylation, leading to the over-activation of signaling in tumors [81].

The *SWI/SNF* (Switch/Sucrose Non-Fermentable) is the complex for chromatin remodeling and gene transcription regulating, and its aberrations are found in probably 25% of cancers [82]. The *ARID1A* is the subunit gene of *SWI/SNF*, with the highest mutation rate among other components of the complex [83]. *ARID1A* plays a critical role in the regulation of gene expression, which is essential for driving either oncogenesis or tumor suppression [84]. During tumor initiation time, *ARID1A* shows the ability to promote HCC through CYP450-mediated oxidative stress, whereas it acts as an inhibitory factor for metastasis in established tumor time [85].

2.4. Research aim

So far, the number of mutations identified in the ctDNA of HCC patients is still limited, and a lack of studies evaluated the influence of ctDNA mutation measurements on established diagnostic techniques for HCC. The accuracy of tumor information obtained in ctDNA mutation and the clinical impact on the degree of accuracy have also been scarcely discussed. Additionally, current research on ctDNA and AFP levels merely compares the two methods or

combines ctDNA concentrations and AFP. No studies have specifically explored the potential impact of combining AFP with ctDNA mutations for HCC diagnosis. Therefore a more in-depth analysis needs to evaluate the clinical implication of these mutations.

In our research, a multi-mutation NGS panel with 100 common genes of HCC was designed for the following research aims:

a. Detecting more mutation targets to discover the genetic profiling of ctDNA in HCC patients.
b. Exploring the effect of ctDNA mutations in HCC diagnosis by combining ctDNA and the established HCC diagnostic biomarker (AFP).
c. Assessing the concordance between plasma gene profiling and tumor tissue and the corresponding clinical factors.
d. Evaluating the implication of genetic profiling in ctDNA for the progression and survival time of HCC.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Human samples

Blood samples

The human blood samples were constituted of 30 patients with a definite HCC diagnosis from 2016–2019 and 10 patients with benign liver disease from the Department of General, visceral, tumor, and transplant surgery in the University Hospital of Cologne. The diagnosis of HCC and liver benign disease were mainly established by histopathologic evidence. The study was undertaken by the Declaration of Helsinki (1975), and the University of Cologne ethics committee has also approved it (Biological Material Collection for Optimisation ID: 13-091).

Paraffin samples

Thirty paraffin wax samples of our HCC cohort were provided by the Department of Pathology at the University Hospital of Cologne. All the patients recruited were informed of consent before participating.

3.1.2. Materials of ctDNA/cfDNA extraction

Name	Company
QIAamp MinElute ccfDNA Midi Kit (50)	Cat No./ID: 55284, Qiagen, Germany
QIAamp DNA Blood Mini Kit (50)	Cat No./ID: 51104, Qiagen, Germany
QIAamp DNA FFPE Tissue Kit (50)	Cat No./ID: 56404, Qiagen, Germany

3.1.3. Materials of ctDNA/cfDNA quantification

Name	Company
Cell-free DNA ScreenTape	Cat No./ID: 5067-5630, Agilent Technologies, USA
Cell-free DNA Reagents	Cat No./ID: 5067-5630, Agilent Technologies, USA

3.1.4. Materials for library preparation and NGS

Name	Company
SureSelect XT HS and XT Low Input Library Preparation Kit for ILM (Pre PCR)	Cat # 5500-0140, Agilent Technologies, USA
SureSelect XT Low Input Index Primers 1-96 for ILM (Pre PCR)	Cat # 5190-6444, Agilent Technologies, USA
SureSelect Target Enrichment Kit, ILM Hyb Module Box 1 (Post PCR)	Cat # 5190-9687, Agilent Technologies, USA
SureSelect XT HS and XT Low Input Target Enrichment Kit ILM Hyb Module Box 2 (Post PCR)	Cat # 5190-9686, Agilent Technologies, USA
SureSelect Custom Tier 2 0.5Mb-2.9Mb Probe (up to 120k oligos) sufficient for post-capture processing of 96 samples	Cat # 5191-6906, Agilent Technologies, USA
SureSelect Custom Tier 2 0.5Mb-2.9Mb Probe (up to 120k oligos) sufficient for post-capture processing of 16 samples.	Cat # 5191-6905, Agilent Technologies, USA
SureSelect XT Low Input Reagent Kit with indexes 1-96, 96rxn kit	Cat # G9703A, Agilent Technologies, USA
SureSelect XT HS Reagent Kit with indexes 1-16, 16rxn kit	Cat # G9702A, Agilent Technologies, USA

3.1.5. Laboratory equipment

Name	Company
Centrifuge	Megafuge 1.0R, Heraeus, Germany
Microcentrifuge	Thermo Scientific™, Germany

Automatic pipettes	Eppendorf, Germany
Vortex	Lab dancer, VWR, Germany
Fridge 4°C	Liebherr, Germany
Freezer -20°C	Bosch, Germany
Freezer -80°C	Sanyo, Japan
Freezer -150°C	Sanyo, Japan
Microtome	33699 TechnoMed GmbH，Germany
Thermocycler	Tpersonal, Biometra, Germany
Thermomixer	ThermoMixer C, Eppendorf, Germany
Real time PCR	QuantStudio 7 flex, Applied Biosystems, USA

3.1.6. Consumable materials

Name	**Company**
Tube (15 mL)	Sarstedt, Germany
EDTA tube	Sarstedt, Germany
SafeSeal tube	Sarstedt, Germany
Centrifuge tube (15 and 50ml)	Sarstedt, Germany
Cryotubes (1.8 mL)	Sarstedt, Germany
Serological pipettes (5, 10 and 25mL)	Sarstedt, Germany
Pipette tips (10, 200 and 1000µL)	Sarstedt, Germany

3.1.7. Software

Name	Company
Microsoft Office	Microsoft Corporation, USA
SPSS Statistics 26	IBM, USA
Graphpad Prism 8	GraphPad Software, Inc., USA
Endnote X9	Thomson Reuter, USA
Gene set enrichment analysis (Version: 4.1.0)	Broad Institute, USA
Bcl2fastq2 Conversion Software v2.20	Illumina, USA
Agilent Genomics NextGen Toolkit (AGeNT)	Agilent, USA
Burrows Wheeler Aligner (BWA, v0.7.17)	Software package
Samtools (v1.14)	Genome Research Limited (reg no. 2742969), England
ComplexHeatmap	R-package, USA

3.2. Methods

3.2.1. Patients

A total of 30 patients with HCC and ten patients with benign liver disease were enrolled in the study. Both the experimental and control groups applied the following criteria:

a. Age> 18 years with HCC or benign liver disease
b. Without prior liver surgery or systemic therapy for HCC
c. Without any previous history of other tumors

Before the surgery, 30 ml of perivenous blood was collected directly from the patient in EDTA tubes and transported to the reception laboratory within 4 hours [86]. The plasma and interphase from the cellular component of the whole blood were isolated by centrifuging at 4000 rpm

(3488g) for 10 minutes at room temperature. Subsequently, the plasma and buffy coat samples were aliquoted into RNA-free Eppendorf microtubes and stored at -80° C for further use.

At the time of blood sample collection, clinical and pathological data of all patients were recorded, including gender, age, HCC stage (TNM and BCLC stage), Child-Turcotte-Pugh (CTP) score, AFP level, concomitant liver diseases such as alcohol liver disease, hepatitis, and non-alcoholic fatty liver disease, presence of cirrhosis, portal venous thrombosis, vascular invasion, and metastases. The follow-up included monitoring progression-free survival (PFS) and overall survival (OS).

3.2.2. CfDNA extraction

To extract cfDNA, 7 ml of plasma was used with the QIAamp MinElute ccfDNA Midi Kit (50) (Qiagen, Germany), following the manufacturer's instructions. In brief, the cell-free nucleic acids in plasma could be pre-concentrated in the magnetic beads and eluted from columns in a spin procedure. The extracted cfDNA was stored at -20°C for future use.

3.2.3. CfDNA quantification and quality appraisal

We checked the quality and quantification of cfDNA, using cell-free DNA screen tape assay (Agilent Technologies). In the cfDNA screen tape, the extracted cfDNA samples were separated by automated electrophoresis and displayed a prominent peak at nearly 150-200bp. Furthermore, this table could show the cfDNA concentration and the quality metric (% cfDNA), which means the percentage of cfDNA subcomponents.

3.2.4. Germline-DNA extraction

The buffy coat samples were used to extract germline DNA (gDNA) by the QIAamp DNA Blood Mini Kit (50) (Qiagen, Germany). In short, blood DNA from 600µl buffy coat was purified with the procedures of fast spin-column, vacuum, and centrifugation. Finally, the gDNA samples were stored in a -20°C refrigerator.

3.2.5. Tumor-DNA extraction

The HCC tumor tissue samples were fixed with formalin and embedded in paraffin with the help of the Institute of Pathology of the University Hospital of Cologne. From archival blocks, matched formalin-fixed paraffin-embedded tissue (FFPE) tumor samples were obtained, and tumor DNA extraction was completed with the QIAamp DNA FFPE Tissue Kit (50) (Qiagen,

Germany), according to the manufacturer's instructions. The extracted tDNA samples were also stored in a -20°C fridge.

3.2.6. HCC sequencing panel design

To optimize mutation detection rates while minimizing panel size, we designed a next-generation sequencing panel for HCC. We reviewed the genomic profiles of HCC using cBioPortal, including the TCGA database and additional HCC clinical data [87] [88,89] [90]. The final design (HCC_Panel_v1.1) targeted the 100 most frequently mutated genes in HCC, covering all exonic domains, with a total size of 692 kbp (Table 2).

Table 2. NGS panel of HCC mutations

HCC mutation genes	Frequency	Oncogene or tumor suppressor gene
CTNNB1	29.70%	Oncogene
TP53	28.90%	Tumor suppressor gene
TERT	19.50%	Oncogene
ALB	10.60%	Tumor suppressor gene
ARID1A	9.50%	Tumor suppressor gene
PCLO	8.60%	Unknown
AXIN1	7.30%	Tumor suppressor gene
LRP1B	7.30%	Tumor suppressor gene
KMT2D	5.70%	Tumor suppressor gene
ARID2	5.70%	Tumor suppressor gene
PREX2	4.80%	Oncogene
RB1	4.70%	Tumor suppressor gene
BAP1	4.60%	Tumor suppressor gene
NFE2L2	4.50%	Oncogene
TSC2	4.30%	Tumor suppressor gene
KMT2C	4.20%	Tumor suppressor gene
KEAP1	4.10%	Tumor suppressor gene
FAT4	4.10%	Tumor suppressor gene
SETD2	4.00%	Tumor suppressor gene
KMT2B	3.70%	Both
ATM	3.50%	Tumor suppressor gene

PTPRB	3.40%	Tumor suppressor gene
KMT2A	3.30%	Both
NF1	3.20%	Tumor suppressor gene
PRKDC	3.20%	Oncogene
APC	3.00%	Tumor suppressor gene
JAK1	2.90%	Both
PIK3CA	2.90%	Oncogene
ACVR2A	2.80%	Tumor suppressor gene
ATR	2.80%	Tumor suppressor gene
NCOR2	2.80%	Unknown
SF3B1	2.80%	Oncogene
CDKN2A	2.70%	Tumor suppressor gene
NOTCH3	2.70%	Oncogene
PTPN13	2.70%	Tumor suppressor gene
POLQ	2.70%	Both
PTPRT	2.70%	Tumor suppressor gene
IGF1R	2.60%	Oncogene
SMARCA4	2.60%	Tumor suppressor gene
ARID1B	2.60%	Both
IL6ST	2.60%	Tumor suppressor gene
RELN	2.60%	Tumor suppressor gene
SPEN	2.60%	Tumor suppressor gene
KIT	2.60%	Oncogene
ZFHX3	2.50%	Tumor suppressor gene
BRCA2	2.50%	Tumor suppressor gene
EP300	2.50%	Both
EPHA3	2.50%	Tumor suppressor gene
PDE4DIP	2.50%	Both
LRRK2	2.50%	Tumor suppressor gene
NOTCH4	2.40%	Oncogene
HNF1A	2.40%	Both
RNF213	2.40%	Tumor suppressor gene
CREBBP	2.30%	Tumor suppressor gene
MKI67	2.30%	Unknown
ROBO1	2.30%	Both
ZNF521	2.30%	Tumor suppressor gene
FLT4	2.20%	Oncogene

NOTCH1	2.20%	Oncogene
PTEN	2.20%	Tumor suppressor gene
RANBP2	2.20%	Tumor suppressor gene
TSC1	2.20%	Tumor suppressor gene
MGA	2.20%	Oncogene
PBRM1	2.20%	Tumor suppressor gene
PAK5	2.20%	Oncogene
CARD11	2.20%	Oncogene
ERBB4	2.10%	Both
MECOM	2.10%	Oncogene
MTOR	2.10%	Oncogene
HGF	2.10%	Both
MGAM	2.10%	Unknown
CDKN1A	2.00%	Both
NTRK3	2.00%	Both
NUP214	2.00%	Tumor suppressor gene
MYO18A	2.00%	Oncogene
NCOR1	1.90%	Oncogene
TAF1	1.80%	Unknown
NTRK2	1.80%	Both
ASXL1	1.60%	Tumor suppressor gene
EGFR	1.60%	Oncogene
TRRAP	1.50%	Unknown
MSH6	1.50%	Both
PDGFRA	1.50%	Oncogene
MED12	1.40%	Oncogene
BRCA1	1.30%	Tumor suppressor gene
CHD2	1.20%	Tumor suppressor gene
BCORL1	1.20%	Oncogene
WRN	1.10%	Tumor suppressor gene
BLM	1.10%	Tumor suppressor gene
PIK3CG	1.00%	Oncogene
FGFR2	1.00%	Oncogene
FGFR4	1.00%	Oncogene
IDH2	1.00%	Oncogene
RAD50	0.90%	Oncogene
KRAS	0.80%	Oncogene

NRAS	0.80%	Oncogene
ERCC5	0.80%	Unknown
ERBB2	0.70%	Oncogene
MET	0.60%	Oncogene
MYC	0.40%	Oncogene

3.2.7. Library preparation and NGS

To accomplish library preparation for targeted gene panel sequencing, we utilized a minimum of 25ng DNA for plasma cfDNA and 200ng DNA for FFPE tDNA per sample. We performed sequencing on plasma cfDNA, germline DNA, and tumor DNA using a targeted gene panel approach.

DNA quantification was performed using the TapeStation 2200 System (Agilent). Additionally, size distribution was assessed for plasma cfDNA, requiring a minimum of 25% of total input DNA to be in the size window of interest for cfDNA (50-700bp). Library preparation was performed using the Agilent Sure SelectXT Low Input protocol, including enzymatic fragmentation (for gDNA and tDNA samples), end-repair, adapter ligation, index PCR, enrichment with the NGS panel we designed (Genepanel Design ID: 3311801 (0,5Mb-2,9Mb)), and post-enrichment PCR (12 cycles). Different circles were used for different types of DNA fragments for index PCR: 10 cycles for cfDNA, eight cycles for gDNA samples, and 11 cycles for tDNA samples. Subsequently, libraries were quantified (Qubit, Tape Station), pooled equimolarly and sequenced on a NovaSeq 6000 device with paired-end, 2x100bp sequencing protocol. We targeted 5 Gb data output for plasma cfDNA and FFPE tDNA, and 1 Gb output for gDNA.

3.2.8. Basic data processing, variant calling and filtering

We performed basic data processing using AGeNT Trimmer (Agilent), Bcl2fastq2 (v2.20.0.422), Samtools (v1.14), and Burrows-Wheeler Aligner (BWA, v0.7.17), as previously reported [91]. We used comparative error suppression (CES) to improve sensitivity and specificity to call somatic single-base substitutions, as previously reported [92]. And we visualized genetic aberrations and clinical annotations using the R-package 'ComplexHeatmap' [92].

3.2.9. Gene Ontology (GO) enrichment

All the ctDNA mutation targets were enriched in DAVID Bioinformatics Resource (https://david.ncifcrf.gov/), with the p-value ≤ 0.05. Then the results were performed by the website (www.bioinformatics.com.cn) and displayed with bubble dot diagrams.

3.2.10. Statistical analyses

We conducted statistical analyses using Excel spreadsheets, SPSS Statistics 26.0, Origin 2021 and GraphPad Prism 8. Clinical variables were reported as median (interquartile range [IQR]) or mean ± standard deviation. Progression-free survival and overall survival were presented with Kaplan-Meier plots. To assess correlations in cfDNA concentration, genetic mutation in ctDNA, and clinical variables, we used appropriate Fisher's exact test, nonparametric tests, Wilcoxon rank-sum test, or receiver operating characteristic (ROC) curves. All statistical analyses in our study were performed with a significance level of 5%.

4. RESULTS

4.1. Clinical characteristics of enrolled HCC patients

Firstly, the primary liver cancer samples from BioMASOTA were checked, which had been established several years before. The study design workflow, including details, is presented in Fig 4. We enrolled 30 HCC patients and 10 patients with liver benign diseases. Table 3 lists the clinical information for HCC patients with long-term follow-ups, reaching a mean age of 69. There are 12 females and 18 males in our HCC cohort, revealing that 40% (12/30) were in the early stage (BCLC stage 0/A), 40% were in the intermediate stage (BCLC stage B), and 20% (6/30) were in the advanced stage. Thirteen cases had viral hepatitis, with one patient having hepatitis B and twelve having hepatitis C. The HCC cohort consisted of 6 patients with alcoholic liver disease and 12 with non-alcoholic fatty liver disease. Cirrhosis was present in 56.7% (17/30) of the cases, while vascular invasion was observed in 53.3% of the HCC patients (16/30), affecting both macrovascular and microvascular. The metastasis rate among the HCC cohort was 13.3% (4/30). The median diameter of the largest tumor nodule was 41.5mm (IQR 22.8-60.8), and portal vein thrombosis was observed in two patients, accounting for 23.0% of the cohort. Our cohort showed a significant decline in both overall survival (OS) and progression-free survival (PFS) as it progressed from TNM stage 1 and stage 2-4 (Fig. 5A). Moreover, the prognosis for HCC patients in the advanced stages was notably poorer in comparison to those in the early and intermediate stages (Fig. 5B).

The clinical information of patients in the control group is listed in Table 4. The control cohort was constituted of 5 females and 5 males, with a mean age of 60. All patients in the control group were free of viral hepatitis and had good liver function before surgery (CTP class A). 30% (3/10) of cases were associated with liver cirrhosis and one patient had non-alcoholic fatty liver disease. All patients were alive and not lost during the follow-up time.

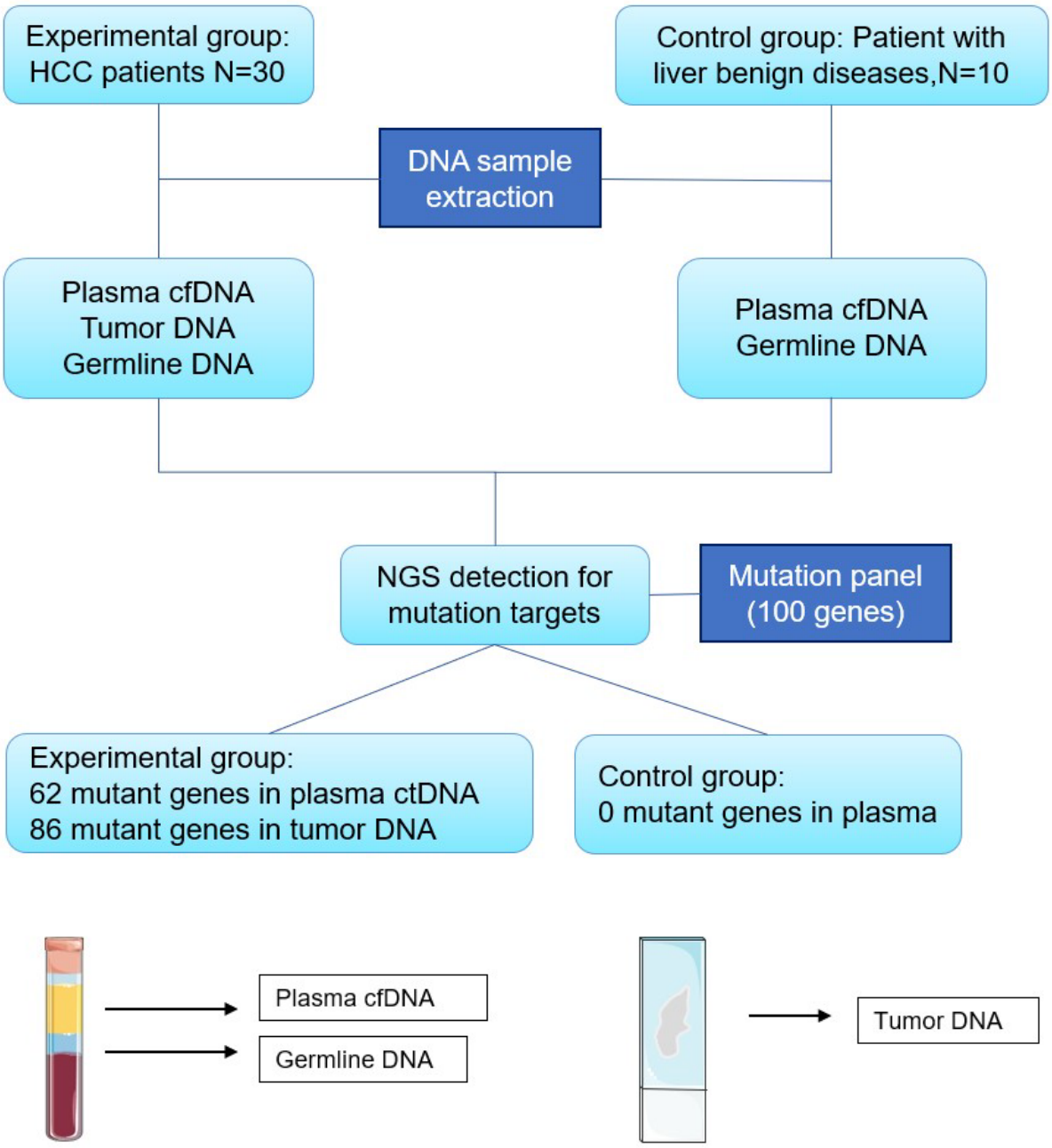


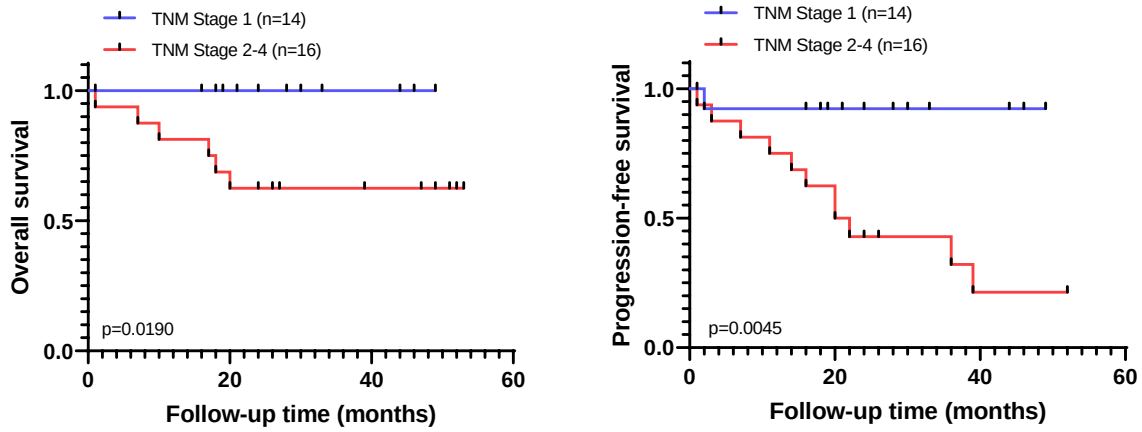
Figure 4. Workflow chart of data generation and analysis: Totally, 40 patients with liver diseases from the BIOMASOTA underwent plasma cfDNA and gDNA extraction, but the tDNA was taken from FFPE (Formalin-fixed paraffin-embedded tissue) samples only in the HCC cohort. In our cohort, 30 HCC patients were the experimental group and 10 patients with benign liver lesions were the control group. Following NGS testing and analysis of all DNA samples, several mutations were identified in the HCC group, whereas no mutations were detected in the control group.

Table 3. Clinical variables among HCC cohorts (N=30).

Clinical variable	Number (%)
Gender	
Male	18 (60%)
Female	12 (40%)
Mean age (years ± SD)	69.0 ± 9.0
Associated with HBV	1 (3.3%)
Associated with HCV	12 (40%)
Associated with cirrhosis	17 (56.7%)
Alcoholic liver disease	6 (20%)
Non-alcoholic fatty liver disease	12 (40%)
Largest tumor diameter (mm)	216
Median largest tumor diameter (mm)	41.5 (IQR 22.8-60.8)
Median AFP pre-operative (ng/ml, N=22)	9.5 (IQR 4.8-47.5)
Macrovascular invasion	3 (10%)
Microvascular invasion	13 (43.3%)
Portal vein thrombosis	2 (6.7%)
Presence of metastasis	4 (13.3%)
CTP classification	
A	27 (90%)
B	3 (10%)
C	0
BCLC classification	
0	5 (16.7%)
A	7 (26.7%)
B	12 (40%)
C	6 (20%)

D	0
pTNM classification	
1	14 (46.7%)
2	6 (20%)
3	5 (16.7%)
4	5 (16.7%)

A.



B.

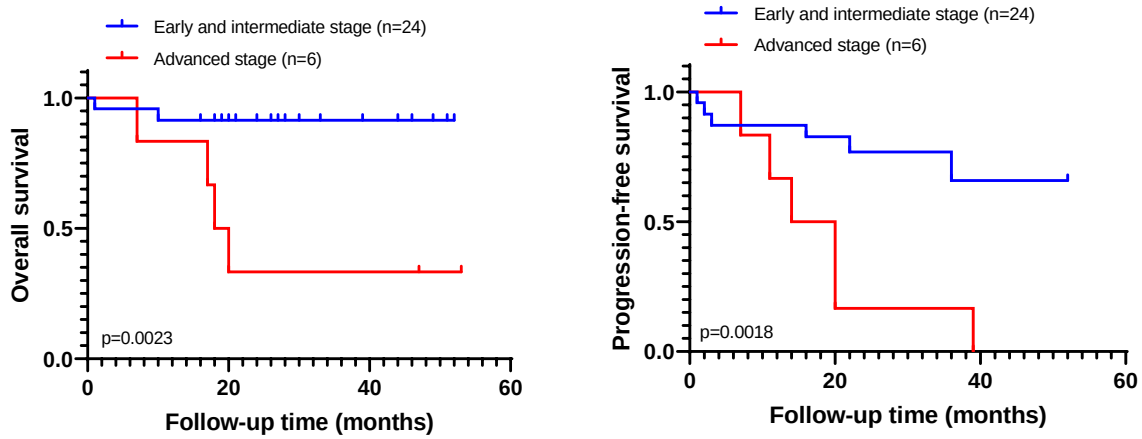


Figure 5. A. Kaplan–Meier analysis for overall survival (p=0.019) and progression-free survival (p=0.0045) between TNM stage 1 and 2-4. B. Kaplan–Meier analysis for overall survival (p=0.0023) and progression-free survival (p=0.0018) in different BCLC stages.

Table 4. Clinical variables among the control group (N=10).

Clinical variable	Number (%)
Gender	
Male	5 (50%)
Female	5 (50%)
Mean age (years ± SD)	60.0 ± 15.5
Etiology of chronic liver disease	
HBV	0
HCV	0
Cirrhosis	3 (30%)
Alcoholic liver disease	1 (10%)
Non-alcoholic fatty liver disease	4 (40%)
Hepatic echinococcosis	3 (10%)
Hepatic adenomas	1 (10%)
hepatic hemangiomas	1 (10%)
Focal nodular hyperplasia	1 (10%)
Median AFP pre-operative (ng/ml, N=5)	9.5 (IQR 2.4-7.3)
CTP classification	
A	10 (100%)
B	0
C	0

4.2. The quality test for cfDNA of HCC

We selected plasma samples from HCC patients at three different time periods during 2016-2018 and subjected them to circulating cfDNA testing. All samples exhibited a high concentration of cfDNA with good quality (Table 5). The high cfDNA concentration suggests the presence of plasma-free DNA. The length of the cfDNA we detected ranged from 100 to 200 base pairs (Fig. 4), consistent with published data on cfDNA.

Table 5. CfDNA concentration in the tested sample

HCC sample	%cfDNA	cfDNA concentration (ng/ml)
A	95	25800
B	78	1940
C	92	2370

Sample A.

Sample B.

Sample C.

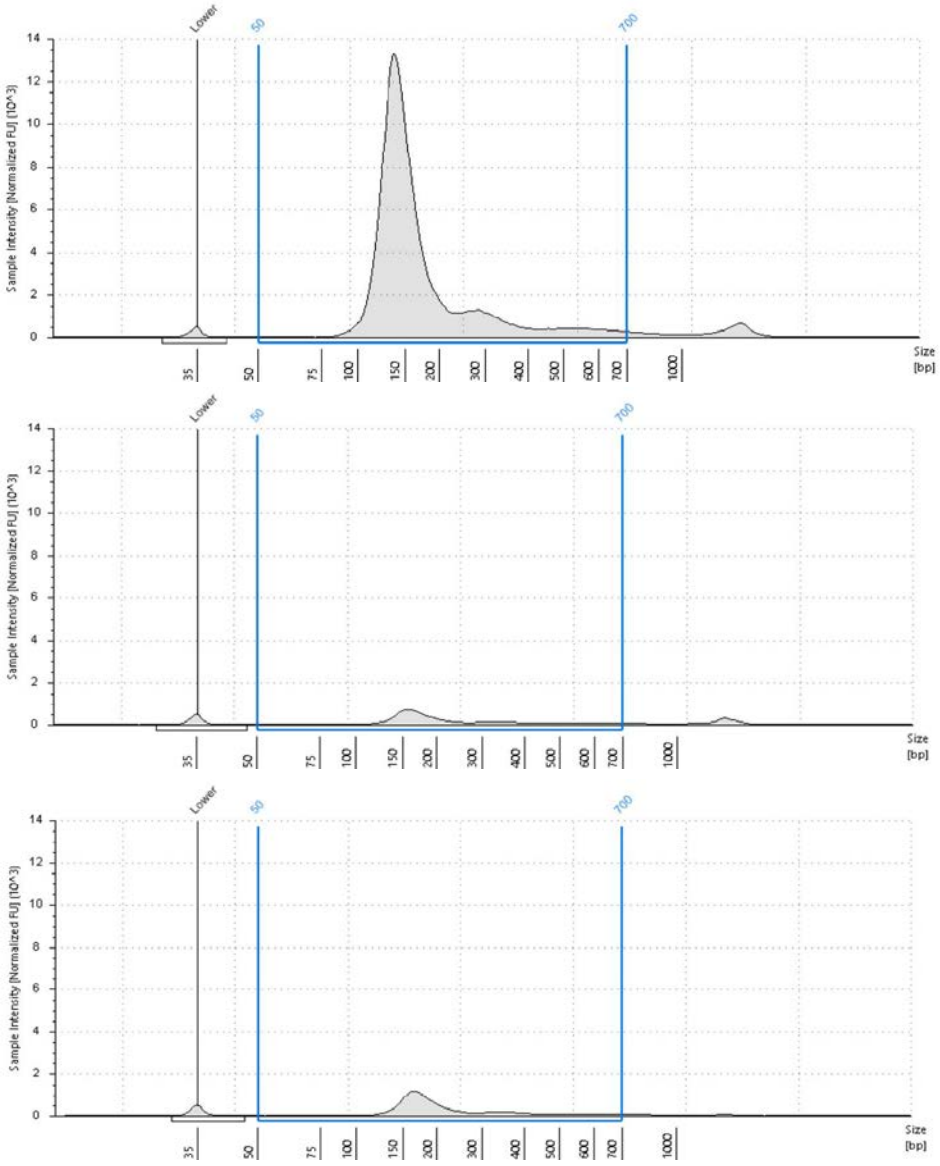


Figure 6. All three cfDNA from plasma show similar size distributions, peaking at 150 to 200 bp in length, which was consistent with the normal size of cfDNA in literature.

4.3. Genetic mutation of ctDNA showed an effect in distinguishing ctDNA and cfDNA

Cell-free DNA could be detected in the plasma samples from both the experimental and control groups (n=40). The median concentration of cfDNA in the HCC group was 10.4ng/ml (IQR 3.4-17.8ng/ml), indicating a slightly higher trend compared to the control group (6.3ng/ml, IQR 3.0-10.5ng/ml). However, the cfDNA concentration between the two groups was not significantly different (Fig. 8A). Subsequently, we focused on the genetic mutations, excluding silent mutations and retaining only functional ones. The genetic mutations detected in ctDNA encompass various types: exonic mutations, intronic mutations, intergenic mutations, splice site mutations, and mutations in the 3'-untranslated region and 5'-untranslated region. Among these mutations, the exonic part constituted the largest proportion, accounting for 77.5% of the total (Fig. 8B). To establish the baseline, several criteria were applied to the filter parameters of mutation genes in both ctDNA and tDNA.

These criteria include: (1) mutations that must be located in exonic regions. (2) intronic and intergenic mutations were excluded. (3) mutations without a clear function were excluded. (4) synonymous and unknown mutations of the exon were excluded.

The final screening revealed at least one mutation gene could be detected in 20 out of 30 individuals' plasma in the HCC group, accounting for 66.7% (Table 5). However, no mutation genes were detected in the control group. Thus, a significant difference in the proportion of patients with mutations was observed between the experimental and control groups (Fig. 8C). The identified mutations in HCCs consisted of 49 eligible mutant genes, encompassing 91 exons in ctDNA, and 72 eligible mutant genes, including 171 exons in tDNA. The mean mutated allele frequency (mAF) of ctDNA was 5.2%, while the ctDNA concentration was calculated at 2.2 log10 [haploid genome equivalents/ml] of plasma. Genetic mutations proved to be more effective in distinguishing ctDNA from normal cfDNA than cfDNA concentration in plasma.

The mutations detected in ctDNA were fewer than those in tDNA across various parameters, including the number of patients with the mutation, types of mutant genes, mutant exons, and the mean number of exonic mutations per patient (Table 6). Additionally, the median mutant allelic frequency of ctDNA was 2 % (IQR 1 %-10 %; n = 20), which was lower than the corresponding value of 0.16 detected in tDNA (IQR 0.07-0.27; n = 30) (Fig. 8D).

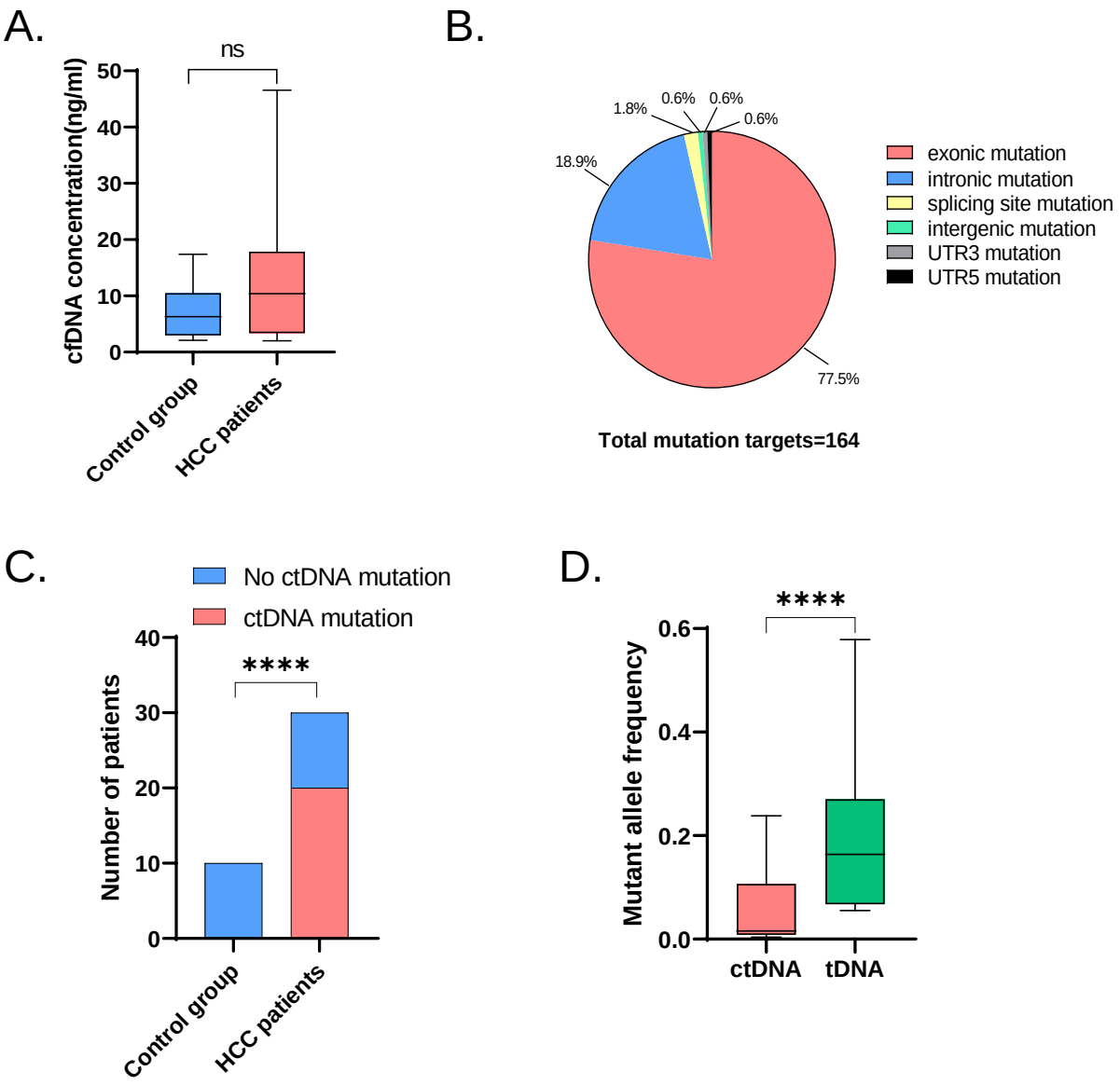


Figure 8. CfDNA concentration and genetic mutation of ctDNA: A. Comparison of cfDNA concentrations in HCC and control group. There was no significant difference. **B.** The 6 different kinds of mutations in ctDNA. **C.** Comparison of mutation ratio in HCC and control group. Fisher's exact test ****p<0.0001. D. Comparison of mutations in plasma ctDNA and tumor FFPE tDNA. Mann Whitney test: **** p<0.0001.

Table 6. Characteristics of mutations in experimental and control groups.

Characteristic	**Experimental group**		**Control group**
	Plasma ctDNA	**Tumor DNA**	**Plasma cfDNA**
Samples volume	30	30	10

Patient with mutation	20/30	30/30	0/10
Genes in panel	49/100	72/100	0/100
Mutant exons	91	170	0
Exonic mutation per sample - Mean	4.6	7.2	0

4.4. Landscapes of mutations of ctDNA and associated signaling pathways

We proceeded to explore the characteristics of ctDNA mutations and conduct a comparison with tDNA mutations, the details of mutation frequencies in ctDNA and tDNA are exhibited in Table 7. Among the 49 ctDNA genes with mutant exons, fifteen were initially detected specifically in free nucleic acid of HCC: *NCOR2, HGF, MECOM, ROBO1, MKI67, PEPN13, RANBP2, RELN, ALB, FAT4, KMT2B, MGAM, PAK5, PTPRB, ZFHX3*. *NCOR2* was the most frequent of these, with a total of 4 patients and a frequency of 13.33 percent.

Table 7. Frequency of mutations in ctDNA and tDNA.

Genes in panel	Patients with plasma ctDNA	ctDNA freq (%)	Patients with Tumor tDNA	tDNA freq (%)
CTNNB1	4	13.33	11	36.67
NCOR2	4	13.33	4	13.33
TP53	3	10.00	4	13.33
ROBO1	3	10.00	2	6.67
PDE4DIP	3	10.00	1	3.33
KMT2C	3	10.00	0	0.00
KMT2D	2	6.67	4	13.33
RNF213	2	6.67	3	10.00
RANBP2	2	6.67	3	10.00
ACVR2A	2	6.67	3	10.00
HGF	2	6.67	2	6.67
ATM	2	6.67	2	6.67
BRCA2	2	6.67	2	6.67

ERBB4	2	6.67	2	6.67
MKI67	2	6.67	1	3.33
IGF1R	2	6.67	0	0.00
MECOM	2	6.67	0	0.00
KMT2B	1	3.33	3	10.00
NOTCH1	1	3.33	5	16.67
PCLO	1	3.33	4	13.33
ALB	1	3.33	3	10.00
CARD11	1	3.33	3	10.00
MGAM	1	3.33	2	6.67
PIK3CA	1	3.33	3	10.00
FAT4	1	3.33	2	6.67
KMT2A	1	3.33	2	6.67
NOTCH4	1	3.33	2	6.67
PTPRB	1	3.33	2	6.67
RAD50	1	3.33	2	6.67
TSC1	1	3.33	2	6.67
ZFHX3	1	3.33	2	6.67
ARID2	1	3.33	1	3.33
AXIN1	1	3.33	1	3.33
BAP1	1	3.33	1	3.33
EGFR	1	3.33	1	3.33
ERBB2	1	3.33	1	3.33
ERCC5	1	3.33	1	3.33
FLT4	1	3.33	1	3.33
MSH6	1	3.33	1	3.33
PAK5	1	3.33	1	3.33
PTPN13	1	3.33	1	3.33
RELN	1	3.33	1	3.33
TAF1	1	3.33	1	3.33
TERT	1	3.33	1	3.33
KIT	1	3.33	0	0.00
NF1	1	3.33	0	0.00
PIK3CG	1	3.33	0	0.00
POLQ	1	3.33	0	0.00
LRP1B	1	3.33	2	6.67
SPEN	0	0.00	1	3.33

PRKDC	0	0.00	0	0.00
CDKN2A	0	0.00	4	13.33
PREX2	0	0.00	3	10.00
PTPRT	0	0.00	3	10.00
TRRAP	0	0.00	3	10.00
ARID1A	0	0.00	2	6.67
ARID1B	0	0.00	2	6.67
MED12	0	0.00	2	6.67
MTOR	0	0.00	2	6.67
NOTCH3	0	0.00	3	10.00
SMARCA4	0	0.00	2	6.67
TSC2	0	0.00	2	6.67
ATR	0	0.00	1	3.33
BCORL1	0	0.00	1	3.33
CHD2	0	0.00	1	3.33
CREBBP	0	0.00	1	3.33
EP300	0	0.00	0	0.00
HNF1A	0	0.00	1	3.33
JAK1	0	0.00	1	3.33
KEAP1	0	0.00	1	3.33
MET	0	0.00	1	3.33
MGA	0	0.00	1	3.33
MYO18A	0	0.00	1	3.33
NFE2L2	0	0.00	1	3.33
NRAS	0	0.00	1	3.33
NTRK2	0	0.00	1	3.33
NUP214	0	0.00	1	3.33
PBRM1	0	0.00	1	3.33
RB1	0	0.00	1	3.33
SETD2	0	0.00	1	3.33
SF3B1	0	0.00	1	3.33
APC	0	0.00	0	0.00
ASXL1	0	0.00	0	0.00
BLM	0	0.00	0	0.00
BRCA1	0	0.00	0	0.00
CDKN1A	0	0.00	0	0.00
EPHA3	0	0.00	0	0.00

FGFR2	0	0.00	0	0.00
FGFR4	0	0.00	0	0.00
IDH2	0	0.00	0	0.00
IL6ST	0	0.00	0	0.00
KRAS	0	0.00	0	0.00
LRRK2	0	0.00	0	0.00
MYC	0	0.00	0	0.00
NCOR1	0	0.00	0	0.00
NTRK3	0	0.00	0	0.00
PDGFRA	0	0.00	0	0.00
PTEN	0	0.00	0	0.00
WRN	0	0.00	0	0.00
ZNF521	0	0.00	0	0.00

4.4.1. Genetic landscape of ctDNA

A genetic heat map was generated to illustrate 32 mutations from ctDNA, including all mutations with a frequency exceeding 10% and a few mutations that we deemed significant despite the frequency of 6.7%. The number of patients harboring mutated exons displayed an upward trend in TNM stage II-IV patients compared to those in TNM stage I, but this difference was not significant (Fig. 9). Moreover, among HCC patients with vascular invasion (macro- and micro-), HBV or HCV infection, cirrhosis, metastasis, and concomitant alcohol liver disease or non-alcoholic fatty liver, a higher probability of detectable mutations in plasma ctDNA was observed, despite these alterations showing only a nonsignificant trend (Fig. 9). The identified mutant exon types included nonsynonymous mutations, stop gain mutations, frameshift deletion mutations, and non-frameshift deletion mutations. The genetic mutation landscape revealed prominent genes such as *CTNNB1* (13.3%), *NCOR2* (13.3%), *TP53* (10%), *PDE4DIP* (10%), *KMT2C* (10%), *ROBO1* (10%), *RANBP2* (6.7%), *ACVR2A* (6.7%), *ATM* (6.7%), and *HGF* (6.7%) (Fig. 9). In the heatmap, it was also the first that *NCOR2*, *ROBO1*, *RANBP2*, *HGF* were identified in ctDNA of liver cancer.

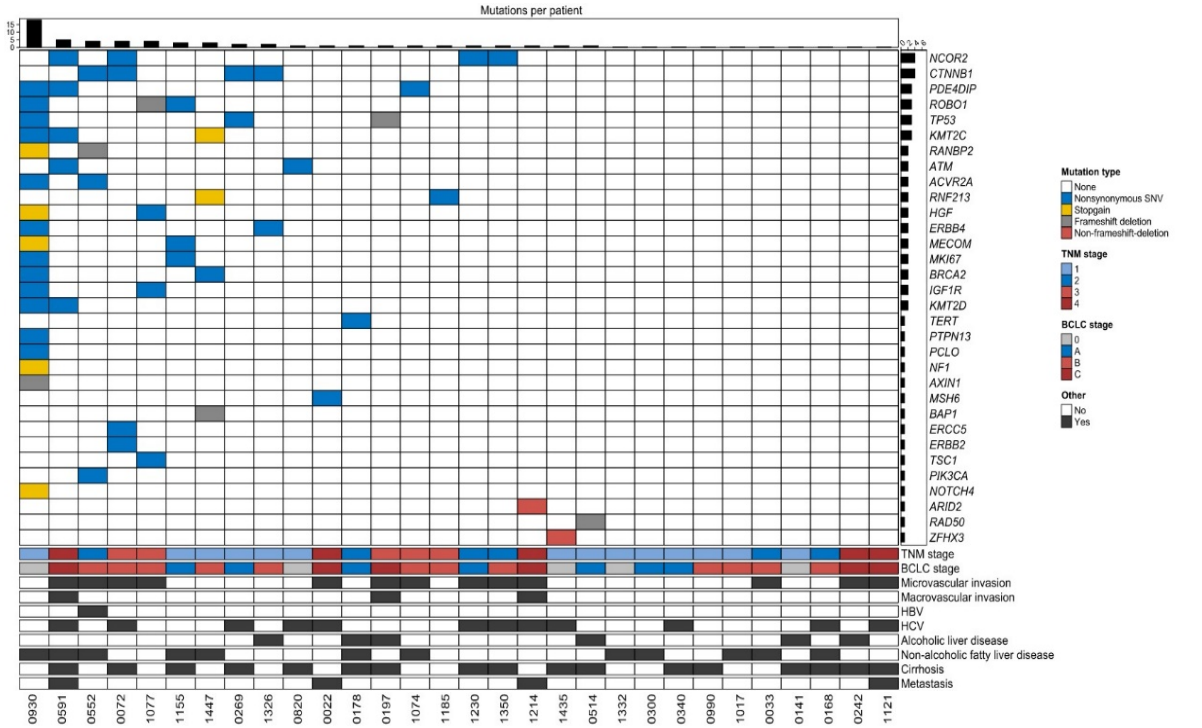


Figure 9. Landscapes of mutations in ctDNA: CtDNA mutation profiling of 30 HCC patients. The number of tumor mutations in each individual is in the top part, and the details of mutations are in the middle. The bottom panel exhibits the TNM stage, BCLC stage, microvascular invasion, macrovascular invasion, HBV, HCV, alcohol liver disease, non-alcohol fatty liver disease, cirrhosis, and metastasis for our HCC cohorts.

4.4.2. Mutant genes in ctDNA and tDNA

Then, we proceeded with a comparative analysis of mutations identified in plasma and their matched tumor tissues. Furthermore, we compared these mutations with the gene frequencies available in publicly accessible databases from cBioportal (N=630), which encompasses the TCGA database and other clinical HCC data (Fig. 10). We focused on 17 genes, as depicted in Figure 9, which exhibited frequencies exceeding 6.7%. Notably, more than 50% of the mutations did not demonstrate significant differences between circulating tumor DNA (ctDNA) and tumor DNA (tDNA), including *NCOR2*, *RANBP2*, *ROBO1*, *ACVR2A*, *ATM*, *HGF*, *RNF213*, *ERBB4*, and *BRCA2*. Conversely, certain genes exhibited markedly distinct mutation frequencies between ctDNA and tDNA, such as *CTNNB1*. Additionally, we observed that the mutation frequencies in ctDNA and tDNA samples from our HCC patient cohort were generally higher than those reported in existing databases, except *CTNNB1* and *TP53*. This discrepancy may be attributed to different sequencing depths.

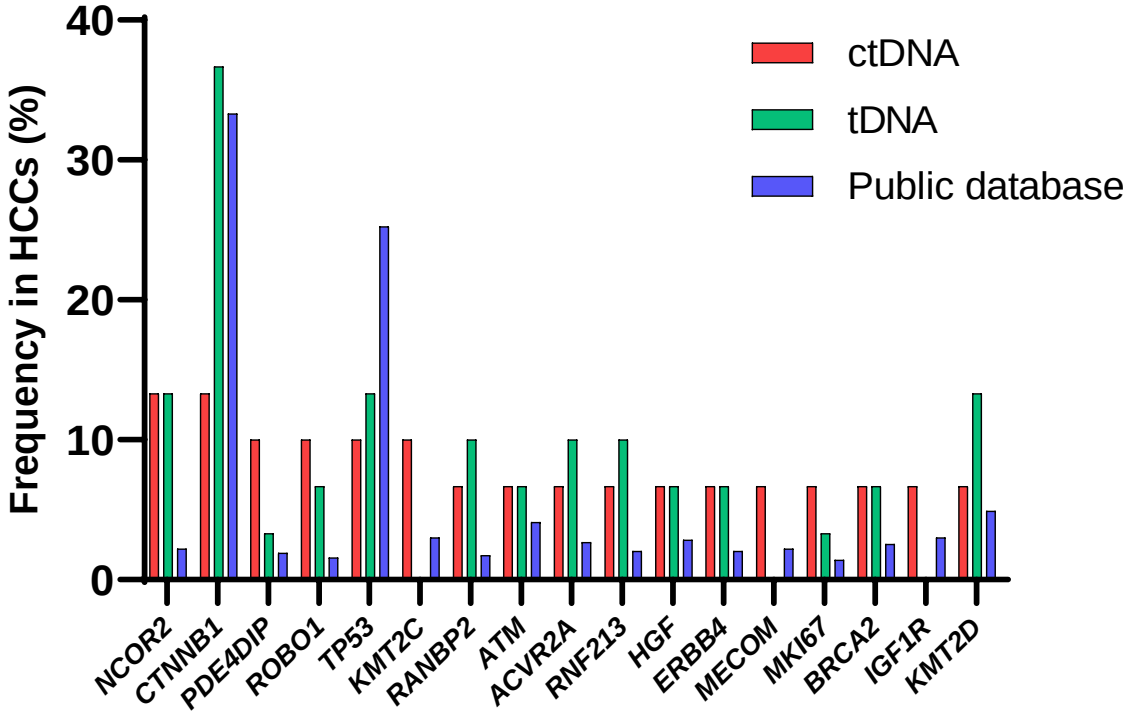
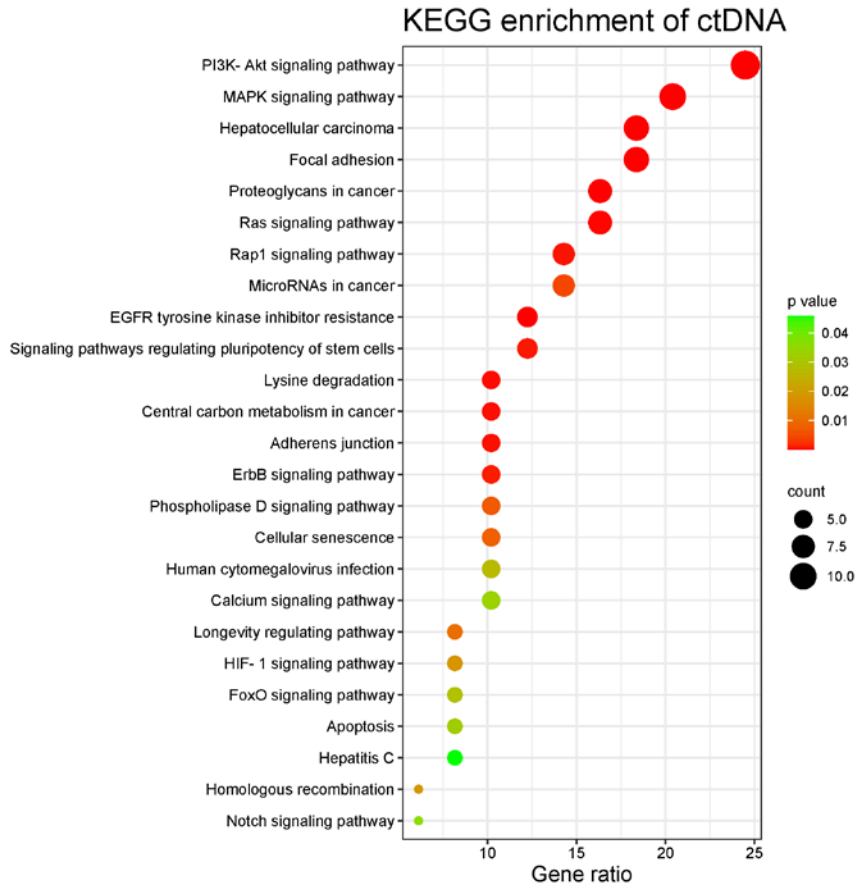


Figure 10. Genetic mutation frequency of HCC in plasma, tumor tissue and public database: Comparative analysis of gene mutation frequency in ctDNA, tDNA and public database (cBioportal website).

4.4.3. Mutation genes and signaling pathways

We conducted KEGG enrichment on the mutant genes identified from both ctDNA and tDNA samples, resulting in several HCC-related signaling pathways (Fig. 11). Among the 25 pathways highly enriched in ctDNA mutations, we observed a stronger enrichment in tDNA as well. In almost all of the pathways, tDNA showed an equal or even higher count of genes than ctDNA, except for the Ras signaling pathway, MAPK signaling pathway and Rap1 signaling pathway. The PI3K-Akt signaling pathway, a representative pathway, showed the highest enrichment in both ctDNA (p=1.06E-06) and tDNA (p=1.36E-07), with 12 and 15 genes, respectively. In ctDNA, the top enriched pathways for the rest included Focal adhesion and MAPK signaling pathways, containing 9 and 10 genes, respectively. Pathways associated with EGFR tyrosine kinase inhibitor resistance (with 6 genes) were relevant to HCC molecularly targeted therapy. Additionally, virus-associated pathways such as hepatitis C (p=0.045) and human cytomegalovirus infection (p=0.027) were enriched in our dataset. Metabolic pathways involving lysine degradation (p=2.84E-04) and central carbon metabolism in cancer (p=4.26E-04) were also found to be relevant. Overall, although some discrepancies were observed in the identified mutations between blood and tissue samples, the distribution and associated pathways exhibited remarkable consistency.

A.


KEGG enrichment of ctDNA
PI3K- Akt signaling pathway
MAPK signaling pathway
Hepatocellular carcinoma
Focal adhesion
Proteoglycans in cancer
Ras signaling pathway
Rap1 signaling pathway
MicroRNAs in cancer
EGFR tyrosine kinase inhibitor resistance
Signaling pathways regulating pluripotency of stem cells
Lysine degradation
Central carbon metabolism in cancer
Adherens junction
ErbB signaling pathway
Phospholipase D signaling pathway
Cellular senescence
Human cytomegalovirus infection
Calcium signaling pathway
Longevity regulating pathway
HIF- 1 signaling pathway
FoxO signaling pathway
Apoptosis
Hepatitis C
Homologous recombination
Notch signaling pathway
10
15
20
25
Gene ratio
p value
0.04
0.03
0.02
0.01
count
5.0
7.5
10.0

B.

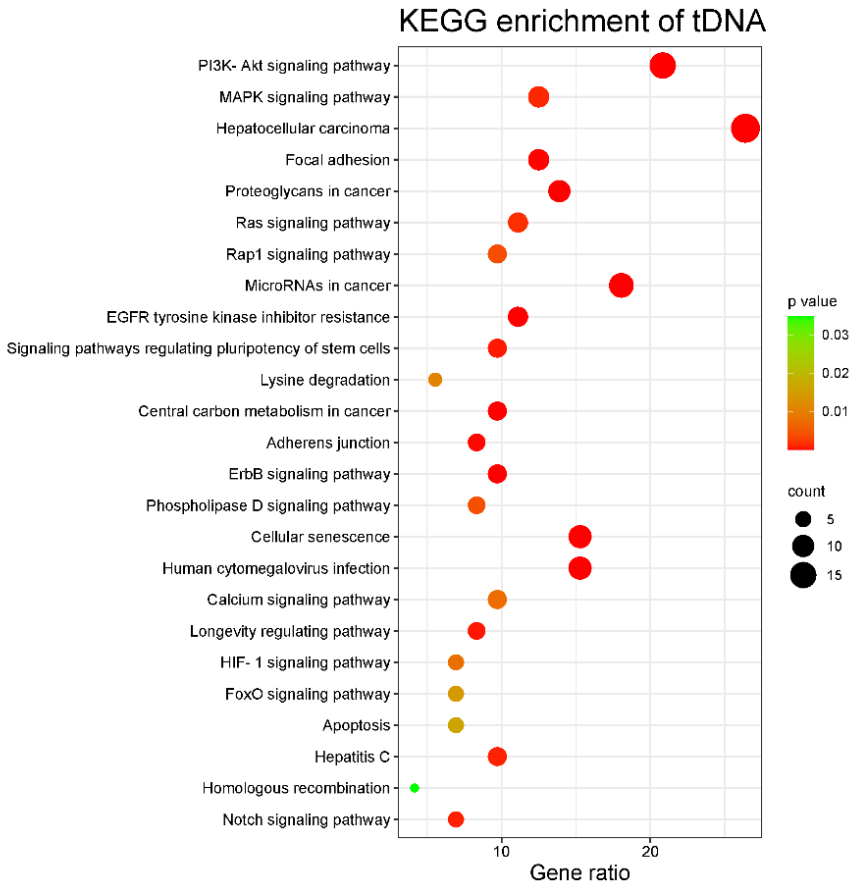

KEGG enrichment of tDNA
PI3K- Akt signaling pathway
MAPK signaling pathway
Hepatocellular carcinoma
Focal adhesion
Proteoglycans in cancer
Ras signaling pathway
Rap1 signaling pathway
MicroRNAs in cancer
EGFR tyrosine kinase inhibitor resistance
Signaling pathways regulating pluripotency of stem cells
Lysine degradation
Central carbon metabolism in cancer
Adherens junction
ErbB signaling pathway
Phospholipase D signaling pathway
Cellular senescence
Human cytomegalovirus infection
Calcium signaling pathway
Longevity regulating pathway
HIF- 1 signaling pathway
FoxO signaling pathway
Apoptosis
Hepatitis C
Homologous recombination
Notch signaling pathway
10
20
Gene ratio
p value
0.03
0.02
0.01
count
5
10
15

Figure 11. Landscapes of mutations of ctDNA and signaling pathway. Totally 25 pathways are selected in ctDNA mutation after enrichment. Each pathway is enriched for at lowest 3 genes with a statistical significance of $p<0.05$. In order to provide a comparative analysis, the enrichment of these pathways in tDNA is also conducted.

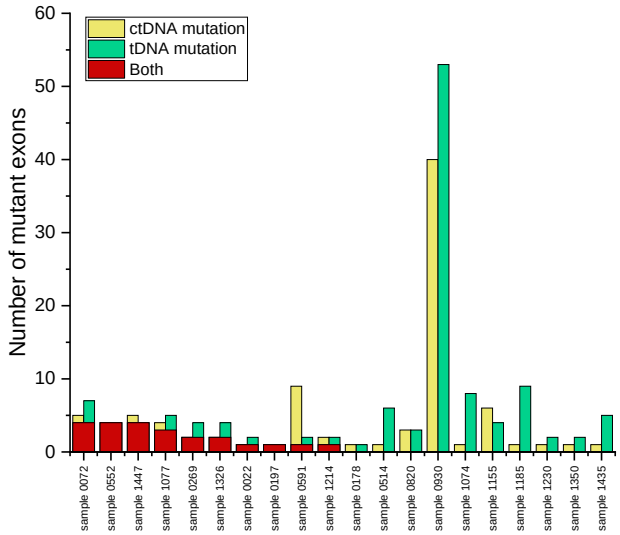
4.5. Concordant mutations and correlation with the clinical factors

After the mutation landscape analysis, our objective was to assess the accuracy of ctDNA in carrying oncogene information. We compared ctDNA and its corresponding matched tDNA extracted from HCC tissue. Then, we checked the concordant mutant genes, which could be consistently detected in both ctDNA and matched tDNA samples simultaneously.

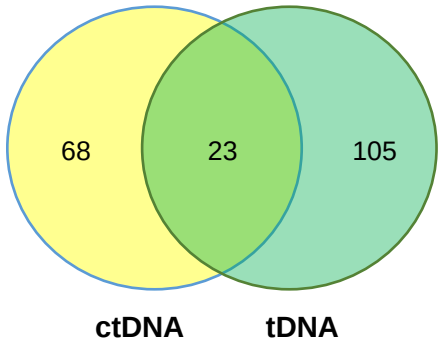
4.5.1. Mutation concordance between plasma ctDNA and matched HCC tDNA

Among the twenty HCCs with at least one functional mutant exon in ctDNA, the concordant mutation genes were detectable in ten, representing 50% of the patients (Fig. 12A). In a single HCC patient, a maximum of 4 concordant mutant exons could be observed. There are a total of 19 genes showing concordant mutations, involving *CTNNB1, TP53, ACVR2A, ALB, ARID2, BAP1, BRCA2, ERBB2, ERBB4, ERCC5, HGF, KMT2D, MSH6, NCOR2, PIK3CA, RANBP2, RNF213, ROBO1,* and *TSC1*. A total of 23 exon sites were contained in these concordant genes, accounting for 25.2% (23/91) of ctDNA mutant exons and 18.0% (23/128) of tDNA mutant exons (Fig. 12B). *CTNNB1* demonstrated the highest frequency in HCC patients, with 4 exons being affected in 40% (4/10) of cases (Fig. 12C). *TP53* was identified in 2 out of the 10 patients, involving 2 exons. The remaining genes had a single mutant exon and were found in only one HCC patient (Table 8). Although the genes showing concordance were not in the majority in either ctDNA or tDNA, concordant mutations were found in half of the patients with plasma mutations, which has profound implications for exploring the clinical value of ctDNA.

A.



B.



C.

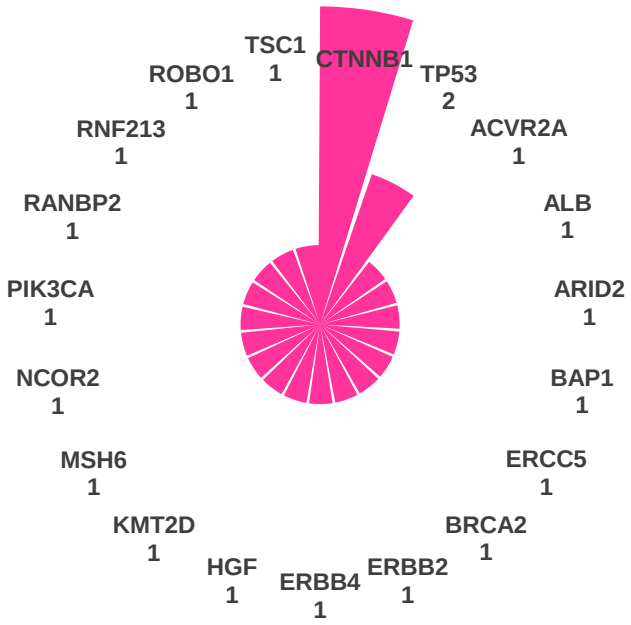


Figure 12. Concordance between mutations identified in ctDNA and matched HCC tDNA. **A.** Our HCC cohort displays a comparison in the number of mutations between ctDNA and tDNA, with concordant mutant genes found in 10 cases. Patient samples were arranged for decreasing concordant mutation. **B.** The Venn diagram illustrates mutant genes' overlapping and distinct portions in ctDNA and tDNA, indicating their concordant and mutually exclusive regions. C. Concodant mutations with different frequencies.

Table 8. Details of the mutations identified both in ctDNA and matched tDNA from HCC patients.

Sample ID	Gene	Alteration type	CytoBand	Position	DNA change	Exonic function
0072	*CTNNB1*	SNP	3p22.1	3:41266113	C > T	nonsynonymous SNV
0269	*CTNNB1*	SNP	3p22.1	3:41266098	A > G	nonsynonymous SNV
0552	*CTNNB1*	SNP	3p22.1	3:41266137	C > T	nonsynonymous SNV
1326	*CTNNB1*	SNP	3p22.1	3:41266100	T > C	nonsynonymous SNV
0197	*TP53*	InDel	17p13.1	17:7578505	GGGCAGGTCT TGGCCAG > -	frameshift deletion
0269	*TP53*	SNP	17p13.1	17:7577580	T > C	nonsynonymous SNV
0552	*ACVR2A*	SNP	2q22.3	2:148683651	C > T	nonsynonymous SNV
1447	*ALB*	InDel	4q13.3	4:74275076	T > -	frameshift deletion
1214	*ARID2*	InDel	12q12	12:46245643	AGG > -	nonframeshift deletion
1447	*BAP1*	InDel	3p21.1	3:52437589	G > -	frameshift deletion
1447	*BRCA2*	SNP	13q13.1	13:32907114	G > T	nonsynonymous SNV
0072	*ERBB2*	SNP	17q12	17:37882817	T > A	nonsynonymous SNV
1326	*ERBB4*	SNP	2q34	2:212248585	A > C	nonsynonymous SNV
0072	*ERCC5*	SNP	13q33.1	13:103515307	T > C	nonsynonymous SNV
1077	*HGF*	SNP	7q21.11	7:81359077	T > G	nonsynonymous SNV
0591	*KMT2D*	SNP	12q13.12	12:49443732	C > A	nonsynonymous SNV

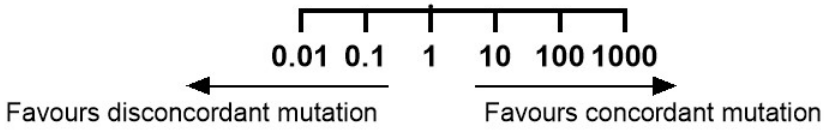
0022	*MSH6*	SNP	2p16.3	2:48027428	T > G	nonsynonymous SNV
0072	*NCOR2*	SNP	12q24.31	12:124841250	C > G	nonsynonymous SNV
0552	*PIK3CA*	SNP	3q26.32	3:178952085	A > G	nonsynonymous SNV
0552	*RANBP2*	InDel	2q12.3	2:109352117	A > -	frameshift deletion
1447	*RNF213*	SNP	17q25.3	17:78265553	C > G	stopgain
1077	*ROBO1*	InDel	3p12.3	3:79174643	G > -	frameshift deletion
1077	*TSC1*	SNP	9q34.13	9:135801023	T > C	nonsynonymous SNV

4.5.2. Clinical variable and concordant mutations

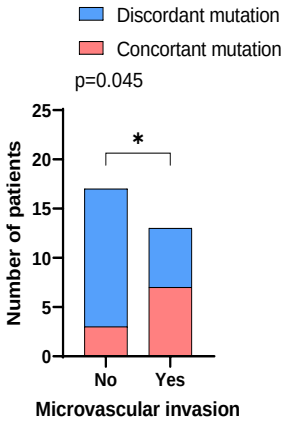
To assess the correlation between concordant mutations and clinic, we conducted the analysis combining the concordant mutations and the clinical data of HCC patients. Our findings revealed a significant association between concordant mutation and vascular invasion as well as BCLC stage (Fig. 13A). Tumor metastasis might be able to influence the probability of the presence of concordant genes in plasma, but there was no statistical difference. HCC patients with tumor vascular invasion exhibited a higher possibility of concordant mutations, both in cases of microvascular invasion (p=0.045) and macrovascular (p=0.030) (Fig. 13B). Moreover, a higher proportion of concordant mutations was observed in patients with BCLC stage B&C compared to those in stages 0&A (p=0.021) (Fig. 13C).

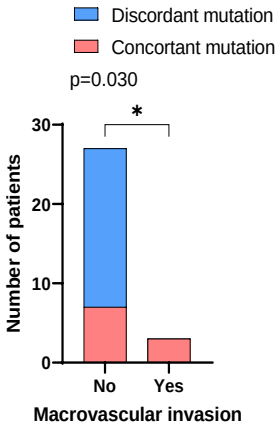
A.

Clinical variable	Odds ratio (Log 10 scale)	OR (95% CI)	P-value
Vascular invasion		5.44 (1.1 – 23.18)	0.045
BCLC stage (B&C)		11 (1.37 - 129.8)	0.021
Metastasis		8.14 (0.98 - 110.0)	0.095
Viral hepatitis		2.79 (0.60 – 10.93)	0.181
TNM stage (3&4)		2.67 (0.47 – 14.60)	0.251
Portal vein thrombosis		2.11 (0.10 – 41.91)	0.302
Gender (male)		1 (0.21 – 3.95)	>0.100
Alcohol liver disease		1 (0.16 – 5.68)	>0.100
Non-alcoholic fatty liver disease		0.52 (0.12 – 2.48)	0.694
Cirrhosis		0.36 (0.09 – 1.67)	0.181



B.





C.

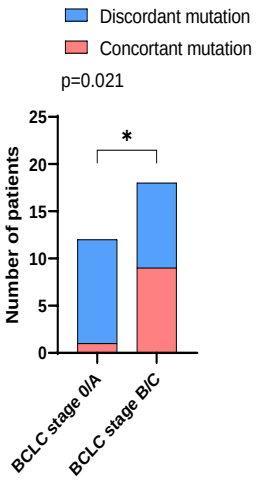


Figure 13. Concordant mutation ratio and HCC patients: A. The association between the clinical variable and concordant mutations is shown in the forest plot. **B.** The stacked charts show the comparison of concordant mutation ratio and the clinical variable with $p<0.05$, including micro/macrovascular invasion, Fisher's exact test, microvascular invasion, $p=0.045$, and macrovascular invasion, $p=0.030$. **C.** Comparison of concordant mutation ratio in patients with different BCLC stages. Fisher's exact test, $p=0.021$.

4.6. CtDNA and HCC diagnosis and prognosis

4.6.1. Combination of ctDNA and AFP for HCC diagnosis

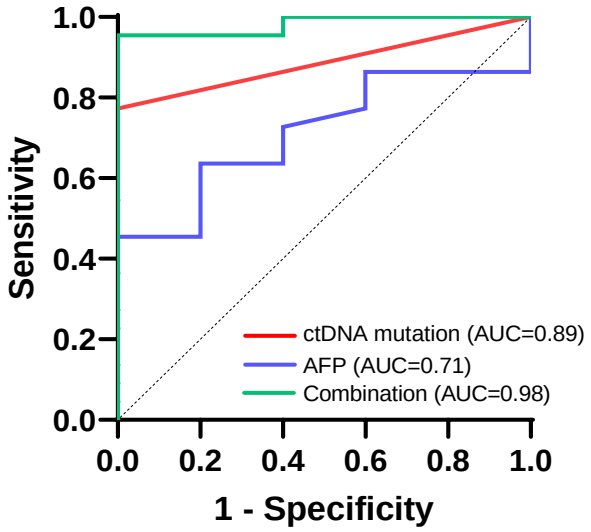
Subsequently, we evaluated the effectiveness of ctDNA mutation for HCC diagnosis and compared it with the conventional HCC biomarker: AFP. The cut-off value of AFP was set at 20ng after reference to the literature [93]. In our cohort, 27 patients took the AFP test, with 22 HCC patients and 5 control group patients. A total of 21 HCC patients were found with elevated AFP or detectable ctDNA mutations in blood samples, including 17 HCCs with ctDNA mutations and 7 were positive for AFP (Table 9).

Table 9. The 2 × 2 table of comparing diagnostics accuracy.

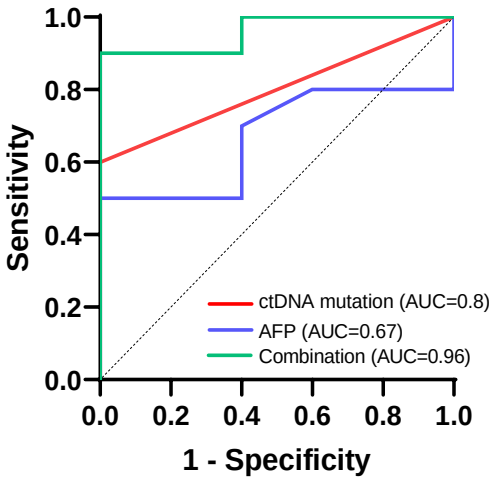
	ctDNA mutation		Total
	Yes	No	
AFP>20ng/ml	3	4	7
AFP≤20ng/ml	14	1	15
Total	17	5	22

In the ROC curve analysis, ctDNA mutation displayed an AUC area of 0.89, indicating a modest increase over AFP level (AUC=0.71). This difference was not statistically significant ($p=0.172$) (Fig. 14A). However, when combining ctDNA mutation with AFP levels, the AUC reached 0.98, revealing a substantial enhancement in diagnostic effect compared to ctDNA mutation or AFP alone. The p-values for the combined approach compared with mutation and AFP were 0.028 and 0.009, respectively (Fig. 14A). Then we also checked the diagnosis accuracy of the ctDNA mutation combined with AFP level for HCC patients in the early stage (BCLC stage 0-A and TNM stage 1) and got the similar results in the ROC curve (Fig. 14B, Fig 14C).

A.



B.



C.

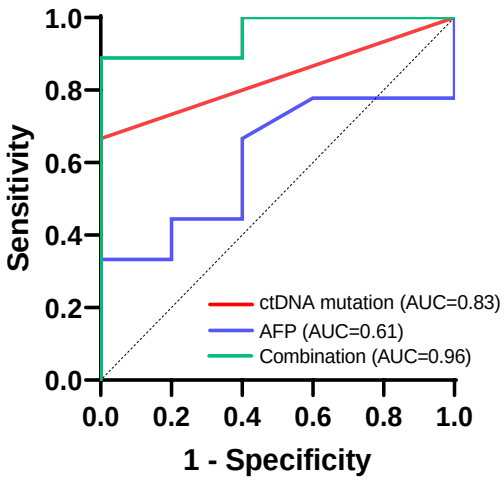


Figure 14: CtDNA mutation and AFP for HCC diagnosis. A. The AUC of ctDNA mutation in the diagnosis ROC curve is 0.89, and the AUC of AFP was 0.71. The change in AUC between ctDNA and AFP levels is not statistically significant (p=0.172). In ROC analysis, The AUC of a combination of ctDNA mutation and AFP is 0.98, with a significant increase in both ctDNA (p=0.028) and AFP (p=0.009). **B.** For HCCs in BCLC stage 0-A, combination AUC showed a significant increase in ctDNA mutation AUC (p=0.042). or AFP AUC (p=0.038). **C.** For HCCs in TNM stage 1, the combination AUC expressed a more significant period than AFP AUC (p=0.025), with no significant difference in ctDNA mutation AUC (p=0.113).

4.6.2. The specific mutation set in ctDNA and prediction of HCC survival

Next, we proceeded with the analysis of the correlation between specific mutations and patient prognosis. Based on the TNM stages, we categorized mutant genes in ctDNA into three mutation sets: those present exclusively in TNM stages 2-4, those appearing only in TNM stage 1, and those occurring in both stages (Fig. 15A). Given that TNM stages 2-4 exhibited poorer overall survival (OS) and progression-free survival (PFS) in our HCC cohort (Fig. 7A), we directed our attention to mutation set A, which comprised 9 patients and 12 genes (*ARID2, ERBB4, ERCC5, KMT2A, MSH6, NCOR2, PIK3CA, PIK3CG, POLQ, PEPRB, TERT,* and *TSC1*) (Table 10).

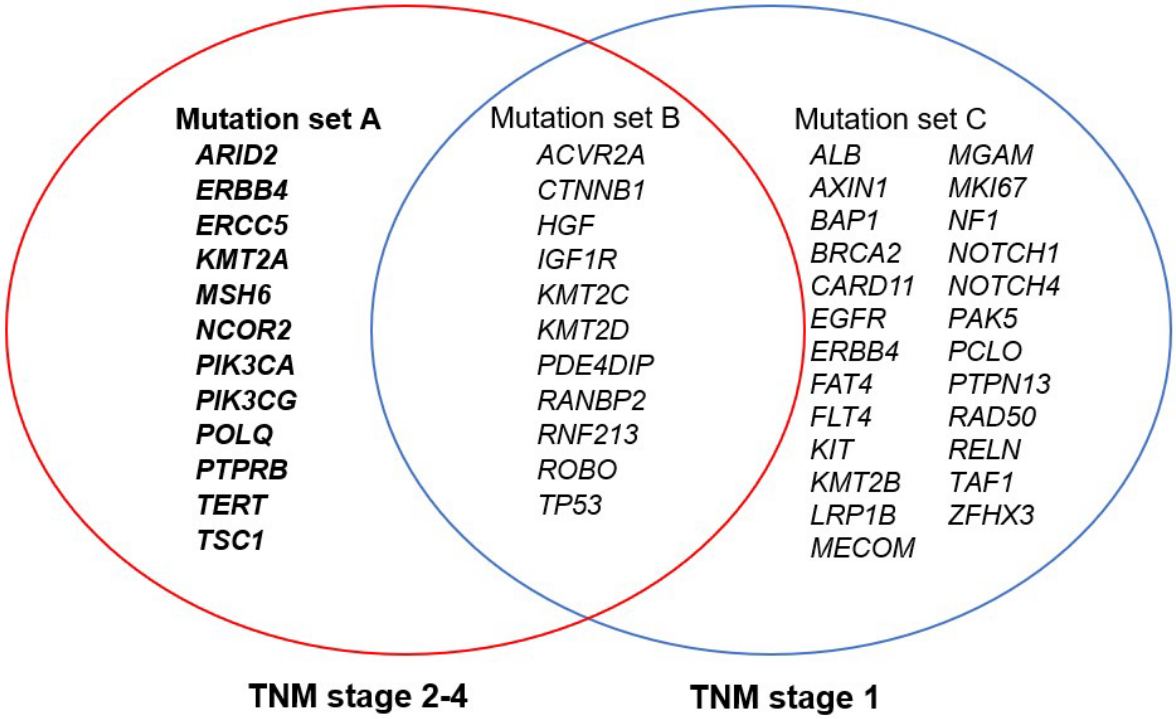


Figure 15. Mutation set of ctDNA and survival: Mutant genes detected only in TNM stages 2-4 are clustered as mutation set A, mutant genes in the overlap part are defined as mutation set B, and mutations only in TNM stage 1 are pooled as mutation set C. The mutation set A includes 12 mutant genes.

Table 10. The patients in TNM stages 2-4 carried genes in mutation set A in our HCC cohort.

	Mutation set A		Total
	Yes	No	
TNM stage 1	0	14	14
TNM stages 2-4	9	7	15
Total	9	21	30

Then we also made the Cox proportional hazards analysis for our HCC cohort (Table 11, Table 12). Univariate analysis showed that late BCLC stage, metastasis, and mutation set A was significantly associated with shorter OS, while late BCLC stage, metastasis, vascular invasion, metastasis, and mutation set A associated with shorter PFS. The mutation set A was an independent risk factor for poor OS and PFS in multivariate analysis.

Table 11. Cox hazard analysis for the prediction of overall survival.

Factor	**Univariate analysis**		**Multivariate analysis**	
	Hazard Ratio (95% CI)	**p-value**	**Hazard Ratio (95% CI)**	**p-value**
Gender (male vs. female)	0.65 (0.12-3.57)	0.624		
Age	0.98 (0.91-1.07)	0.767		
BCLC stage (0-B vs. C)	8.94 (1.63-49.04)	0.012		
Viral hepatitis (yes vs. no)	2.66 (0.49-14.55)	0.259		
Cirrhosis (yes vs. no)	4.14 (0.48-35.47)	0.195	4.58 (0.53-39.44)	0.167

Vascular invasion (yes vs. no)	6.43 (0.75-55.05)	0.089		
Metastasis (yes vs. no)	8.87 (1.733-45.36)	0.009		
Alcohol liver disease (yes vs. no)	2.94 (0.53-16.15)	0.216		
Non-alcoholic fatty liver disease (yes vs. no)	0.69 (0.13-3.76)	0.664		
cfDNA concentration	0.996 (0.93-1.06)	0.916		
ctDNA mutation (yes vs. no)	3.04 (0.35-26.1)	0.31		
Mutation set A (yes vs. no)	5.66 (1.03-30.98)	0.046	6.102 (1.11-33.56)	0.038

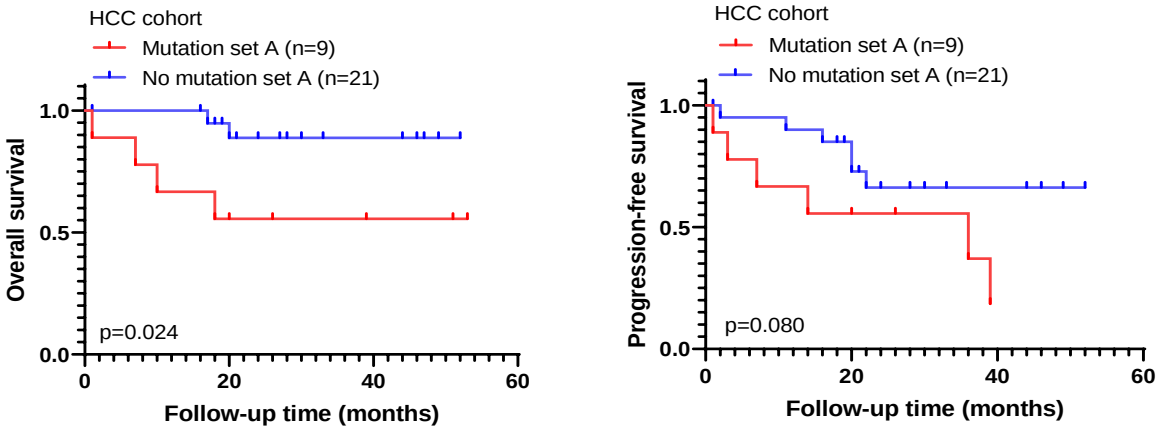
Table 12. Cox hazard analysis for the prediction of progression-free survival.

Factor	**Univariate analysis**		**Multivariate analysis**	
	Hazard Ratio (95% CI)	**p-value**	**Hazard Ratio (95% CI)**	**p-value**
Gender (male vs. female)	0.73 (0.22-2.44)	0.61		
Age	1.01 (0.95-1.07)	0.793		
BCLC stage (0-B vs. C)	5.16 (1.63-16.32)	0.005		
Viral hepatitis (yes vs. no)	1.62 (0.51-5.14)	0.416		
Cirrhosis (yes vs. no)	2.95 (0.80-10.91)	0.106	3.58 (0.94-13.67)	0.062

Vascular invasion (yes vs. no)	3.97 (1.07-14.73)	0.039		
Metastasis (yes vs. no)	4.45 (1.30-15.21)	0.017		
Alcohol liver disease (yes vs. no)	1.91 (0.51-7.09)	0.336		
Non-alcoholic fatty liver disease (yes vs. no)	0.683 (0.21-2.27)	0.534		
cfDNA concentration	0.98 (0.93-1.03)	0.49		
ctDNA mutation (yes vs. no)	1.31 (0.39-4.40)	0.666		
Mutation set A (yes vs. no)	2.64 (0.85-8.23)	0.094	3.22 (1.01-10.24)	0.048

Mutation set A exhibited a strong association with poorer OS (p=0.024) in our HCC cohort and showed a tendency to relate with worse PFS (0.08) (Fig. 16A). To validate the relationship between specific mutation set A and HCC prognosis, we also verified the effect of mutation set A in the TCGA database, which encompassed 174 HCCs for OS and 175 HCCs for PFS. Because our HCC patient population is white, we also select white data in the TCGA database. HCC patients with mutation set A demonstrated significantly poorer PFS (p=0.023) and a negative trend in OS (p=0.074) (Fig. 16B). Importantly, our prognosis data align well with the results observed in the TCGA cohort, affirming the consistency of our findings.

A.



B.

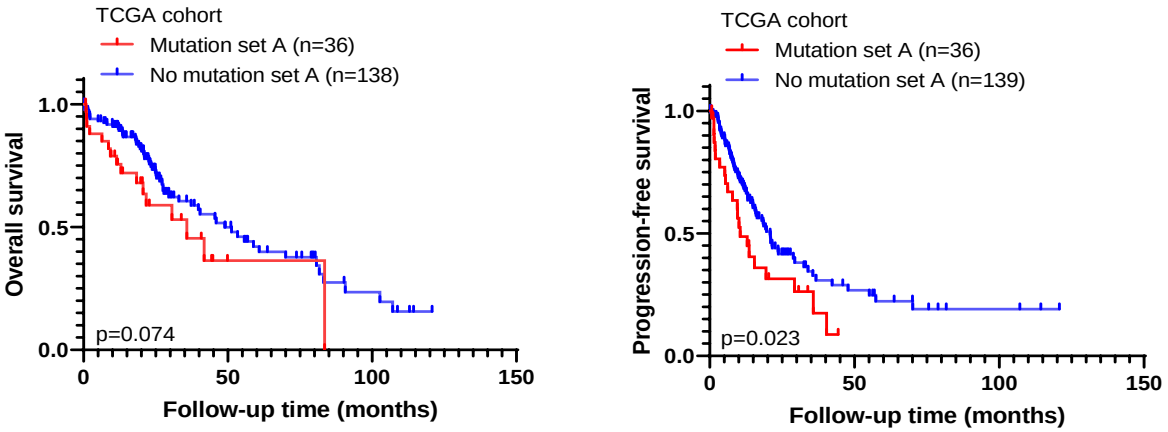


Figure 16. Mutation set A of ctDNA and HCC survival: A. Kaplan–Meier analysis for OS (p=0.024) and PFS (p=0.080) in our HCC cohort. **B.** Kaplan–Meier analysis for OS (p=0.074) and PFS (p=0.023) in the TCGA cohort. P-values were calculated with the log-rank test.

5. DISCUSSION

With the advancement of molecular biology, ctDNA has emerged as a promising tool for diagnosing and monitoring malignant carcinoma in liquid biopsy [94]. CtDNA shares the inherent benefits of liquid biopsy, such as non-invasiveness, real-time, and repeatability analysis. The non-invasive sample collection method could reduce the challenges of obtaining samples multiply. These findings greatly support the feasibility of implementing real-time molecular monitoring for cancer. Furthermore, ctDNA possesses a unique advantage in carrying genetic information from the tumor, enabling precise diagnosis and prognosis [95]. However, despite the potential of ctDNA for the diagnosis and prognosis of HCC, scientists still face many challenges in its application. One of the significant issues is the approach for minimizing the interference of normal cfDNA during ctDNA detection. Current research has explored the feasibility of gene mutations in ctDNA as specific biomarkers. The earliest studies on ctDNA mutations date back to 2000, specifically examining the Ser-249 *p53* mutation in ctDNA of HCC patients in Gambia [96]. Subsequent investigations detected ctDNA mutations in HCCs, revealing more mutated sites beyond *p53*. For instance, Ao and Huang et al. examined Chinese HCC cohorts and identified mutation genes, such as *CTNNTB1*, *Axin1*, *ARID1A*, and *TP53*, in 38.6% of the patients [63]. In ctDNA, common mutant genes associated with HCC could be detected, and genes with low frequencies, such as RAS mutation [64]. Although ctDNA mutations are potential specific markers, the identified mutant targets are still largely random. In our speculation, this randomness may be attributed to tumor heterogeneity [97] and the low mutation specificity of HCC [56].

In the exploratory study, we designed a sequence panel specifically targeting 100 common mutations of HCC. In this panel, 49 genes were in ctDNA mutations, while 72 were detected in tDNA, suggesting a good coverage of HCC-specific mutations. These findings demonstrated the excellent effect of the NGS panel. In addition, more functional mutant sites were detected in tDNA than ctDNA, which means the genetic information of HCC might be more extensive in paraffin tissue than in plasma. Nevertheless, some mutations were only identified in ctDNA, demonstrating that the genetic profile in ctDNA was not entirely encompassed in tDNA. It is probably attributed to the heterogeneity of the tumor.

The cell-free DNA was identified in all the plasma samples from the participants, no matter HCC or benign liver diseases. However, the mutations could only be found in the free nucleic acids of the HCC group and not in the control group. Measurement of concentrations was a typical means of assessing ctDNA/cfDNA, which malignant tumors could elevate [98,99]. In our patient cohort, cfDNA concentration only slightly increased in the HCC group, which might be related to the fact that our control group did not consist of healthy individuals. Still, it implied

that the ctDNA mutation was a better biomarker for distinguishing ctDNA from cfDNA than total cfDNA concentration.

Based on the results, the genetic profiles of ctDNA exhibited unique features. New ctDNA mutations of HCC have been discovered in our cohort, involving *NCOR2, ROBO1, RANBP2, HGF, MECOM, MKI67, PTPN13,* and *ZFHX3.* These mutations mentioned were previously detected in HCC TCGA database; however, they represented the first instances of occurrence in ctDNA. The *TERT* mutation was frequently observed in the HCC database, but its occurrence was relatively low within our cohort. This disparity was attributed to our panel covering the exon region, whereas *TERT* mutations predominantly occur at the promoter site [100]. Within our HCC ctDNA results, *NCOR2* and *CTNNB1* emerged as the two genes with the highest mutation frequency (13.3%). The mutation rate of *NCOR2* aligned with its frequency in paraffin tissue and was higher than reported in other studies [56]. On the contrary, the mutation rate of *CTNNB1* in ctDNA was comparatively lower than it was observed in organized tumor samples, as well as other literature sources [56]. A similar pattern of *CTNNB1* was also observed for *TP53*, another gene frequently mutated in HCC. While some discrepancies existed between plasma and tumor tissues for the identified mutations, the overall enrichment of mutant genes in ctDNA remained consistent with the mutant genes of tDNA. The genes were enriched, focusing on the classical HCC signaling pathways.

Genetic concordance represents whether the genes carried in ctDNA accurately display tumor information. After removing the interference of germline mutations, we confirmed a high concordance in 23 mutant exons between ctDNA and the corresponding tumor tissues. These concordant mutations were found in ten patients, constituting 50% (10/20) of the patients with functional mutations detected in plasma. Howell et al. have checked the concordance between ctDNA and matched tDNA before [61]. However, our study went beyond identifying genetic concordance and explored a relationship between mutational concordance and clinical information. We explored several clinical factors that may influence the accuracy of ctDNA in capturing the genetic information of HCC tissues. Interestingly, mutation concordance strongly correlated with vascular invasion, encompassing both micro and macrovascular invasion. This discovery confirmed that ctDNA could convey genetic information from primary tumors, with enhanced accuracy in the presence of vascular invasion and a high tumor burden.

In addition, our findings revealed the vital role of ctDNA mutations in the diagnosis and predicting prognosis of HCC. Recently, a novel detection tool named “CancerSEEK” was developed by Cohen JD and colleagues, which combines ctDNA mutations with circulating proteins for tumor detection [101]. Then, a positive impact of combining ctDNA mutations with Des-Gamma-Carboxy Prothrombin (DCP) has been previously demonstrated [102]. While

previous studies have tried to explore the combination of AFP levels and ctDNA, they mainly focus on copy number variation or ctDNA methylation [103,104]. Although a liquid biopsy assay incorporating ctDNA mutations, AFP, and DCP has been developed for diagnosing liver cancer, its application remains limited to HBV-associated HCC [105]. Hence, our study takes a distinctive approach by combining ctDNA mutations with AFP to evaluate their collective efficacy for HCC diagnosis, marking a significant advancement in this field. In our data, mutation of ctDNA exhibited only a modest performance improvement compared to AFP. However, when combined, these two biomarkers demonstrated a significant advantage in diagnosing HCC for early diagnosis and tumors up to 2 cm in diameter. The diagnostic accuracy of ctDNA mutation combining AFP level surpassed either biomarker alone.

Furthermore, existing studies have emphasized the predictive value of ctDNA mutations for HCC prognosis, but most focus on individual mutant genes [102 106]. Our study goes beyond this by demonstrating that grouping specific mutant genes into mutation sets could enhance prognosis prediction in HCC patients. The newly detected mutation of ctDNA (*NCOR2*) was also enrolled in the mutation set, indirectly illustrating the impact of *NCOR2* on HCC prognosis. Given the relatively low frequency of single mutant genes in HCC (with *TP53* and *CTNNB1*, the most commonly mutated genes occurring in less than 40% of cases), our research assesses the value of the mutation set approach.

We also acknowledge some limitations of our study. Firstly, although we tried our best to maximize the inclusion of the common genes of HCC in the sequencing panel, there were still many random specific mutation genes in the result. It's probably due to our small sample size of the patient cohort. Secondly, whereas genetic mutations have been found in patients with benign liver diseases (such as hepatocellular adenomas and liver cirrhosis) in several studies, we did not detect any mutations in the control group [107 108]. Finally, our single time-point test of ctDNA limited the study's power to observe the dynamic changes in HCC. Ideally, multiple time-point sampling allowed a clear view of changes in ctDNA throughout HCC progression, including preoperative, postoperative, and recurrence.

Our study has provided compelling evidence demonstrating the strong potential of ctDNA mutations as a specific biomarker for liquid biopsy in HCC. Importantly, we have shown that ctDNA mutations offer a higher ability to discriminate between tumors and benign diseases than cfDNA concentration. Then, we also observed the rate of concordant mutations related to tumor burden, especially vascular invasion. In addition, co-diagnosis of ctDNA mutation and AFP shows excellent potential in HCC. The identification of specific mutation sets in ctDNA holds promise for improving prognosis predictions. Although limited by sample capacity, our findings strongly support the use of ctDNA mutations in advancing precision medicine.

Pursuing this direction and expanding the sample size will yield valuable insights into the clinical value associated with ctDNA mutations.

In conclusion, ctDNA mutation from plasma exhibits the capacity to distinguish malignant and benign liver diseases. As the tumor progresses, ctDNA would capture more accurate genetic information from tumors in HCC patients, particularly with vascular invasion. Combined diagnosis of ctDNA mutation and AFP will generate remarkable outcomes in HCC. Moreover, the specific mutation set in ctDNA strongly predicts the HCC patient's prognosis. Our research reveals that ctDNA carries the genetic information consistent with tumors, demonstrating the massive potential of ctDNA mutations as a novel biomarker for diagnosing and monitoring HCC (Fig. 17).

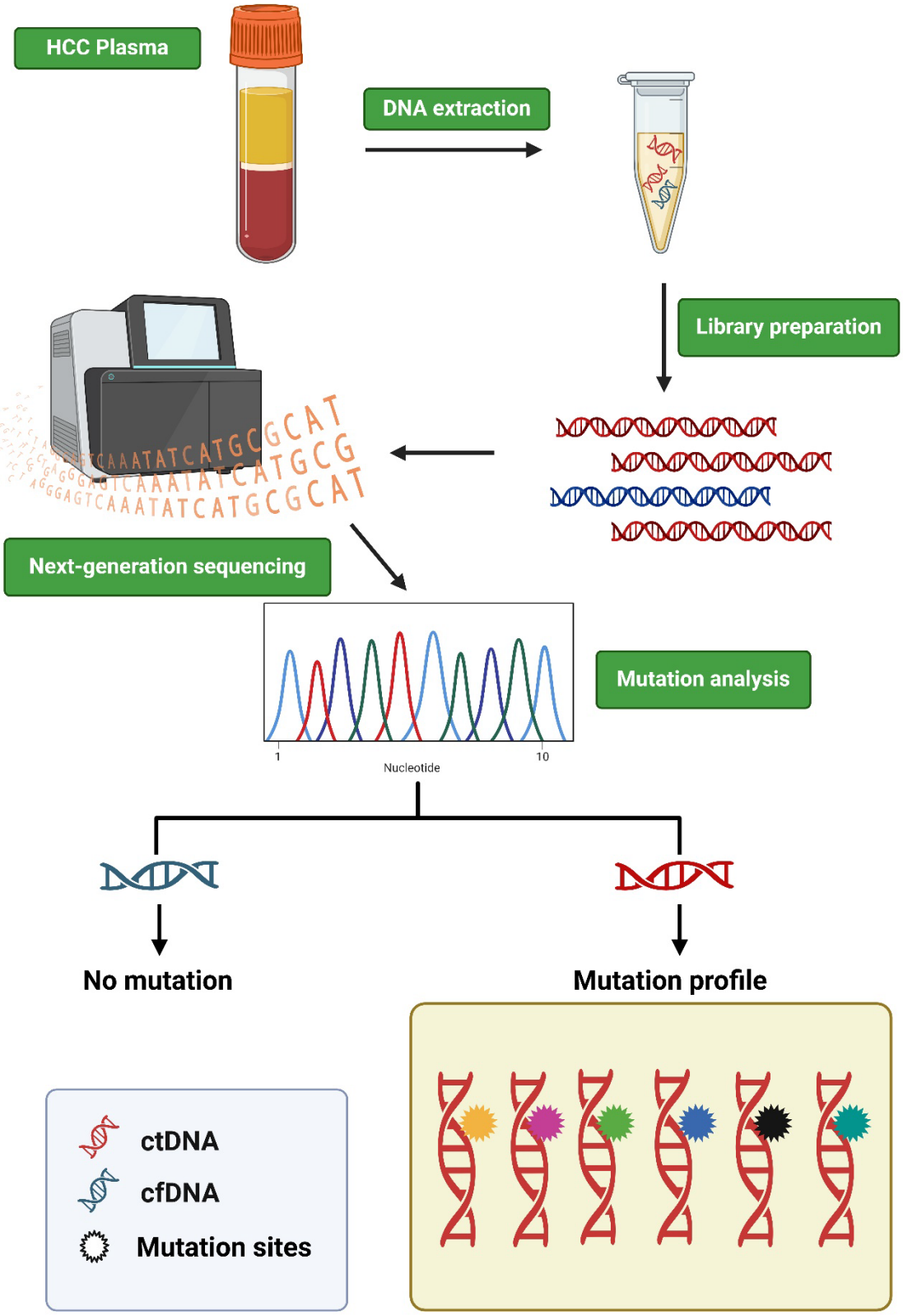


Figure 17. The free nucleic acid fragments (ctDNA and cfDNA) are extracted from plasma of HCC patients and checked in next-generation sequencing. Mutations in ctDNA (stars of different colors) could be a valuable biomarker for the distinction between ctDNA and cfDNA. Various mutations are carried in ctDNA, constituting a unique mutation profile and a strong association with clinical feature.

6. REFERENCE

1. Singal AG, Lok AS, Feng Z, Kanwal F, Parikh ND. Conceptual model for the hepatocellular carcinoma screening continuum: Current status and research agenda. *Clin Gastroenterol Hepatol* 2022; **20**(1): 9–18.
2. Villanueva A. Hepatocellular carcinoma. *The New England Journal Of Medicine* 2019; **380**(15): 1450-62.
3. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinician* 2021; **71**(3): 209-49.
4. Josep M. Llovet, Kelley RK, Villanueva A, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers* 2021; **7**(1): 6.
5. Liver EAftSot. EASL Clinical practice guidelines management of hepatocellular carcinoma. *Journal of Hepatology* 2018; **69**(1): 182-236.
6. Collaboration GBoDLC. The burden of primary liver cancer and underlying etiologies from 1990 to 2015 at the global, regional, and national level results from the global burden of disease study 2015. *JAMA Oncology* 2017; **3**(12): 1683-91.
7. McGlynn1 KA, Petrick JL, El-Serag HB. Epidemiology of hepatocellular carcinoma. *Hepatology* 2021; **73()**(Suppl 1): 4–13.
8. Axley P, Ahmed Z, Ravi S, Singal AK. Hepatitis c virus and hepatocellular carcinoma a narrative review. *Journal of Clinical and Translational Hepatology* 2018; **6**(1): 79–84.
9. Huang DQ, Mathurin P, Cortez-Pinto H, Loomba R. Global epidemiology of alcohol-associated cirrhosis and HCC: trends, projections and risk factors. *Nature reviews Gastroenterology & hepatology* 2023; **20**(1): 37-49.
10. Anstee QM, Reeves HL, Kotsiliti E, Govaere O, Heikenwalder M. From NASH to HCC current concepts and future challenges. *Nature reviews Gastroenterology & hepatology* 2019; **16**(7): 411-28.
11. Zhu Q, Ma Y, Liang J, et al. AHR mediates the aflatoxin B1 toxicity associated with hepatocellular carcinoma. *Signal Transduction and Targeted Therapy* 2021; **6**(1): 299.
12. toni EnD, raab AS-, Fuchs M, et al. Age independent survival benefit for patients with hepatocellular carcinoma (HCC) without metastases at diagnosis a population- based study. *Gut* 2020; **69**(1): 168-76.
13. Manieri E, Herrera-Melle L, Mora A, et al. Adiponectin accounts for gender differences in hepatocellular carcinoma incidence. *Journal of experimental medicine* 2019; **216**(5): 1108-19.
14. Tsuchiya N, Sawada Y, Endo I, Saito K, Uemura Y, Nakatsura T. Biomarkers for the early diagnosis of hepatocellular carcinoma. *World J Gastroenterol* 2015 October 7; **21(37)**.

15. Roberts LR, Sirlin CB, Zaiem F, et al. Imaging for the diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis. *HEPATOLOGY* 2018; **67**(1): 401-21.
16. Trevisani F, D'Intino PE, Morselli-Labate AM, et al. Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. *Journal of Hepatology* 2001; **34**(4): 570-5.
17. Schlageter M, Terracciano LM, D'Angelo S, Sorrentino P. Histopathology of hepatocellular carcinoma. *World journal of gastroenterology* 2014; **20**(43): 15955-64.
18. ND P, N T, AG S. Blood-based biomarkers for hepatocellular carcinoma screening: Approaching the end of the ultrasound era? . *J Hepatol* 2023; **78**(1): 207-16.
19. JA M, Z F, Y W, et al. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprote in in early hepatocellular carcinoma. *Gastroenterology* 2009; **137**(1): 110-8.
20. Shen Q, Fan J, Yang X-R, et al. Serum DKK1 as a protein biomarker for the diagnosis of hepatocellular carcinoma: a large-scale, multicentre study. *Lancet Oncol* 2012; **13**(8): 817-26.
21. JA M, PR R, O N, et al. GP73, a resident Golgi glycoprotein, is a novel serum marker for hepatocellular carcinoma. *J Hepatol* 2005; **43**(6): 1007-12.
22. S S, A P, S G, et al. Identification of osteopontin as a novel marker for early hepatocellular carcinoma. *Hepatology* 2012; **55**(2): 483-90.
23. WW Z, JJ G, L G, et al. Evaluation of midkine as a diagnostic serum biomarker in hepatocellular carcinoma. *Clin Cancer Res* 2013; **19**(14): 3944-54.
24. Y J, ES J, YS C, JW K, SH J. Glypican-3 level assessed by the enzyme-linked immunosorbent assay is inferior to alpha-fetoprotein level for hepatocellular carcinoma diagnosis. *Clin Mol Hepatol* 2016; **22**(3): 359-65.
25. Alix-Panabières C, Pantel K. Liquid biopsy: From discovery to clinical application. *Cancer discovery* 2021; **11**(4): 858-73.
26. Nikanjam M, Kato S, Kurzrock R. Liquid biopsy: current technology and clinical applications. *Journal of Hematology & Oncology* 2022; **15**(1): 131.
27. Ye Q, Ling S, Zheng S, Xu X. Liquid biopsy in hepatocellular carcinoma circulating tumor cells and circulating tumor DNA. *Molecular Cancer* 2019; **18**(1): 114.
28. Lin D, Shen L, Luo M, et al. Circulating tumor cells: biology and clinical significance. *Signal transduction and targeted therapy* 2021; **6**(1): 404.
29. Debbi L, Guo S, Safina D, Levenberg S. Boosting extracellular vesicle secretion. *Biotechnology Advances* 2022; **59**: 107983.
30. Lampignano R, Kloten V, Krahn T, Schlange T. Integrating circulating mirna analysis in the clinical management of lung cancer: present or future? *Molecular Aspects of Medicine* 2020; **72**: 100844.

31. Sasimol Udomruk, Santhasiri Orrapin, Dumnoensun Pruksakorn, Chaiyawat P. Size distribution of cell-free DNA in oncology. *Critical Reviews in Oncology / Hematology* 2021; **166: 103455**.

32. P M, P M. Nuclear acids in human blood plasma. *C R Seances Soc Biol Fil* 1948; **142**(3-4): 241-3.

33. Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the serum of cancer patients and the effect of therapy. *CANCER RESEARCH* 1977; **37**(3): 646-50.

34. Sanchez C, Roch B, Mazard T, et al. Circulating nuclear DNA structural features, origins, and complete size profile revealed by fragmentomics. *JCI Insight* 2021; **6**(7).

35. Corcoran RB, Chabner BA. Application of cell-free DNA analysis to cancer treatment. *The New England Journal of Medicine* 2018; **379**(18): 1754-65.

36. Kaseb AO, Sánchez NS, Sen S, et al. Molecular profiling of hepatocellular carcinoma using circulating cell-free DNA. *Clinical Cancer Research* 2019; **25**(20): 6107-18.

37. García-Pardo M, Makarem M, Li JJN, Kelly D, Leighl NB. Integrating circulating-free DNA (cfDNA) analysis into clinical practice: opportunities and challenges. *British Journal of Cancer* 2022; **127**(4): 592-602.

38. Rolfo C, Mack PC, Scagliotti GV, et al. Liquid biopsy for advanced non-small cell lung cancer (nsclc): A statement paper from the IASLC. *Journal of Thoracic Oncology* 2018; **13**(9): 1248-68.

39. Moding EJ, Nabet BY, Alizadeh AA, Diehn M. Detecting liquid remnants of solid tumors: Circulating tumor DNA minimal residual disease. *Cancer discovery* 2021; **11**(12): 2968-86.

40. Shendure J, Balasubramanian S, Church GM, et al. DNA sequencing at 40: past, present and future. *Nature* 2017; **550**(7676): 345-53.

41. Diekstra A, Bosgoed E, Rikken A, et al. Translating sanger-based routine DNA diagnostics into generic massive parallel ion semiconductor sequencing. *Clinical Chemistry* 2015; **61**(1): 154-62.

42. Shendure J, Ji H. Next-generation DNA sequencing. *Nature Biotechnology* 2008; **26**(10): 1135-45.

43. Dijk ELv, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends in Genetics* 2014; **30**(9): 418-26.

44. Hindson BJ, Ness KD, Masquelier DA, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Analytical Chemistry* 2011; **83**(22): 8604-10.

45. Palacín-Aliana I, García-Romero N, Asensi-Puig A, Carrión-Navarro J, González-Rumayor V, Ayuso-Sacido Á. Clinical utility of liquid biopsy-based actionable mutations detected via ddPCR. *Biomedicines* 2021; **9**(8): 906.

46. Barbany G, Arthur C, Liedén A, et al. Cell-free tumour DNA testing for early detection of cancer - a potential future tool. 2019; **286**(2): 118-36.

47. Xu H, Zhu X, Xu Z, et al. Non-invasive analysis of genomic copy number variation in patients with hepatocellular carcinoma by next generation DNA sequencing. *Journal of Cancer* 2015; **6**(3): 247-53.

48. Jiang P, Sun K, Tong YK, et al. Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma. *Proceedings of the national academy of sciences of the United States of America* 2018; **115**(46).

49. Piciocchi M, Cardin R, Vitale A, et al. Circulating free DNA in the progression of liver damage to hepatocellular carcinoma. *Hepatology International* 2013; **7**(4): 1050-7.

50. Xu RH, Wei W, Krawczyk M, et al. Circulating tumour DNA methylation markers for diagnosis and prognosis for diagnosisandprognosis of hepatocellular carcinoma. *Nature Materials* 2017; **16**.

51. Wu X, Li J, Gassa A, et al. Circulating tumor DNA as an emerging liquid biopsy biomarker for early diagnosis and therapeutic monitoring in hepatocellular carcinoma. *Int J Biol Sci* 2020; **16**(9): 1551-62.

52. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability-an evolving hallmark of cancer. *Nature reviews Molecular cell biology* 2010; **11**(3): 220-8.

53. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. *Science* 2015; **349**(6255): 1483-9.

54. Rebouissou S, Nault J-C. Advances in molecular classification and precision oncology in hepatocellular carcinoma. *Journal of Hepatology* 2020; **72**(2): 215-29.

55. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature* 2013; **500**(7463): 415-21.

56. Network. CGAR. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* 2017; **169**(7): 1327-41.

57. Nakamura Y, Taniguchi H, Ikeda M, et al. Clinical utility of circulating tumor DNA sequencing in advanced gastrointestinal cancer SCRUM-Japan GI-SCREEN and GOZILA studies. *Nature Medicine* 2020; **26**(12): 1859-64.

58. Phallen J, Sausen M, Adleff V, et al. Direct detection of early-stage cancers using circulating tumor DNA. *Science translational medicine* 2017; (9): eaan2415

59. Dang DK, Park BH. Circulating tumor DNA: current challenges for clinical utility. *The Journal of Clinical Investigation* 2022; **132**(12).

60. Ng CKY, Di Costanzo GG, Tosti N, et al. Genetic profiling using plasma-derived cell-free DNA in therapy-naive hepatocellular carcinoma patients: a pilot study. *Annals of Oncology* 2018; **29**(5): 1286-91.

61. Howell J, Atkinson SR, Pinato DJ, et al. Identification of mutations in circulating cell-free tumour DNA as a biomarker in hepatocellular carcinoma. *European Journal of Cancer* 2019; **116** 56-66.

62. Felden Jv, Craig AJ, Garcia-Lezana T, et al. Mutations in circulating tumor DNA predict primary resistance to systemic therapies in advanced hepatocellular carcinoma. *Oncogene* 2021; **40**(1): 140-51.
63. Huang A, Zhao X, Yang XR, et al. Circumventing intratumoral heterogeneity to identify potential therapeutic targets in hepatocellular carcinoma. *Journal of Hepatology* 2017; **67**(2): 293-301.
64. Lim HY, Merle P, Weiss KH, et al. Phase II Studies with Refametinib or Refametinib plus Sorafenib in Patients with RAS-Mutated Hepatocellular Carcinoma. *Clinical Cancer Research* 2018; **24**(19): 4650-61.
65. Yuan X, Larsson C, Xu D. Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players. *Oncogene* 2019; **38**(34): 6172-83.
66. Rao CV, Asch AS, Yamada HY. Frequently mutated genes/pathways and genomic instability as prevention targets in liver cancer. *Carcinogenesis* 2017; **38**(1): 2-11.
67. Qiao Y, Wang J, Karagoz E, et al. Axis inhibition protein 1 (Axin1) deletion-induced hepatocarcinogenesis requires intact β-catenin but not notch cascade in mice. *Hepatology* 2019; **70**(6): 2003-17.
68. Huang L, Guo Z, Wang F, Fu L. KRAS mutation: from undruggable to druggable in cancer. *Signal Transduction and Targeted Therapy* 2021; **6**(1): 386.
69. Punekar SR, Velcheti V, Neel BG, Wong K-K. The current state of the art and future trends in RAS-targeted cancer therapies. *Nature reviews Clinical oncology* 2022; **19**(10): 637-55.
70. Neuzillet C, Tijeras-Raballand A, Mestier Ld, Cros J, Faivre S, Raymond E. MEK in cancer and cancer therapy. *Pharmacology & Therapeutics* 2014; **141**(2): 160-71.
71. Huynh H, Nguyen TTT, Chow K-HP, Tan PH, Soo KC, Tran E. Over-expression of the mitogen-activated protein kinase (MAPK) kinase (MEK)-MAPK in hepatocellular carcinoma its role in tumor progression and apoptosis. *BMC Gastroenterology* 2003; **3**: 19.
72. Ito Y, Sasaki Y, Horimoto M, et al. Activation of mitogen-activated protein kinases/extracellular signal-regulated kinases in human hepatocellular carcinoma. *Hepatology* 1998; **27**(4): 951-8.
73. Dratwa M, Wysoczańska B, Łacina P, Kubik T, Bogunia-Kubik K. TERT—regulation and roles in cancer formation. *Frontiers in immunology* 2020; **11**: 589929.
74. Nault J-C, Ningarhari M, Rebouissou S, Zucman-Rossi J. The role of telomeres and telomerase in cirrhosis and liver cancer. *Nature Reviews Gastroenterology & Hepatology* 2019; **16**(9): 544-58.
75. Pinyol R, Tovar V, Llovet JM. TERT promoter mutations: gatekeeper and driver of hepatocellular carcinoma. *Journal of hepatology* 2014; **61**(3): 685-7.

76. Williams AB, Schumacher B. p53 in the DNA-Damage-Repair Process. *Cold Spring Harbor Perspectives in Medicine* 2016; **6**(5): a026070.
77. Hussain SP, Schwank J, Staib F, Wang XW, Harris CC. TP53 mutations and hepatocellular carcinoma: insights into the etiology and pathogenesis of liver cancer. *Oncogene* 2007; **26**(15): 2166-76.
78. Perugorria MJ, Olaizola P, Labiano I, et al. Wnt-β-catenin signalling in liver development, health and disease. *Nature reviews Gastroenterology & hepatology* 2019; **16**(2): 121-36.
79. Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. *Oncogene* 2017; **36**(11): 1461-73.
80. Xu C, Xu Z, Zhang Y, Evert M, Calvisi DF, Chen X. β-Catenin signaling in hepatocellular carcinoma. *The journal of clinical investigation* 2022; **132** (4): e154515.
81. Song X, Wang S, Li L. New insights into the regulation of Axin function in canonical Wnt signaling pathway. *Protein Cell* 2014; **5**(3): 186-93.
82. Mittal P, Roberts CWM. The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nature reviews Clinical oncology* 2020; **17**(7): 435-48.
83. Mathur R. ARID1A loss in cancer: Towards a mechanistic understanding. *Pharmacology & Therapeutics* 2018; **190**: 15-23.
84. Mullen J, Kato S, Sicklick JK, Kurzrock R. Targeting ARID1A mutations in cancer. *Cancer treatment reviews* 2021; **100**: 102287.
85. Sun X, Wang SC, Wei Y, et al. Arid1a Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer Cell* 2017; **32**(5): 574-89 e6.
86. van Dessel LF, Beije N, Helmijr JC, et al. Application of circulating tumor DNA in prospective clinical oncology trials - standardization of preanalytical conditions. *Mol Oncol* 2017; **11**(3): 295-304.
87. Harding JJ, Nandakumar S, Armenia J, et al. Prospective genotyping of hepatocellular carcinoma clinical implications of next-generation sequencing for matching patients to targeted and immune therapies. *Clinical Cancer Research* 2019; **25**(7): 2116-26.
88. Schulze K, Imbeaud S, Letouzé E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nature Genetics* 2015; **47**(5): 505–11.
89. Ahn S-M, Jang SJ, Shim JH, et al. Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification.pdf. *Journal of Hepatology* 2014; **60**(6): 1972-82.
90. Fujimoto A, Totoki Y, Abe T, et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nature Genetics* 2012; **44**(7): 760-4.

91. Sobesky S, Mammadova L, Cirillo M, et al. In-depth cell-free DNA sequencing reveals genomic landscape of Hodgkin's lymphoma and facilitates ultrasensitive residual disease detection. *Med* 2021; **2**(10): 1171–93.
92. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016; **32**(18): 2847–9.
93. Ding Y, Feng M, Ma D, et al. The 20 years transition of clinical characteristics and metabolic risk factors in primary liver cancer patients from China. *Front Oncol* 2023; **13**: 1109980.
94. Felden Jv, Lezana TG-, Schulze K, Losic B, Villanueva A. Liquid biopsy in the clinical management of hepatocellular carcinoma. *Gut* 2020; **69**(11): 2025-34.
95. Zhang Y, Yao Y, Xu Y, et al. Pan-cancer circulating tumor DNA detection in over 10,000 chinese patients. *NATURE COMMUNICATIONS* 2021; **12**(1): 11.
96. Kirk GD, Anne-Marie, Camus-Randon, et al. Ser-249 p53 mutations in plasma DNA of patients with hepatocellular carcinoma from The Gambia. *Journal of the National Cancer Institute* 2000; **92**(2): 148-53.
97. Fridland S, Choi J, Nam M, et al. Assessing tumor heterogeneity integrating tissue and circulating tumor DNA (ctDNA) analysis in the era of immuno-oncology-blood TMB is not the same as tissue TMB. *Journal of immunotherapy of cancer* 2021; **9**(8).
98. Alborelli I, Generali D, Jermann P, et al. Cell-free DNA analysis in healthy individuals by next-generation sequencing: a proof of concept and technical validation study. *Cell death & disease* 2019; **10**(7): 534.
99. Chen Y, Li X, Liu G, et al. ctdna concentration, miki67 mutations and hyper-progressive disease related gene mutations are prognostic markers for camrelizumab and apatinib combined multiline treatment in advanced NSCLC. *Frontiers in oncology* 2020; **10**: 1706.
100. Hirai M, Kinugasa H, Nouso K, et al. Prediction of the prognosis of advanced hepatocellular carcinoma by TERT promoter mutations in circulating tumor DNA. *Journal of Gastroenterology and Hepatology* 2021; **36**(4): 1118-25.
101. Cohen JD, Li L, Wang YX, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* 2018; **359**(6378): 926–30.
102. Cai Z, Chen G, Zeng Y, et al. Comprehensive liquid profiling of circulating tumor DNA and protein biomarkers in long-term follow-up patients with hepatocellular carcinoma. *Clinical Cancer Research* 2019; **25**(17): 5284-94.
103. Meng Z, Ren Q, Zhong G, et al. Noninvasive detection of hepatocellular carcinoma with circulating tumor DNA features and α-fetoprotein. *The Journal of Molecular Diagnostics* 2021; **23**(9): 1174-84.

104. Chalasani NP, Porter K, Bhattacharya A, et al. Validation of a Novel Multitarget Blood Test Shows High Sensitivity to Detect Early Stage Hepatocellular Carcinoma. *Clinical Gastroenterology and Hepatology* 2022; **20**(1): 173–82.
105. Qu C, Wang Y, Wang P, et al. Detection of early-stage hepatocellular carcinoma in asymptomatic HBsAg-seropositive individuals by liquid biopsy. *PNAS* 2019; **116**(13): 6308-12.
106. Matsumae T, Kodama T, Myojin Y, et al. Circulating cell-free DNA profiling predicts the therapeutic outcome in advanced hepatocellular carcinoma patients treated with combination immunotherapy. *Cancers* 2022; **14**(14): 3367.
107. Raft MB, Jørgensen EoN, Vainer B. Gene mutations in hepatocellular adenomas. *Histopathology* 2015; **66**(7): 910–21.
108. Müller M, Bird TG, Nault J-C. The landscape of gene mutations in cirrhosis and hepatocellular carcinoma. *Journal of Hepatology* 2020; **72**(5): 990-1002.