Spring 9-8-2022

Finding Biomarkers from Long-Non-Coding RNAs in Serum of Patients with Non-Alcoholic Fatty Liver Disease

Nouran Yonis
The American University in Cairo AUC, nouranyonis@aucegypt.edu

THE AMERICAN UNIVERSITY IN CAIRO
الجامعة الأمريكية بالقاهرة

Finding biomarkers from long-non coding RNAs in serum of patients with non-alcoholic fatty liver disease

A Thesis Submitted to
Institute of Global Health and Human Ecology
In partial fulfillment of the requirements for
The degree of Master of Science
By:

Nouran Samy Ahmed Morsy Yonis

Under the supervision of:

Dr. Anwar Abd Elnaser (Advisor)

Assistant Professor,
Institute of Global Health and Human Ecology (IGHHE)
The American University of Cairo

Spring 2022

The American University in Cairo

School of Science and Engineering

Finding biomarkers from long-non coding RNAs in serum of patients with non-alcoholic fatty liver disease

Thesis Submitted by

Nouran Samy Ahmed Morsy Yonis

To the Institute of Global Health and Human Ecology

Spring 2022

In partial fulfillment of the requirements for

the degree of Master of Science in Global Public Health and Human Ecology

Has been approved by

_____ _____ _____

Dept. Chair/ Director Date Dean Date

DEDICATION

This thesis work is dedicated
to my Grandmother, whom I
wish she was here to share
another success with her.
my father, Samy Yonis
my mother, Nahla Mohamed
and my beloved sisters, Rana and
Hadeer for their love and support

ACKNOWLEDGEMENTS

My sincere appreciation and gratitude go first and foremost, to God for allowing me to have such a fruitful experience at AUC.

I would like to express my gratitude to my advisor, Ass. Prof. Anwar Abdelnaser, for his continuous supervision, guidance, and support throughout my thesis. I would like to extend my special thanks to Professor Hassan El-Fawal, Founder of the Institute of Global Health and Human Ecology at AUC, whom without I would have never been a part of this place. His valuable feedback enriched my knowledge, and guided me in choosing the right direction to complete my MPH degree successfully.

I would like to extend my gratitude my AUC professors, who taught me the academic skills and the basics of scientific research. A very special thanks goes to Dr.Mohamed Salama for his continuous support and motivation and Dr.Doaa for her insightful feedback that helped me expanding my skills in my area of research.

I'm deeply thankful to Dr.Abdelmonim and Dr.Yousra for providing us with the needed patient samples from the National Hepatologyand Tropical Medicine Research Institute (NHTMRI). I also would like to thank Prof. Nashwa El-Khazragy, at Clinical Pathology-Hematology Department, Faculty of Medicine, Ain Shams University, for providing us the healthy control samples necessary for my research.

I would like to thank Hana Abdelzaher and Areeg Dabbish for helping me with the practical experience and statistical analysis and patiently answering and explaining all of my questions. Special thanks to Ahmed Elghoneimy in performing the bioinformatic analysis of the project.

This project would not have been possible without the alliance of The American University in Cairo (AUC). Thank you to AUC for providing the research grant which aided in funding this project. Finally, I would like to thank everyone who contributed and helped with this project including staff and lab members at AUC.

Finally, I couldn't have achieved any progress without the support and motivation of my dear family for their continuous love, support and guidance throughout my master's journey. Thank

you Dad, Mum, my sisters Rana and Hadeer, my brother -in law Moataz, and my cutest loved niece, Laila. Without your prayers, I would have never reached that far. I would like to extend my gratefulness to my aunts and uncles, cousins and amazing friends whom support always pushed me to do better.

ABSTRACT

Nonalcoholic fatty liver disease (NAFLD) is becoming a serious global public health problem as it is a highly prevalent condition, that begins as simple liver steatosis and develops into an inflammatory steatosis form known as non-alcoholic steatohepatitis (NASH). NAFLD has been linked with multiple comorbidities, which in certain cases predisposes to mortality secondary to extrahepatic neoplasms and coronary diseases. On the other hand, NASH leaves patients at risk of developing hepatocellular carcinoma (HCC). The incidence of NAFLD-based HCC is steadily increasing and is projected to surpass the incidence of HCC from viral origins. Long non-coding RNAs (lncRNAs) have been recently used as biomarkers for predicting various diseases. This research study intended to use bioinformatics computational analysis of microarray and sequencing databases to identify a differentially expressed panel of lncRNAs in the serum of NAFLD patients relative to samples from healthy individuals. Five differentially expressed lncRNAs (SNHG17, H19, MEG3, DUBR, and PVT1) were obtained from the bioinformatic analysis. Real-time polymerase chain reaction (PCR) was used to measure the levels of expression of these lncRNAs in the serum of 20 healthy individuals, 62 NAFL patients, and 30 NASH patients. The results of this study showed a highly significant increase in serum concentration of H19, MEG3, DUBR, and PVT1 in NAFL patients compared to healthy individuals (P <0.001, P=0.034, P= 0.003, P <0.001 respectively). Similarly, the expression of H19, DUBR, and PVT1 was significantly altered in NASH patients compared to healthy individuals (P <0.001, P=0.055, P=0.004, and P < 0.001 respectively)**.** PVT1 had diagnostic power in differentiating NAFL from NASH with sensitivity and specificity (80% and 73.81 respectively).

Using a combined panel of four lncRNAs improved the overall sensitivity and specificity of NAFL detection to 73.81% and 100% respectively, and NASH diagnosis to 73.33% and 95% respectively. Moreover, a combined panel of H19 and PVT1 possess potential diagnostic power for early detection of NAFL and NASH with an accuracy of 98.39%, and 88% respectively. Further investigation is required to assess the therapeutic effects of these lncRNAs in NAFLD. In conclusion: Using a combined lncRNAs panel (H19, MEG3, DUBR, and PVT1) could serve as a non-invasive biomarker for early detection of NAFLD. A novel lncRNA panel (H19,

MEG3, DUBR, and PVT1) was introduced to establish a potentially definitive and minimally invasive diagnostic test that can detect the disease at an early stage of its progression.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ALT	Alanine aminotransferase
ANOVA	Analysis of variance
AST	Aspartate aminotransferase
AUC	Area under the curve
BMI	Body mass index
CBC	Complete blood count
cDNA	Complementary deoxyribonucleic acid
Ct	Cycle threshold
DE-lncRNAs	Differentially expressed lncRNAs
DM	Diabetes Mellitus
ds	Double stranded
GEO	Gene expression omnibus
GGT	Gamma-glutamyl transferase
GO	Gene Ontology
HCC	Hepatocellular carcinoma
IRB	Institutional Review Board
miRNA	Micro ribonucleic acid
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NCBI	National center for biotechnology information
NHTMRI	National Hepatology and Tropical Medicine Research Institute
NPV	Negative predictive value
ORF	Open reading frame
PCR	Polymerase chain reaction

PPV	Positive predictive value
qRT-PCR	Quantitative real-time polymerase chain reaction
RBCs	Red blood cells
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RQ	Relative quantification
RT	Reverse transcription
SNORD	Small nucleolar RNA
T2DM	Type 2 diabetes mellitus
US	Ultrasound
PAI-1	Plasminogen activator inhibitor 1

1. CHAPTER 1: LITERATURE REVIEW

1.1. Introduction

Nonalcoholic fatty liver disease (NAFLD) has become a serious global public health problem as it is one of the causes of fibrosis and cirrhosis that leads to hepatocellular carcinoma (Nishida & Goel, 2011). By 2030, the United States is expected to have over 100 million NAFLD patients, causing a greater prevalence of decompensated cirrhosis, hepatocellular carcinoma (HCC), and liver-related deaths (Estes, Razavi, Loomba, Younossi, & Sanyal, 2018). Hepatocellular carcinoma (HCC) is one of the big issues we have in Egypt and the sixth most common cancer worldwide (McGlynn, Petrick, & London, 2015). Moreover, NAFLD is recognized as the most common form of chronic liver disease in many parts of the world (Elzafir Elsheikh; Linda L Henry; Zobair M Younossi, 2013). The prevalence of NAFLD worldwide is about 25% (Araújo, Rosso, Bedogni, Tiribelli, & Bellentani, 2018). The highest prevalence of NAFLD reported in the Middle East was (31.79%), while Africa reported the lowest prevalence rate (13.48%) (Younossi et al., 2016b). The Prevalence among adults for non-alcoholic fatty liver disease (NAFLD) was reported that 31.6 % people suffered from steatosis (Tomah et al., 2021), yet according to a study conducted by (Alkassabany, Farghaly, & El-Ghitany, 2014), the prevalence rate among school children was 15.8 %. Massive efforts and extensive research from the government are currently taking place to detect NAFLD at an early stage and prevent its consequences.

Diagnosis of NAFLD has changed over the years to switch from invasive or imaging based-techniques to non-invasive methods. Although liver biopsy is the gold standard for determining the degree of fibrosis. However, because of the price, error variability, and risk for problems such as pain and bleeding, makes its use is limited. Furthermore, despite that ultrasonography is the first-line imaging technology in clinical practice, it has low sensitivity (Anstee, Targher, & Day, 2013) and is unable to distinguish between simple fatty liver and NASH. Therefore combining biomarkers could improve the sensitivity and precision of NAFLD detection.

Previous research focused on the interpretation of coding genes, even though the non-coding regions make up around 99% of the human genome. Long noncoding RNAs (lncRNAs),

a group of RNA molecules longer than 200 nucleotides that are unable to be translated into proteins, have attracted great attention in recent years, especially for their roles in liver diseases. (Gangqing Hu, 2013). Researchers confirmed the involvement of lncRNAs in lipid-related diseases. Furthermore, lncRNAs may influence fatty acid accumulation in the liver by up- or down-regulating downstream molecules in fatty acid metabolism (Sookoian et al., 2017a). Many studies addressed lncRNAs expression in NAFLD. Thus, Studying the long non-coding RNA in non-alcoholic fatty liver disease will help in finding better biomarkers. Such biomarkers will help in early diagnosis and NAFLD prevention, thus improving the NAFLD patient quality of life.

The current research hypothesized that computationally assigned lncRNAs possess differential expression between NAFLD patients and healthy subjects. The lncRNAs panel can predict NAFL and NASH in patients using minimally invasive serum samples. Thus, the current study suggested a novel lncRNAs panel that potentially acts as a prognostic biomarker for NAFLD early detection.

1.2. Literature review

1.2.1. Non-alcoholic fatty liver disease (NAFLD)

1.2.1.1. Definition

Non-alcoholic fatty liver disease (NAFLD) is characterized by an accumulation of fat that reaches 5% of the liver weight and is associated with insulin resistance and metabolic syndrome (Dongiovanni & Valenti, 2017). NAFLD is the accumulation of excess fat in liver cells that is unrelated to alcohol use. The presence of fat in the liver is normal. A fatty liver, on the other hand, is defined as one in which fat accounts for more than 5–10% of the weight of the liver (steatosis). Non-alcoholic fatty liver (NAFL) and non-alcoholic steatohepatitis (NASH) are two types of NAFLD (Cobbina & Akhlaghi, 2017).

Nonalcoholic steatohepatitis (NASH) is a more severe type of NAFLD. The liver swells and becomes damaged as a result of NASH (Andre Paquin, 2021). The NAFLD spectrum ranges from non-alcoholic fatty liver (NAFL), nonalcoholic steatohepatitis (NASH), cirrhosis, and hepatocellular carcinoma. NAFL is the presence of steatosis with no evidence of hepatocellular injury, while NASH is presence of hepatic steatosis, inflammation and scarring. The NAFL and NASH are both reversible , once they develop to cirrhosis and hepatocellular carcinoma, they are irreversible (Sivell, 2019).

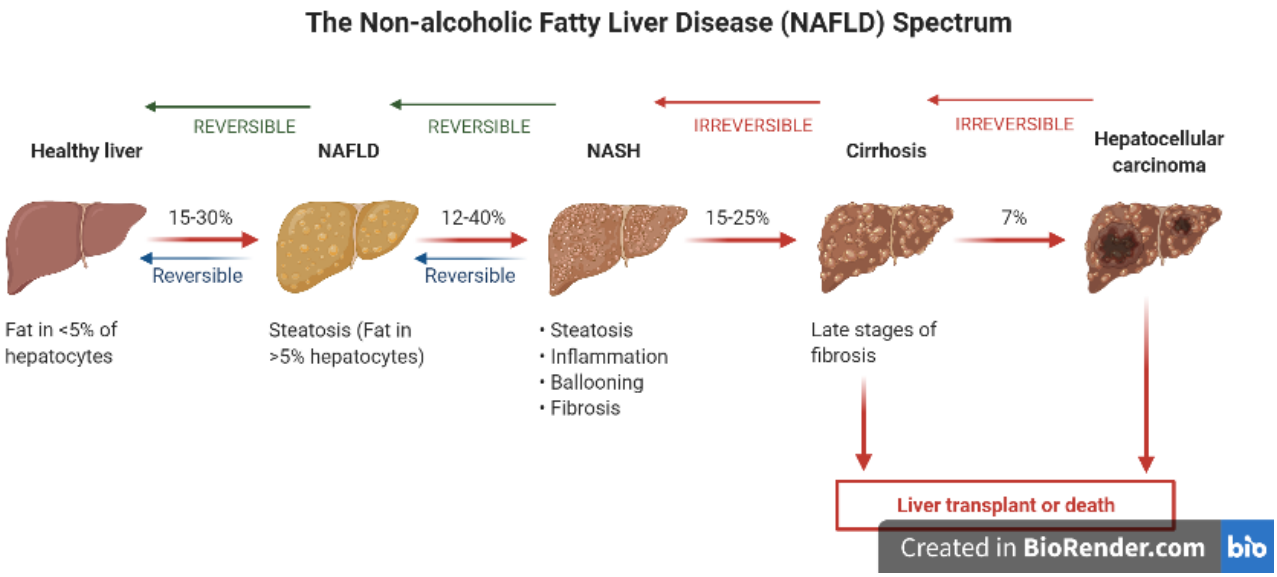


Figure 1. Schematic diagram showing the non-alcoholic fatty liver disease spectrum. Fat deposition in the hepatocytes causes a healthy liver to become fatty, followed by fibrotic liver, in which the nature of the liver cells begins to change into connective tissues, and finally cirrhotic liver, in which the liver begins to restructure and vascular systems are formed with the formation of necrotic areas. Finally, hepatocytes are converted into malignant hepatocytes, resulting in liver carcinogenesis. Adapted from (Turchinovich, Baranova, Drapkina, & Tonevitsky, 2018) Created with BioRender.com.

1.2.1.2. Pathogenesis

The exact pathogenesis of NAFLD is not clearly defined. The imbalance between the uptake and synthesis of fatty acids leads to hepatic fat accumulation. Esterification in the form of triglycerides (TG) is the best way to store fatty acids in the liver. Excess hepatocellular triglycerides are caused by dietary fatty acids, increased peripheral lipolysis, and elevated hepatic *de novo* lipogenesis (DNL). Mitochondrial oxidation and reduction of VLDL play a role in hepatic fat accumulation. Thus, TG accumulation causes second insults, cellular damage, activation of fibrogenesis, and hepatic inflammation. However other organs such as muscle,

intestine, and adipose tissue are also involved in the pathogenesis of NAFLD (Dongiovanni & Valenti, 2017).

Non fatty liver disease is associated with metabolic syndrome symptoms such as diabetes and is one of the hepatic manifestations of obesity (Baffy, Brunt, & Caldwell, 2012). Many extrahepatic morbidities are correlated with NAFLD such as diabetes mellitus, coronary, and chronic kidney diseases (Seko, Yamaguchi, & Itoh, 2018).

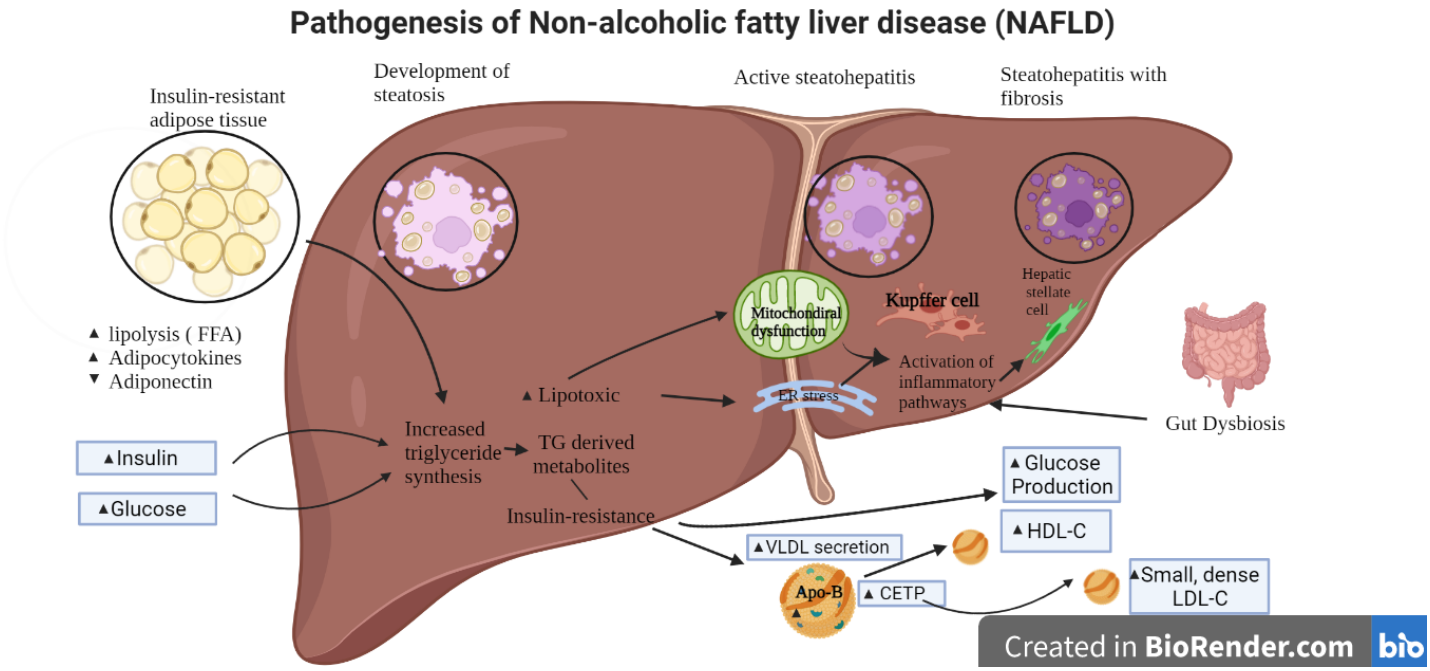


Figure 2. Pathogenesis of NAFLD. The imbalance between the uptake and synthesis of fatty acids leads to hepatic fat accumulation. Thus, causing increased peripheral lipolysis and elevated hepatic de novo lipogenesis (DNL). Mitochondrial oxidation and reduction of VLDL play a role in hepatic fat accumulation. Thus, TG accumulation causes fibrogenesis and hepatic inflammation. Adapted from (Akshintala, Chugh, Amer, & Cusi, 2019). Created with BioRender.com.

1.2.1.3. Epidemiology

A previous study conducted in the United States on 105 patients diagnosed with HCC revealed that 13% of patients with cryptogenic cirrhosis –related HCC are due to NAFLD (Marrero et al.,

2002). Furthermore, a study conducted by (Ratziu et al., 2002) found that obesity–related cirrhosis leads to severe liver disease which led to hepatocellular carcinoma and death.
NAFLD has become the most prevalent chronic liver disease worldwide in the last two decades, affecting an estimated 25.24% of the world's population, with the greatest incidence rates in the Middle East and South America (Younossi et al., 2016b). The prevalence in continents varied from one to another; for example, Africa had a prevalence of 13.5%, Europe 23.7%, North America 24.1%, Asia 27.4%, South America 30.5%, and the Middle East 31.8% (Rinella & Charlton, 2016).
NASH is associated with obesity, type 2 diabetes (T2D), and the male gender. However, because the histological examination is required to establish NASH, the real incidence of NASH in various populations remains uncertain (G Seyda Seydel et al.,2016). It's also worth noting that NAFLD progresses slowly, taking decades to develop. The actual prevalence trends are unknown due to the disease's silent nature. However, given the obesity and T2D pandemics, as well as the rising relative number of liver transplants for NAFL/NASH patients, the prevalence of NAFL and NASH is anticipated to have grown in recent decades and is predicted to continue to rise (Marchesini et al., 2016). The incidence of NAFLD in China is linked to the country's growing obesity rate (the prevalence from approximately 2% in 2000 to 7% in 2014) (Mitra, De, & Chowdhury, 2020; Zhou et al., 2020) comparing to Egypt which is 31.6% where 1 in 3 people had steatosis, and 1 in 20 had moderate-to-advanced fibrosis according to a study conducted by (Tomah et al., 2021).

1.2.2. NAFLD risk factors

1.2.2.1. Environmental risk factors

Exposures to endocrine-disrupting compounds (especially during early life), such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), polychlorinated biphenyls (PCBs), benzo[a]pyrene, bisphenol A (BPA), and phthalates, have been linked to the development of fatty liver disease through mechanisms such as binding to nuclear hormone receptors and epigenetic alterations (Lindsey S Treviño, 2018). Obesity, which is a risk factor for NAFLD, has been linked to many of these chemicals, including BPA and phthalates (Foulds, Treviño, York, & Walker, 2017).

Heavy metals, trihalomethanes, methyl tertiary- butyl ether, and selenium were associated with NAFLD (VoPham, 2019). Other environmental factors that impact NAFLD development include gut microbiota, alcohol consumption, toxic compounds, and drugs (Arciello, Gori, & Balsano, 2013; Haque & Barritt, 2016; Ihunnah, Jiang, & Xie, 2011; Marchesini et al., 2016).

1.2.2.2. Genetic risk factors

The prevalence of NAFLD varies widely over the world (Rinella & Charlton, 2016). Furthermore, regardless of BMI, insulin resistance, alcohol intake, or medication usage, the prevalence of hepatosteatosis differs dramatically by ethnicity (Hispanics, whites, and blacks, respectively) in the United States (Kahali, Halligan, & Speliotes, 2015). Genetic differences account for the majority of the variance between ancestries. The heritable component of NAFLD is predicted to be 22–38% based on ancestry (Locke et al., 2015). In recent years, there has been a lot of research on the genetic variables that predispose to NAFLD. GWAS has identified the three best-verified genetic variants that affect NAFLD: patatin-like phospholipase domain-containing 3 (PNPLA3), transmembrane 6 superfamily member 2 E167K variant (TM6SF2) (Macaluso et al., 2015), and membrane-bound O-acyltransferase domain containing 7 (MBOAT7) (Macaluso et al., 2015). (MBOAT7) (Mancina et al., 2016).

The variation has been linked to higher ALT levels, imaging-based hepatosteatosis, and histologic NAFLD, such as NASH, fibrosis, and cirrhosis (Locke et al., 2015). GWAS has also discovered several more genes that may be linked to NAFLD which are PNPLA3, TM6SF2, and MBOAT7 are three of these genes (Dongiovanni, Romeo, & Valenti, 2015; Mancina et al., 2016; Sookoian & Pirola, 2011).

1.3. NAFLD diagnosis

NAFLD is asymptomatic until the advanced stage of the disease. Thus, the use of biomarkers as early predictors needs to be studied (Baranova & Younossi, 2008). A liver biopsy is a gold standard for determining the degree of fibrosis. However, because of the price, error variability, and risk for problems such as pain and bleeding makes its use is limited. As a result, doing a liver biopsy on every person with NAFLD, which affects nearly a quarter of the world's population, is impracticable (Younossi et al., 2016).Therefore, the need to have reliable biomarkers is crucial. A biomarker is defined as an indicator of a normal or indication of abnormality or disease. The currently available biomarkers are inadequate.

Several techniques are used to evaluate the diagnosis of NAFLD. The most commonly reliable used ones in Egypt and worldwide are described as follows.

1.3.1. Abnormal liver function tests

A recent study conducted in Egypt identified abnormal liver function tests such as (AST, ALT, ɣ-glutamyltransferase GGT, total bilirubin), kidney functions, and CBC among the most reliable tools for diagnosis. Other blood tests, serological tests, and viral markers can be used to exclude other diseases (Hemida, Haroun, Mahmoud, & Mohamed, 2021). AST and ALT are the most commonly used serum markers to exclude NAFLD. Patients with NASH have elevated ALT than patients with NAFL Meanwhile, relying only on these two serum markers can be common with other conditions (Pouwels et al., 2022).

1.3.2. Imaging techniques

Ultrasonography is thus the first-line imaging technology in clinical practice, despite its low sensitivity (Anstee et al., 2013). Ultrasonography of the liver is a generally safe and affordable procedure. The sensitivity and specificity in patients with at least mild steatosis are 89% - 91% and 82% - 93% respectively (Palmentieri et al., 2006).

In a more recent investigation employing cutting-edge technology, remarkable sensitivity (100%) for identifying mild to severe steatosis was observed. Furthermore, ultrasonography frequently fails to detect steatosis of lesser severity. When patients with mild steatosis grades were included, sensitivity declined from 91% to 64% (Saadeh et al., 2002).

As a result, ultrasonography underestimates the prevalence of fatty liver in prevalence studies. Furthermore, ultrasonography is unable to distinguish between simple fatty liver and NASH. It also lacks the ability to identify fibrosis (Saadeh et al., 2002).

Another imaging technique that can be used is transient elastography (TE), known as fibroscan is a vibration-controlled TE device. The fibroscan is widely used to assess the staging of steatosis based on a scale starting from grade 0. Other imaging techniques include CT, MRI, and magnetic resonance spectroscopy (MRS). Their sensitivity and specificity to detect steatosis aren’t very precise (Dorairaj, Sulaiman, Abu, & Murad, 2021).

1.3.4. Current biomarkers in NAFLD

The use of current biomarkers in the diagnosis of NAFLD is crucial to avoid cirrhosis, fibrosis, and hepatocellular carcinoma. The following (Table 1) retrieved from (Hydes, Brown, Hamid, Bateman, & Cuthbertson, 2021) summarizes some of the current biomarkers used and their accuracy in the diagnosis of NAFL and NASH.

Table 1. Current biomarkers used for NAFLD diagnosis.

NAFL biomarkers	
Biomarker	**Accuracy**
Fatty Liver Index	Sensitivity, 87%; specificity, 86%
Hepatic Steatosis Index	Sensitivity, 93.1%; specificity, 92.4%
NAFLD fatty liver fat score	Sensitivity, 86%; specificity, 71%
NAFL screening score	AUC = 0.83
Lipid accumulation product	For each log unit increase, odds ratio for steatosis = 4.28
Steatotest	Sensitivity, 0.91; specificity
Metabolomic test	AUC = 0.64
NASH biomarkers	
Biomarker	**Accuracy**
ALT	Sensitivity, 64%; specificity, 75%
AST	Sensitivity, 77%; specificity, 62%
CK-18 fragments	Sensitivity, 75%; specificity, 81%
Activated PAI-1	95% CI
oxNASH panel: age, BMI, AST, 13-hydroxyl-octadecadenoic acids, linoleic acid	Sensitivity, 81%: specificity, 97%
FGF-21	Sensitivity, 0.62; Specificity, 0.78; FGF-21

1.3. Drawbacks of current screening techniques

Most of the available screening techniques are specific in their use. Those based on scores or body functions are variable. In addition, the cost, availability, accuracy, and patient acceptability remain important concerns. Because of their indefinite accuracy, make them less reliable to be used in clinical practice. For example, the image tools are useful to detect NAFLD, yet they are not optimized for diagnosing patients who might have type 2 diabetes (T2D). Since T2D is a risk factor for NAFLD, it is important to assess the accuracy of these techniques in diabetic patients (Mantovani, Byrne, Bonora, & Targher, 2018). Consequently, none of these imaging-based techniques are precise, thus, the need to develop alternative noninvasive diagnostic tools is crucial. Long non-coding RNAs being used successfully previously can be a reliable approach to detecting early NAFLD.

1.5. Long non-coding RNAs (lncRNAs)

Long-noncoding RNAs are transcripts longer than 200 nucleotides without a lengthy protein-coding open reading frame (ORF) that regulate gene expression at epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels, and are involved in a variety of biological activities (Di Mauro et al., 2021). Furthermore, long non-coding RNAs (lncRNAs) have the potential to control a number of NAFLD-related processes while not producing any protein products. LncRNAs may control the cis- or trans-regulation of gene expression through chromatin remodeling, epigenetic regulation, and transcriptional and post-transcriptional controls. The lncRNAs can serve as sponges to compete for the binding of RNA molecules (miRNAs) to prevent their downstream actions, scaffolds or platforms to offer a space for molecular interactions, and decoys to prevent protein interactions (Di Mauro et al., 2021).

1.5.1. lncRNAs biogenesis

lncRNAs are RNA-type molecules with a 5′ methyl-cytosine cap and a 3′ poly(A) tail that are transcribed by RNA polymerase II (Pol II) (Y. Liu et al., 2021; X. Zhang, Hong, Chen, Xu, & Wang, 2019). LncRNAs are categorized into several distinct categories based on their various features. lncRNAs, for example, can be classified into five categories based on their genetic origins: sense, antisense, bidirectional, intronic, and intergenic as shown in Figure 3. LncRNAs are categorized into three categories based on their function: rRNA, tRNA, and cRNA.

Furthermore, lncRNAs may be classified as nuclear, cytoplasmic, and mitochondrial lncRNAs based on their subcellular location.

(Alessio, Bonadio, Buson, Chemello, & Cagnin, 2020; Y. Liu et al., 2021)

The canonical mode caps, polyadenylates, and splices are the vast majority of lncRNAs (Gourvest, Brousset, & Bousquet, 2019). They can also be processed by noncanonical processes such as ribonuclease P (RNase P) cleavage to create mature 3′ ends, capping by snoRNA-protein (snoRNP) complexes at their ends, and the formation of circular structures. For example, RNase P recognizes and cleaves tRNA-like structures at their 3′ ends, resulting in the development of mature 3′ ends for certain lncRNAs (e.g., MALAT1 and Menβ) (Wu, Yang, & Chen, 2017).

The conserved nuclear RNA family includes snoRNA as one of its members. snoRNAs play crucial roles during ribosome subunit maturation as guide RNAs can be capped by snoRNPs at both ends, leading to their enhanced stability (L. L. Chen, 2016).

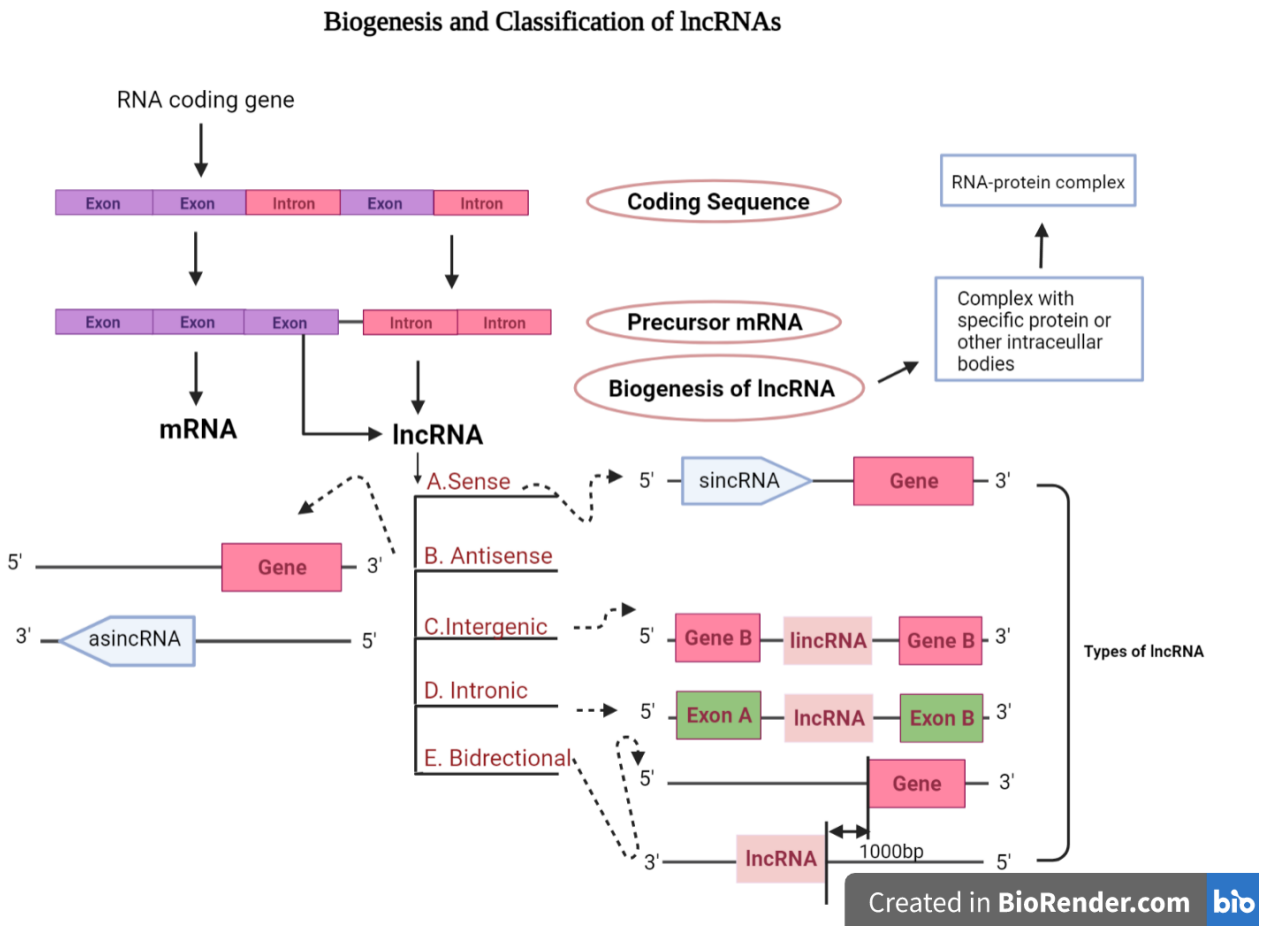


Figure 3. Biogenesis and Classification of lncRNAs. Types of lncRNAs include sense, antisense, intergenic, intronic, and bidirectional. Adapted from (Dhanoa, Sethi, Verma, Arora, & Mukhopadhyay, 2018; T. N. Zhang, Wang, Huang, & Gao, 2021). Created with BioRender.com.

1.5.2. lncRNAs as biomarkers in NAFLD

Researchers confirmed the involvement of lncRNAs in triglyceride, cholesterol, and lipoprotein metabolism. Furthermore, lncRNAs may influence fatty acid accumulation in the liver by up- or down-regulating downstream molecules in fatty acid metabolism (Sookoian et al., 2017a). Many studies addressed lncRNAs expression in NAFLD. Some lnRNAs are predominant in the hepatic stellate cells such as Alu-mediated p21 transcriptional regulator (*APTR*). Serum APTR levels were detected to be higher in cirrhotic patients in comparison to healthy individuals. Such results support that APTR can have a diagnostic value (Distefano & Gerhard, 2022; Sukowati, Cabral, Tiribelli,

& Pascut, 2021). Homeobox (*HOX*) transcript antisense RNA (*HOTAIR*) and Liver fibrosis-associated lncRNA 1 (*LFAR1*) , Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), and Nuclear-enriched abundant transcript 1 (NEAT1),showed promising results in microarray analysis to profile lncRNAs in CCl_4-treated mice. (Yu, Chen, Dong, & Zheng, 2017; Yu, Lu, et al., 2015)

The majority of the lncRNAs linked to NAFLD have been discovered in animal models of hepatic fibrosis, primarily CCl4-induced fibrosis. CCl4 is a typical approach for producing liver fibrosis, and it promotes HSC activation, dysregulated extracellular matrix formation and breakdown, and progressive hepatic fibrosis, just as fibrogenesis is linked to NAFLD in patients. However, CCl4 promotes hepatic inflammation, and the compound's intrinsic toxicity affects liver function in ways that do not mimic NAFLD fibrogenesis in people. Despite these limitations, the repeatability of findings in CCl4-treated animals and people with hepatic fibrosis in a variety of studies suggests a promising function for these lncRNAs in the diagnosis of NAFLD patients. (Distefano & Gerhard, 2022).

1.6. Rationale

The current techniques lack sensitivity and specificity in NAFLD detection. Therefore, the rationale of the current study is to identify a prognostic lncRNAs panel capable of distinguishing NAFLD patients in a more accurate, sensitive, and specific approach at an early stage of the disease.

1.7. Hypothesis

The computationally assigned lncRNAs possess a prognostic potential in detecting NAFLD in patients using minimally invasive serum samples.

1.8. Objectives and aims

The objective of this research is to identify a panel of lncRNAs with a differential expression between NAFLD patients and healthy individuals that could serve as a signature for early detection of NAFLD. While the aims of the current study are :

The specific aims of the current study are to :

1) Identify a panel of lncRNAs differentially expressed between NAFLD patients and healthy individuals through:

a) Reviewing the literature of the NCBI (national center for biotechnology information) database to determine the candidate lncRNAs.

b) Computational analysis of microarray and sequencing databases published in NCBI's Gene Expression Omnibus (GEO).

2) Collect serum samples of NAFLD patients as well as healthy subjects with subsequent isolation of lncRNAs from these samples.

3) Measure the expression levels of the identified lncRNAs from aim 1 using qPCR.

1.9. Novelty of this research

The novelty of this project is deciphered in its ability to answer an important question with regards to the capabilities of the identified panel of lncRNAs to use in the early detection of non-alcoholic fatty liver disease. It is anticipated that the identified panel of lncRNAs will have a differential expression pattern between NAFLD patients and healthy individuals that could serve as a signature for early detection at the steatosis stage which will aid in future diagnosis. Furthermore, the results of this thesis showed for the first time a combined panel of lncRNAs (H19, MEG3, DUBR, and PVT1), discovering the lncRNA PVT1 is an important biomarker that differentiates between NAFL and NASH patients.

2. CHAPTER 2: MATERIALS AND METHODS

2.1. Bioinformatic analysis

2.1.1. Data Acquisition

lncRNA microarray data from the Gene Expression Omnibus (GEO) database from two studies on NAFLD patients (GSE107231 and GSE72756) and two studies on NAFLD and High-fat diet (HFD) fed mice models (GSE108228 and GSE94790) were analyzed. The upregulated and downregulated lncRNAs of each set were determined with a cutoff log Fold Change (logFC) >1 and logFC <-1 respectively, and a p-value < 0.05. The probe sequences were mapped using the ensemble.org Blast tool which facilitated the annotation process as the arrays used nonconventional IDs. The intersection of the datasets was done using this Venn diagram tool (Bioinformatics & Evolutionary Genomics, n.d.). We then retrieved the sequences of selected lncRNAs from RNA Central (Sweeney et al., 2020) (Supplementary 1).

We also surveyed the literature for reported lncRNAs associated with NAFLD (Table 8). We selected a panel of 5 lncRNAs based on the intersection of the datasets and the availability of supporting literature. For those lncRNAs, the experimentally validated miRNA interactions were identified from Literature. The interaction domains of the lncRNA and the miRNA were identified using lncBase V2 (Maria D. Paraskevopoulou et al., 2016). The site of interaction was flanked with 100 nucleotides up and downstream for primer design.

2.1.2. SYBR® Green-based RT-qPCR primer design

SYBR® Green has been used in RT-qPCR to amplify lncRNA uc.372 (Guo et al., 2018). The primers were designed to flank the probes from the Human Microarray Datasets (GSE107231 and GSE72756). Using Primer3 (Lander, 2019). The primer design specs were optimized to yield the best performance in terms of efficiency and specificity when used with SYBR® Green RT-qPCR (Thornton & Basu, 2011). Primer dimerization was checked using an oligoanalyzer for self-dimers, heterodimers and hairpins (Integrated DNA Technologies, 2022).

In brief, primers were rejected if they 1) form hairpins at the Melting Temperature (Tm), or 2) have 3 base matches at the 3' end with $\Delta G < -3.5$ kcal/mol since these would be extended by the DNA polymerase and form primer dimers. The PCR amplicons were checked using UCSC in-silico PCR (Kent, n.d.) and primers were rejected if they have off-targets on the Dec. 2013 (GRCh38/hg38) version. GC content of the PCR amplicons was analyzed using GC Calculator

(Buddies, n.d.). In general, we selected primers that yield amplicons with GC content [40-60%]. Finally, the PCR amplicons were analyzed with UNAfold -- M.Zuker, 2003
for a Tm below 60 C ([Na] = 50 nM, [Mg] = 3 nM). Primers are listed in the Supplementary file.

2.2. Patients and samples

Blood samples have been collected from each patient after signing a written consent form (appendix 1,2) approved by the Institutional Review Board (IRB) of the American University in Cairo (case number **2020-2021-018**) in accordance with the Helsinki Declaration. The sampling procedure was held in collaboration with the liver transplantation clinic at the National Hepatology and Tropical Medicine Research Institute (NHTMRI) of Cairo University. We screened over 100 patients: this study included 62 Egyptian patients with NAFL, 30 Egyptian patients with NASH, and 20 healthy.

2.2.1. Inclusion and exclusion criteria

Complete Laboratory work data was collected from patients. The following were the intended battery of tests (Sookoian et al., 2017): Fasting plasma glucose, Total Lipid Profile: Serum total cholesterol, HDL, LDL, and triglycerides after overnight fasting, Liver function Tests: Serum Bilirubin, AST, and ALT, and albumin. Data regarding Abdominal Ultrasonography, Anthropometric measurement (Height (m), Weight (Kg), Waist circumference (cm)) were also collected along with laboratory work up to identify cases with metabolic syndrome. A diagnosis of metabolic syndrome is to be established per the NCEP Adult Treatment Panel III, according to which a diagnosis of metabolic syndrome is made if one had at least three of the following five components: Waist circumference > 102 cm in men or > 88 cm in women, Serum Fasting Triglycerides > 150 mg/dl, Serum Fasting HDL < 40 mg/dl in men and < 50 mg/dl in women, Blood pressure > 130/85 mmHg or receiving treatment, Fasting plasma glucose ≥ 100 mg/dl (Sharda, Yagnik, Soni, & Nigam, 2015).

All diseased patients were having NAFLD or NASH. Patients were excluded from the study if they are <18 years, those with missing fibro scan data. To determine metabolic syndrome components, non-fasting participants at the time of blood collection, immigrants, those with secondary causes of steatosis, participants with self-reported alcohol abuse (≥ 30 g alcohol daily for men and ≥ 20 g for women), total parenteral nutrition, hepatitis B and hepatitis C virus

infection, and the use of drugs (heparin therapy) known to precipitate steatosis. By using standard clinical and laboratory evaluation, as well as liver biopsy was used as a gold standard method to differentiate between NAFL and NASH patients, autoimmune liver disease, metabolic liver disease, Wilson's disease, and a-1-antitrypsin deficiency were likewise ruled out in all patients (Sookoian et al., 2017).

2.2.2. Sampling and serum preparation

Five mL of blood was withdrawn from each patient into a labeled disposable serum collection tube (global roll gel and clot activator tube). For complete clotting, blood samples were kept for one hour at room temperature (15–25°C), then samples were processed for serum separation following miRNeasy serum/plasma (miRNeasy M. Handbook, 2013).

2.3.RNA isolation

The miRNeasy serum /plasma extraction kit (Qiagen, Valencia, CA, USA) was used to extract total RNA, which the manufacturer's protocol (miRNeasy M. Handbook, 2013) requires the use of a QIAzol lysis reagent. After serum preparation, 5 volumes of QIAzol Lysis Reagent were added to the serum samples (see table 1 for guidelines). Mixed by vortexing or pipetting up and down. After leaving the homogenate on the benchtop at room temperature (15–25 °C) for 5 min, 1 volume of chloroform was added to the tube containing the homogenate (see table 1 for guidelines). The tube was vortexed vigorously for 15 s and the homogenate was placed on the benchtop at room temperature for 2–3 min.

The serum was then centrifuged using Microcentrifuge(s) (with rotor for 2 ml tubes) for centrifugation at 4°C and at room temperature (15–25°C) was used during the purification of the total RNA Including lncRNA from Serum for 15 minutes at 12,000 xg at 4˚C. The upper aqueous phase which contains the RNA was transferred to a new collection tube and 1.5 volumes of 100% ethanol were added and mixed in by pipetting up and down several times. Up to 700 µL of the sample was then pipetted into a RNeasy Mini spin column in a 2 ml collection tube and centrifuged at ≥8000 x *g* (≥10,000 rpm) for 15 s at room temperature (15–25°C). The flow-through was discarded and this step was repeated using the remainder of the samples. 700µL of

RWT buffer and 500µL of RPE buffer were added and centrifuged for 15 seconds at ≥8000 x *g* (≥10,000 rpm) to wash the column respectively. 500µL of 80% ethanol was added and centrifuged for 2 minutes at ≥8000 x *g* (≥10,000 rpm). The spin column was placed in a new 2 mL collection tube and centrifuged, the lid was open and left to air dry for 5 min. The spin column was then placed in a new 1.5 mL tube and 14µL of RNase-free water was added and centrifuged at full speed for 1 minute to elute the RNA. RNA concentration was determined using the NanoDrop2000 (Thermo scientific, USA). The extracted serum samples were then stored -at -80°C for further applications.

Table 2. Reagent volumes for various starting volumes of serum/plasma

Serum/plasma (µL)	QIAzol Lysis Reagent (µL)	Chloroform (µL)	Approx. volume of upper aqueous phase (µL)	100% Ethanol (µL)
≤50	250	50	150	225
100	500	100	300	450
200	1000	200	600	900

2.4.Reverse transcription and cDNA synthesis

For the lncRNA, polyadenylation and reverse transcription are performed in parallel in the same reaction. Polyadenylation and reverse transcription of lncRNA occur together in the same process. The oligo-dT primers have a 3' degenerate anchor and a universal tag sequence on the 5' end, which allows amplification of mature lncRNA in the real-time PCR. cDNA synthesis was performed using a miScript II RT kit (Qiagen, Hilden, Germany) according to the manufacturer's

instructions (R. lncRNA P. A. Handbook, 2014). 50 ng of RNA from each sample was diluted with a nuclease-free water variable according to the RNA volume taken (Table 2). Then, 2 µL of Buffer GE for a total volume of 10 µL. Then, add 10 µL of Reverse Transcription mix (Table 3) was added to each sample for a final reaction volume of 20 µL. The reaction contains 4 µL of 5x Buffer BC3, 1 µL of Control P2, 2 µL of Reverse Transcriptase mix, and 3 µL of Nuclease – Free Water. This method synthesized cDNA from mature long non-coding RNAs. The RT reaction was incubated at 37°C for 60 min followed by inactivation of the RT enzyme at 95°C for 5 min. The completed reaction was stored at -20°C to prevent cDNA degradation. For real-time PCR reaction, 2 µL were taken from the diluted cDNA.

Table 3. Genomic DNA elimination mix

Component	**Amount**
RNA	25 ng – 5 µg
Buffer GE	2 µL
Nuclease – Free Water	Variable
Total Volume	10 µL

Table 4. Reverse-transcription mix

Component	**Volume for 1 reaction**	**Volume for 6 reactions**	**Volume for 24 reactions**
5x Buffer BC3	4 µL	24 µL	96 µL
Control P2	1 µL	6 µL	24 µL
RE3 Reverse Transcriptase Mix	2 µL	12 µL	48 µL

Nuclease- Free Water	3 µL	18 µL	72 µL
Total Volume	10 µL	60 µL	240

2.5.Real-time PCR amplification of lncRNAs

For lncRNA, RT² lncRNA qPCR Assays and RT2 SYBR Green master mixes (Qiagen, Hilden, Germany) were used. A total reaction volume of 25µL containing 9.5µL of Nuclease-Free Water, 12.5µL of RT² SYBR Green master mix, 1µL of RT² lncRNA qPCR Assay, and 2µL of cDNA. All lncRNA primers were selected based on previously published reports and purchased from Qiagen.

A standard protocol designed by Qiagen was followed as in (Table 5). Fluorescence data collection was performed at the end of each elongation step. Rotor-Gene Q machine was used during Real-time PCR reactions.

Table 5. Cycling conditions

Cycles	Duration	Temperature	Comments
1	10 min	95°C	HotStart DNA Taq Polymerase is activated by this heating step.
40	15 s 30-40 s	95°C 55°C	Perform the fluorescence data collection. Different cycles need different lengths of time to

			detect fluorescent signals. Choose the appropriate time for the annealing step (55°C) for your cycler.
	30 s	72°C	

Specific forward lncRNAs primers are described in the (Table 6) below were custom-made available at Qiagen and purchased as forward primers.

Table 6. List of mature lncRNA primers used for qRT-PCR

lncRNAs	Expression Level	Forward Primer (5'→3')	Reverse Primer (5'→3')	Support
PVT1	Up	ATCCGTGTCT GGGAGAAACC	CCCCACAGTC ATACCCGTAA	Literature and Shared among all datasets analyzed (GSE107231, GSE72756, GSE108228 and GSE94790)
H19	Up	ACTCACGCAC ACTCGTACTG	CAAAGCCTCC ACGACTCTGT	Literature and Mice Data (GSE108228 and GSE94790)

MEG3	Down in Mice data and literature Up in some literature	ACCAAATCATTT CTGTGCCACTTC	CCTCATGTTTG CCTGGCATGG	Literature and Mice Data (GSE108228 and GSE94790)
DUBR	Down	AGTGTCTTCAGA ACCAAATGAGT	ACTGGAACAC TATCAGGGCT	Shared among all datasets analyzed (GSE107231, GSE72756, GSE108228 and GSE94790)
SNHG17	Down	ATCCTCACCTC TCCCTACGTC	TGAGAGGTGTT GCAGGAGAGT	Shared among all datasets analyzed (GSE107231, GSE72756, GSE108228 and GSE94790)
GAPDH	Housekeeping gene	GAACGGGAAGCTCAC TGG	GCCTGCTTCACCACCT TCT	Literature

2.6.Data analysis

The ΔΔCq method was recommended for data analysis (R. lncRNA P. A. Handbook, 2014). For each sample, a threshold cycle (CT) was calculated based on the time (measured by the number of PCR cycles) at which the reporter fluorescence emission increased beyond a threshold level (based on the background fluorescence of the system). The measurements for each sample were averaged to give an average CT value for each group, after removing the outliers. The CT of samples was observed up to 35 cycles. Results were expressed using the comparative CT method as described in Appendix A (R. lncRNA P. A. Handbook, 2014). Briefly, the ΔCT values were calculated in every sample for each gene of interest as follows: CT gene of interest − CT reference gene, with GAPDH as the reference gene for lncRNA. Calculation of relative changes in the expression level of one specific gene (ΔΔCT) was performed by subtraction of the ΔCT of the control (untreated cells) from the ΔCT of the corresponding treatment groups. The values and ranges given in different figures were determined as follows: $2^{-\Delta\Delta CT}$ represents the fold change in gene expression.

2.7.Statistical analysis

Statistical Package for the Social Science (SPSS) version 25 (IBM SPSS, Chicago, Il, USA) was utilized to analyze the data. The variables were expressed as the mean ± standard deviation (SD). Normality testing was performed using the Anderson-Darling test, D'Agostino and Person test, Shapiro-Wilk test, and Kolmogorov-Smirnov tests using GraphPad Prism (version 8.4.3). The Mann-Whitney U-test was used to analyze and compare the demographic and biochemical data. The Mann-Whitney test was chosen because the data did not have a normal distribution meaning the data was considered to be nonparametric. The analysis of variance (ANOVA) test was used to determine whether there was any significance in the expression of lncRNAs between each group of patients (controls, NAFL and NASH). The expression was considered to be significant if the probability was less than 0.05 ($p < 0.05$). Quantitative data were demonstrated as mean ± SEM, range (minimum–maximum), or number (percentages) as appropriate. For non-normally distributed data, values were analyzed using Mann-Whitney U (for comparison between two

groups), and Kruskal-Wallis H (for comparison among three or more groups), however normally distributed quantitative data was analyzed using one-way ANOVA (analysis of variance) to compare three or more groups. Analysis of qualitative data was performed using the Chi-square test. Spearman's rank correlation was used to study the inter-relation between target lncRNAs. Receiver operating characteristic (ROC) curve analysis was done to evaluate the diagnostic accuracy (sensitivity and specificity) of the 5 lncRNAs. Moreover, the ROC curve was used to calculate the area under the curve (AUC). Figures were designed using SPSS and GraphPad prism. All statistical tests were two-tailed, and P-value ≤ 0.05 was considered statistically significant. The Pearson correlation method was used to correlate lncRNA expression patterns within each group of patients. A minimum p-value < 0.05 was used to define a significant correlation between lncRNAs.

CHAPTER 3: RESULTS

2.8. Microarray bioinformatic analysis

Our analysis (Table7) revealed that human data in the two studies (GSE107231 and GSE72756) intersected at 5 upregulated and 21 downregulated lncRNAs. The mice data (GSE108228 and GSE94790) intersected at 8 upregulated and 28 downregulated lncRNAs. Altogether, the human and mice data intersected at 1 upregulated and 3 downregulated lncRNAs.

Table 7.The intersection of dysregulated lncRNAs from the human and Mice microarray data

Intersection	Number of lncRNAs	lncRNA(s)	Literature Data
Upregulated Human and Human (GSE107231 and GSE72756)	5	LINC01198, LINC00607, MAFTRR, PWRN1, LINC00907	
Downregulated Human and Human (GSE107231 and GSE72756)	21	DUXAP8, LINC00924, KCNMA1-AS1, LINC01255, PPFIA2-AS1, LINC00635, DCTN1-AS1, FZD10-AS1, AL591501.1, CA3-AS1, SPRY4-AS1, LINC00299, TEX41, SOX2-OT, DPP10-AS2, LINC02332, AC116565.1, AC024267.6, AC055854.1,	

		SLIT3-AS1, AC239809.3	
Upregulated Mice and Mice (GSE108228 and GSE94790)	8	GM32017, GM17494, SNHG18, 5430405H02RIK, BACH2OS, H19, GM4651, 9330158H04RIK	H19 is the first reported lncRNA to be associated with liver disease (He et al., 2014) and was also shown to induce steatosis in hepatocytes (H. Wang et al., 2020)
Downregulated Mice and Mice (GSE108228 and GSE94790)	28	9330159M07RIK, SNHG3, D630024D03RIK, A530083I20RIK, GM16028, 4930405D11RIK, 2810410L24RIK, RIAN, GM17206, E330020D12RIK, E330012B07RIK, AC160336.1, SNHG17, 2610016A17RIK, BC048559, GM50166, MEG3, GM17028, GM15651, 4930448E22RIK, 4930467D21RIK, 4933413J09RIK, A530058N18RIK, GM2200, GM17167, GM10874	MEG3 is downregulated in human fibrotic liver as well as CCl4-induced mice models (He et al., 2014), but was shown to be upregulated in the liver of NASH cirrhosis and liver fibrosis patients (Kim, Park, & Lee, 2020; L. Zhang, Yang, Trottier, Barbier, & Wang, 2017b)
Upregulated Human and Mice (GSE107231, GSE72756, GSE108228 and GSE94790)	1	PVT1	PVT1 is upregulated in patients with activated HSCs and fibrotic liver (R. Huang, Duan, Fan, Li, & Wang, 2019; Khalifa, Errafii, Al-Akl, & Arredouani,

			2020a; Lu, Luo, Wang, Ye, & Wang, 2017)
Downregulated Human and Mice (GSE107231, GSE72756, GSE108228 and GSE94790)	3	DUBR, SNHG17, SNHG14	SNHG14 was shown to be upregulated in HCC cell lines and to have a sponge effect on miR-4673 (Pu et al., 2019)

*Highlighted in Yellow are the lncRNAs supported by our analysis and literature

2.9. RNA sequencing bioinformatic analysis

Based on the bioinformatic analysis compiled in Table 8, in this current study, we have selected these five lncRNAs (SNHG17, H19, MEG3, DUBR, and PVT1) because of their potential promising effect in the early detection of non-alcoholic fatty liver disease in a more accurate, sensitive and specific approach (shown in Table 9).

Table 8. lncRNA Associated with NAFLD in Literature

lncRNA	Up	Down	Stage	Targets	Reference
LncARSR			NAFLD	YAP1 IRS2/AKT pathway	(R. Huang et al., 2019; Khalifa, Errafii, Al-Akl, & Arredouani, 2020b)
MEG3		*	NAFLD/NASH Activated HSCs	DNMT and TGFB1 miR-21	(He et al., 2014; P. Huang et al., 2019a; Khalifa et al., 2020b; Kim et al., 2020; L. Zhang, Yang, Trottier, Barbier, & Wang, 2017a)
NEAT1			NAFLD/NASH	ACTA2 and Col1a1	(R. Huang et al., 2019; Khalifa et al., 2020b)

MALAT-1 (Also known as NEAT2)			Fibrosis HCC	CXCL5	(R. Huang et al., 2019; Khalifa et al., 2020b; C. Li et al., 2015; Yu, Jiang, Chen, Dong, & Zheng, 2017)
lnRNA-CoX2			Fibrosis Cirrhosis HCC	PTGS2	(Khalifa et al., 2020b; Tang et al., 2017)
HOTAIR			HCC Fibrosis	PTEN	(Bian et al., 2017; R. Huang et al., 2019; Khalifa et al., 2020a; W. Li, Chen, Lin, & Huang, 2017)
APTR			HCC fibrosis	TGF-β1	(R. Huang et al., 2019; Khalifa et al., 2020b; Yu, Zheng, et al., 2015)
PVT1			HCC	miR-152	(R. Huang et al., 2019; Khalifa et al., 2020b; Lu et al., 2017)
B4GALT1-AS1			NAFLD	hnRNPA1 PI3K/Akt pathway mTOR / SREBP-1C pathway	(R. Huang et al., 2019; J. Wang et al., 2018)
FLRL2			NAFLD	Arntl-Sirt1 axis	(Y. Chen et al., 2019)
H19			NAFLD (in vitro)	MLXIPL miR-130a/PPARγ axis	(J. Liu, Tang, Wang, & Liu, 2019; H. Wang et al., 2020)
MRAK052686 (Mice)			Steatosis	Nrf2 (Implied but not confirmed)	(Yuan et al., 2015)
SRA	*		promoting hepatic steatosis	FoxO1 ATGL	(G. Chen et al., 2016)
SNHG14			HCC	miR-4673/SOCS1	(Pu et al., 2019)
Uc.372			NAFLD	miR-195 miR4668	(Jun Guo, et al., 2018)

*Highlighted in Yellow are the lncRNAs supported by our analysis and literature

* MEG3 is downregulated in CCl4-induced mice models, as well as human fibrotic liver (He et al., 2014) but is upregulated in the liver of NASH cirrhosis and liver fibrosis patients (Kim et al., 2020; L. Zhang et al., 2017b)

*SRA knockout revealed its functional rule in promoting hepatic steatosis by repressing ATGL. No info was provided on its expression level.

Table 9. Recommended Panel of lncRNAs

	Sequence	LncRNA name	Catalog no.
1	AATCATAGACATGTTGGAAGGAGCGTTCTATGGCCTGGATCTCCTGAAGCTACATTCAGT	SNHG17	330701
2	CTTCTGAATTTAATTTGCACTAAGTCATTTGCACTGGTTGGAGTTGTGGAGACGGCCTTG	H19	330001
3	AACATTCATCCTCCACAGCCACGGGGACACCCTGCACCTATTCCCACGGGACAGGCTGGA	MEG3	330701
4	TCACTTCGTGCTAGCATATGGGCAATCTCAATTTATTTCTAATAACTCCCTGTATCTTTC	DUBR	330701
5	GGCCTGGTCTCCATTATTTGAGATGAGTTACATCTTGGAGGTGAGGACGTGCCTCGTGGT	PVT1	330701
6	TTGAGGTCAATGAAGGGGTCGAAGGTGAAGGTCGGAGTCA	Housekeeping gene (GAPDH)	NM_002046

2.10. Study subjects and laboratory testing

2.10.1. Subjects were classified into the following groups

I) Group 1 (Healthy controls)

Serum of 20 normal healthy samples were donated by global research lab, age range from 38-62 years old, the median age was 49 years old, and mean age ± SD was 49.61 ± 5.392.

II) Group 2 (NAFL group)

Blood samples from 62 patients with NAFL were collected. Patients' ages ranged from 22-66 years old, the median age was 46 years old, and mean age ± SD was 46.63 ± 9.884.

III) Group 3 (NASH group)

Blood samples from 30 patients with NASH were collected. Patients' ages range from 23-65 years old, the median age was 42 years old, and mean age ± SD was 42.50 ± 9.339

2.10.2. Clinicopathological and demographic features of the study groups

Demographic and clinical data are compiled in (Table 10). There is a significant difference was observed in the gender distribution between the healthy control, NAFL, and NASH groups. Moreover, there is a statistical difference was observed in age distribution among the healthy individuals, and other groups (P = 0.0.005). Elevated levels of ALT showed highly a significant difference among the NASH groups (P < 0.001). Similarly, AST elevated levels were highly significant in NAFL and NASH groups relative to healthy individuals (P < 0.001). No statistical significance was detected in FBG among the two diseased groups upon comparison with healthy individuals (P =0.376).

Table 10. Clinicopathological and demographic features of the study population

Clinicopathological Features	No. of participants [n= 92]	Groups			Statistics	
		Control [n=20 (% within group)]	NAFL [n=42 (% within group)]	NASH n=30 (% within group)]	$X^{2(a)}$	*P*-value
Age						
Mean age (≤46) (n=18)	45	4 (22.2%)	21(51.2%)	20 (71.4%)	10.632	0.005
Mean age (>46)	42	14 (77.8%)	20 (48.8%)	8 (28.6%)		
Missing	5	2	1	2		
Gender						
Male	23	12 (66.7%)	3 (7.1%)	8 (27.6%)	23.363	<0.001
Female	66	6 (33.3%)	39 (92.9%)	21 (72.4%)		
Missing	3	2	0	1		
BMI						
ALT						
≤ 50 IU/L	63	18 (100%)	38 (100%)	7 (24.1%)	57.318	<0.001
> 50 IU/L	22	0	0	22 (75.9%)		
Missing	7	2	4	1		
AST						
≤ 50 IU/L	61	17 (94.4%)	33 (94.3%)	11 (37.9%)	31.306	<0.001
> 50 IU/L	21	1 (5.6%)	2 (5.7%)	18 (62.1%)		
Missing	10	2	7	1		
FBG						
≤ 99 mg/dl	24	11 (61.1%)	7 (70%)	6 (42.9%)	1.957	0.376
> 99 mg/dl	18	7 (38.9%)	3 (30%)	8 (57.1%)		
Missing	50	2	32	16		

Statistical significance is considered as *P*-value ≤ 0.05.
Statistical analysis was performed using ([a]) Chi-square test

2.11. LncRNAs serum signature in the study groups

The study design relied on the determination of the differential expression signature of the lncRNAs obtained from the bioinformatic analysis. The expression levels of the target lncRNAs were assessed using SYBR Green-based qPCR in the diseased groups compared to healthy individuals as a control group. GAPDH was used as a reference gene for the same sample to calculate ΔCt values. The relative expressions of the candidate lncRNAs were assessed using $2^{-\Delta\Delta Ct}$ method.

Fold changes of the DE-lncRNAs among the study groups were represented in (figure 3)The results of this study revealed that serum levels of the four candidate lncRNAs were differentially expressed in NAFL and NASH patients in comparison to healthy individuals with high statistical significance (*P*-value < 0.05) using Mann-Whitney U statistical test. However, only PVT1 serum level showed statistical significance upon comparing NAFL patients with those having NASH. The expression levels of H19, MEG3, DUBR, and PVT1 were significantly altered in NAFL patients relative to the healthy controls ($P < 0.001$, $P = 0.034$, $P = 0.003$, and $P < 0.001$ respectively). Whereas, H19, DUBR, and PVT1 showed differences in differential expression between NAFL and the control group ($P < 0.001$, $P = 0.004$ and $P < 0.001$ respectively).

Comparison of the mean rank (which represents the arithmetic average of the positions in the list, preferred to be used in non-parametric tests) of the fold change among the study groups (Tables 11-14) showed that the increase in the fold change of H19, MEG3, DUBR was compatible with the disease progression. Whereas, PVT1, showed a reduction in the mean expression in the NAFL and NASH patients. Only SNHG17 showed no statistical difference between NAFL and NASH patients in comparison with healthy groups.

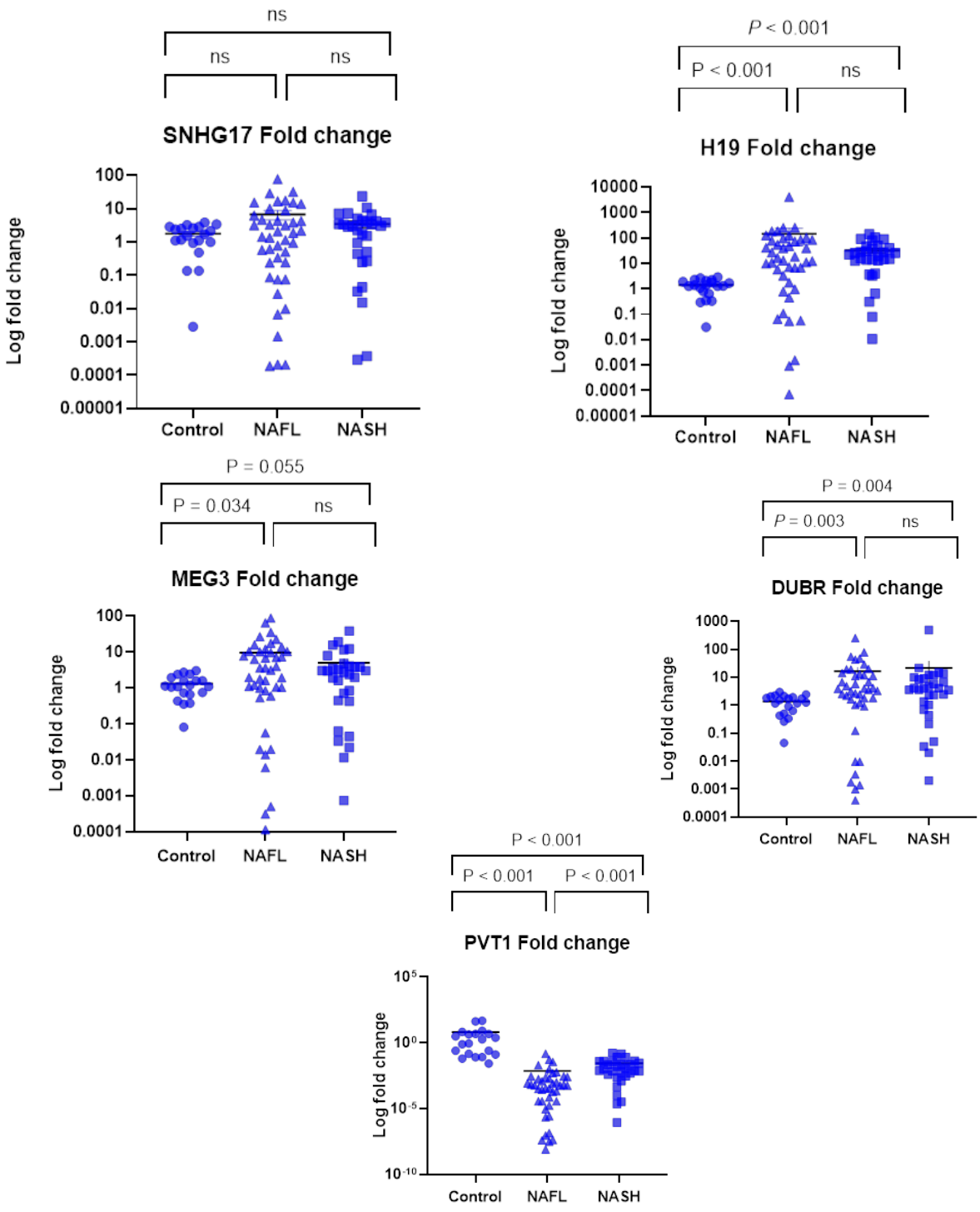


Figure 3. Fold change of the DE-lncRNAs in the study groups. Scatter dot plots demonstrate the fold change of serum expression of the target lncRNAs (SNHG17, H19, MEG3, DUBR, and PVT1) among the study groups. Y-axis represents the log of the fold change of each lncRNA; the X-axis shows the study groups.

Table 11. Fold change mean rank of target lncRNAs among the studied groups.

Target	Groups			Statistics	
	Control fold change mean rank (n= 20)	NAFL fold change mean rank (n= 42)	NASH fold change mean rank (n= 30)	X^2	P-value
SNHG17	42.25	47.26	48.27	0.672	0.715
H19	23.50	51.74	54.50	19.148	<0.001
MEG3	34.60	51.26	47.77	5.376	0.068
DUBR	29.60	51.60	50.63	10.260	0.006
PVT1	81.03	27.79	49.68	54.495	<0.001

Statistical significance is considered as *P*-value ≤ 0.05

Table 12. Fold change mean rank of NAFL and healthy individuals

Target	Groups		Statistics	
	Control fold change mean rank (n= 20)	NAFL fold change mean rank (n= 42)	X^2	P-value
SNHG17	30.10	32.17	392.000	0.673
H19	19.90	37.02	188.000	<0.001
MEG3	24.45	34.86	279.000	0.034
DUBR	21.78	36.13	225.500	0.003
PVT1	52.10	21.69	8.000	<0.001

Table 13. Fold change mean rank of NASH and healthy individuals

	Groups		Statistics	
Target	**Control fold change mean rank (n= 20)**	**NASH fold change mean rank (n= 30)**	**X^2**	**P-value**
SNHG17	22.65	27.40	243.000	0.259
H19	14.10	33.10	72.000	<0.001
MEG3	20.65	28.73	203.000	0.055
DUBR	18.33	30.28	156.500	0.004
PVT1	39.43	16.22	21.500	<0.001

Table 14. Fold change mean rank of NAFL and NASH patients

	Groups		Statistics	
Target	**NAFL fold change mean rank (n= 42)**	**NASH fold change mean rank (n= 30)**	**X^2**	**P-value**
SNHG17	36.60	36.37	626.000	0.964
H19	36.21	36.90	618.000	0.891
MEG3	37.90	34.53	571.000	0.500
DUBR	36.96	35.85	610.500	0.824
PVT1	27.60	48.97	256.000	<0.001

2.12.Receiver operating characteristic (ROC) analysis

Receiver operating characteristic (ROC) analyses were used to evaluate the diagnostic performance of the four potential lncRNAs. The ability of each prospective biomarker to detect the diseased groups was displayed as the area under the ROC curve (AUC), which was constructed using SPSS software version 25. The threshold value for the optimum sensitivity and specificity was determined using ROC analysis based on the relative quantification (RQ) values of the RNAs. True positive samples (sensitivity percent) and false positive samples (1 - specificity) of each RNA's RQ values were calculated at several cutoff points to determine appropriate cutoffs.

Furthermore, cutoff values for each of the RNAs were chosen. If the RQ was greater than or equal to this cutoff value, the sample was considered positive. The Chi-Square test was used to look at the distribution of positive cases of each lncRNA among the three groups. Calculation of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy across different comparative research groups to assess the diagnostic capability of DE-lnRNAs. For better diagnostic accuracy, ROC analysis was implemented for combined panels of statistically significant lncRNAs.

2.13.Diagnostic potential of the DE-lncRNAs

2.13.1.Diagnostic potential of the DE-lncRNAs in NAFL patients compared to healthy individuals

ROC analysis calculations were assessed for the candidate lncRNAs to discriminate NAFL patients from healthy controls (Figures 4-5 and Table 15). The AUC values were 0.776, 0.668, 0.732, and 0.990 corresponding to H19, MEG3, DUBR, and PVT1 respectively with high statistical significance (P-value < 0.001, P-value < 0.007, P-value < 0.006 , and P-value < 0.001). All of the targets showed high sensitivity (ranging from 97.62% to 61.90 %) and accuracy (ranging from 96.77% to 66.13 %) for the diagnosis of NAFL patients. Combining lncRNAs (H19, MEG3, DUBR, and PVT1) in which four tested positive, the whole panel is considered positive have sensitivity, specificity, and accuracy of 73.81%, 100%, and 82.26% respectively with (P-value < 0.001). Moreover, combining lncRNAs (H19, DUBR, and PVT1) in which three tested positive, the whole panel is considered positive and have sensitivity, specificity, and accuracy of 80.95%, 100%, and 100% respectively with (P-value < 0.001). Having a combined panel of the most

significant lncRNAs (H19 and PVT1); in which if 2 lnRNAs tested positive, the whole panel is considered positive; increased the sensitivity, specificity, and accuracy of detection to 100%, 95 %, and 98.39% respectively with high statistical significance (*P*-value < 0.001). Yet, using a combined panel of 2 lncRNAs (H19 and DUBR) had lower sensitivity, specificity, and accuracy of 80.95%, 65%, and 75.81% respectively with (*P*-value < 0.001).

2.13.2. Diagnostic potential of the DE-lncRNAs in NASH patients compared to healthy individuals

To identify NASH patients from healthy individuals, ROC curves were drawn for the candidate lncRNAs (Figures 5-6 and Table 16). The AUC values were 0.880, 0.662, 0.739, and 0.964 corresponding to H19, MEG3, DUBR, and PVT1, respectively with high statistical significance (*P*-value < 0.001, *P*-value = 0.004, *P*-value < 0.001, and *P*-value < 0.001). All of the targets showed high sensitivity (ranging from 100% to 66.67%) and accuracy (ranging from 92 % to 70%) for discrimination of NASH patients. A combined panel of 4 lncRNAs (H19, MEG3, DUBR, and PVT1) had sensitivity, specificity, and accuracy of 73.33%, 95%, and 82% respectively with (*P*-value < 0.001). Furthermore combined panel of 3 lnRNAs (H19, DUBR, and PVT1) improved overall sensitivity, specificity, and accuracy of detection to 88.67%, 95%, and 90% respectively with high statistical significance (*P*-value < 0.001). In addition, a combined panel of 2 lncRNAs (H19 and DUBR) had sensitivity, specificity, and accuracy of 88.67% , 90%, and 88% respectively. Upon combining only 2 lncRNAs (H19 and PVT1), the sensitivity increased to 100% , yet specificity was 70%, and accuracy is the same at 88%.

2.13.3. Diagnostic potential of the DE-lncRNAs in NAFL patients compared to NASH individuals

In a comparison between NAFL patients with NASH patients, AUC was calculated and only one lncRNA PVT1 had a statistically significant value (Figure 8 and Table 17). For PVT1 had AUC equal to 0.797 with high statistical significance (*P*-value < 0.001), and sensitivity, specificity, accuracy (80%, 73.81%, and 67.06) respectively.

2.14.Correlation between the studied lncRNAs

Spearman's correlation test was performed to investigate the correlation between the fold change of expression of each lncRNA and the other lncRNAs. A positive correlation was recorded between the expression of all lncRNAs among the study groups, with high statistical significance ($P < 0.001$) (Table 18). All lncRNAs (SNHG17, H19, MEG3, DUBR, and PVT1) showed a positive correlation. SNHG17 showed positive correlations with all other lncRNAs (H19, MEG3, DUBR, and PVT1). H19 showed a positive correlation with all lncRNAs except PVT1. Similarly , MEG3 and DUBR both showed positive correlations with all other lncRNAs except for PVT1. Consequentially, PVT1 only had a correlation with SNHG17. This table suggests that SNHG17 and PVT1 showed different behavior being the only two down regulators tested.

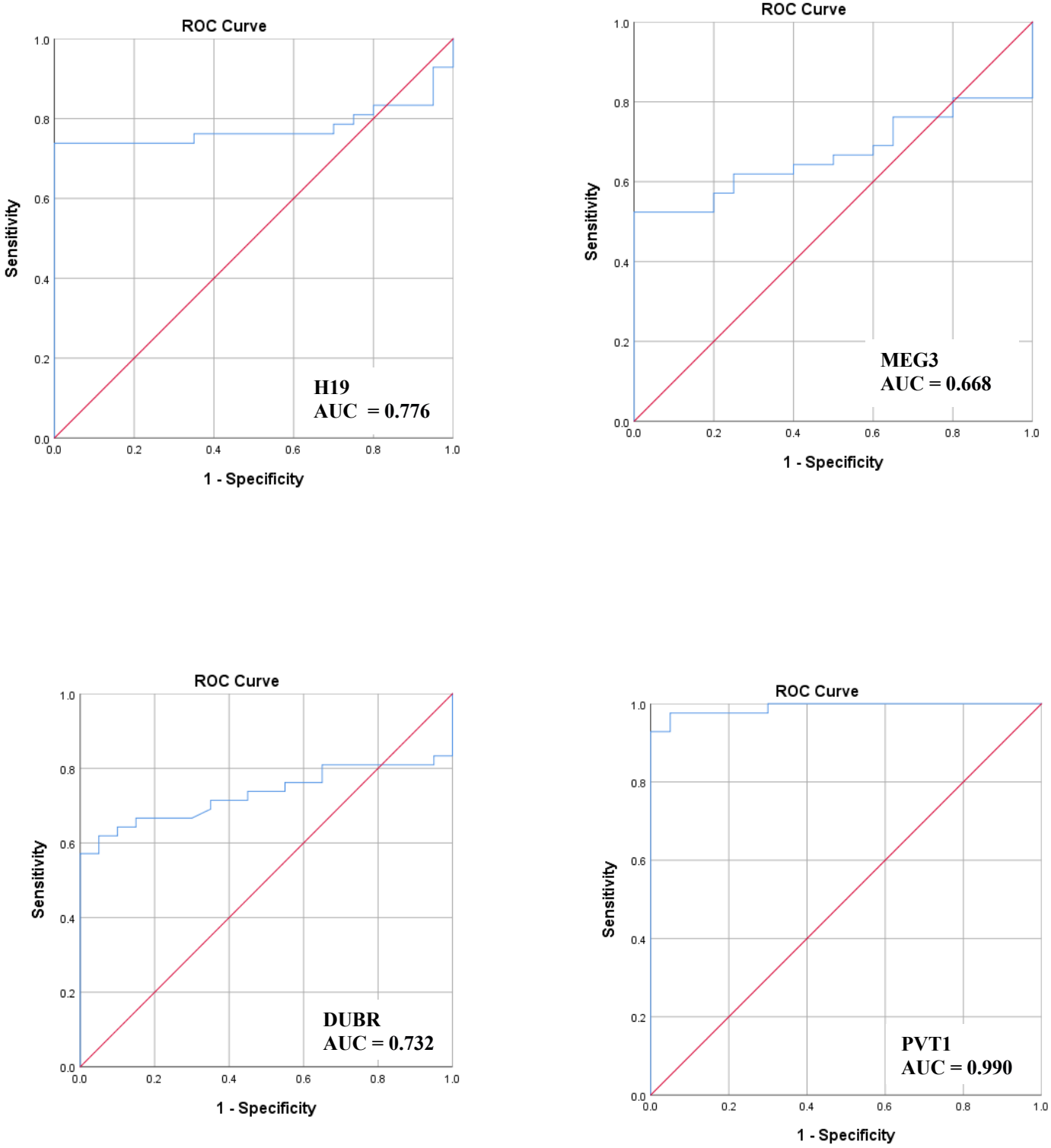
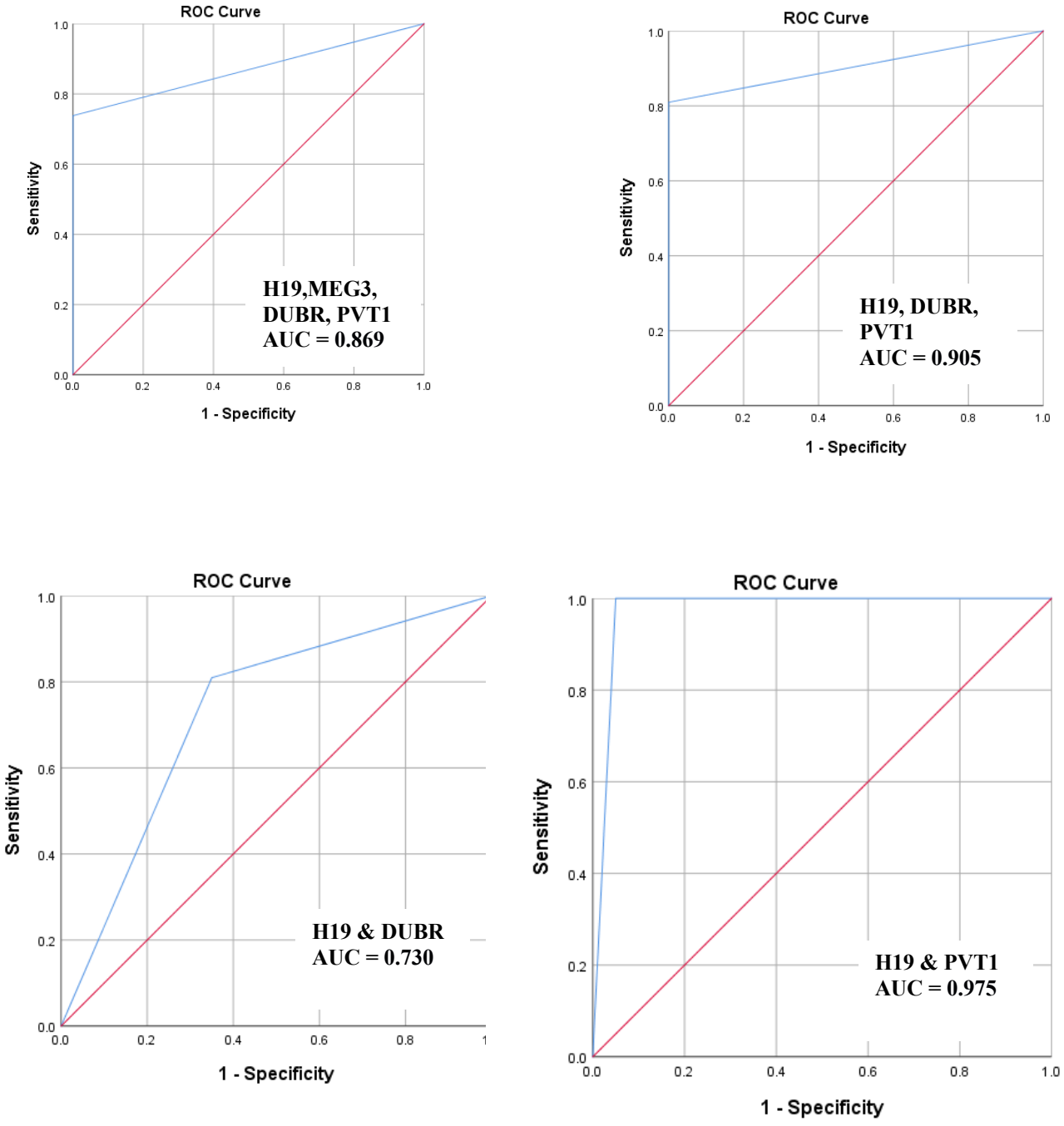


Figure 4. ROC curves and AUC for the diagnostic potential of the differentially expressed serum lncRNAs (H19, MEG3, DUBR, and PVT1) in the differentiation between NAFL patients and healthy individuals.



Figures 5. ROC curves and AUC for the diagnostic potential of the differentially expressed serum lncRNAs (H19, MEG3, DUBR and PVT1) in combined panel of NAFLD biomarkers

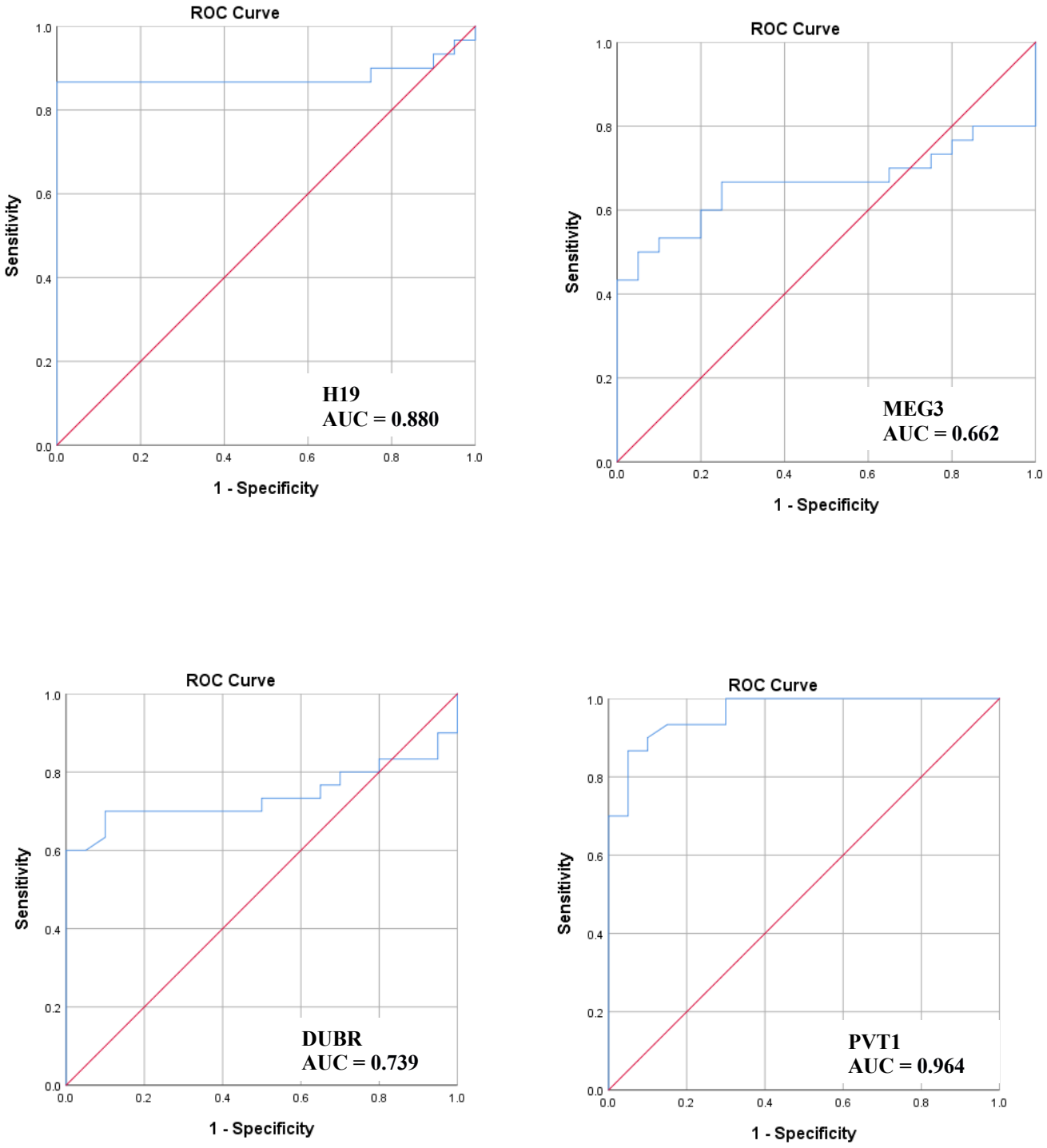
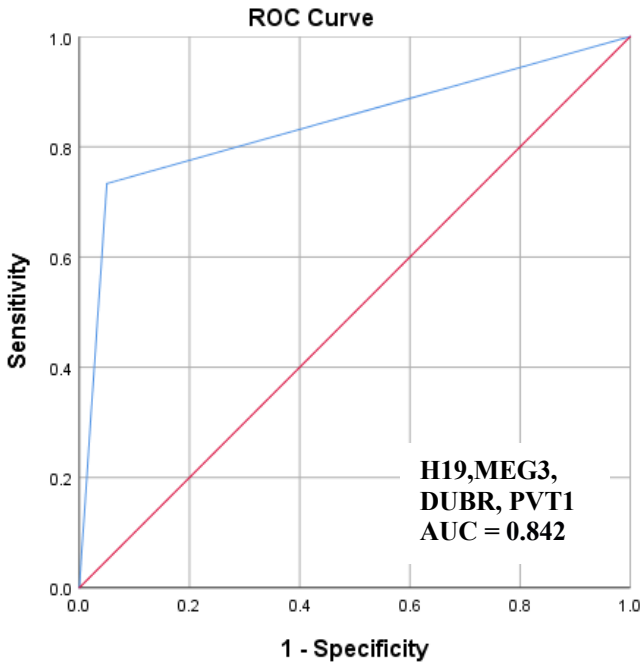
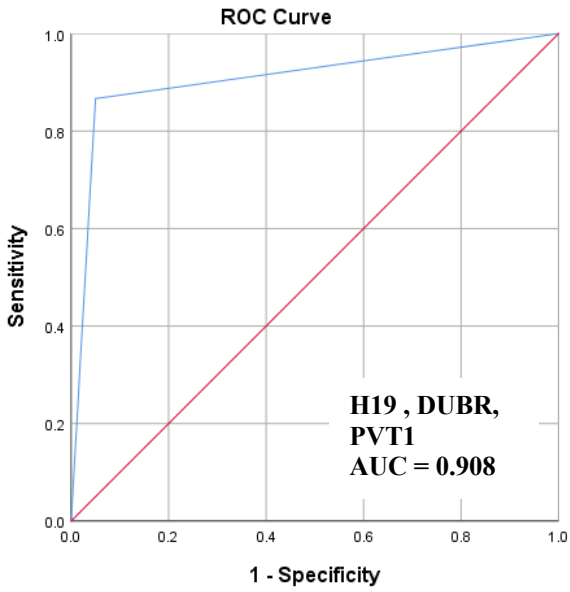
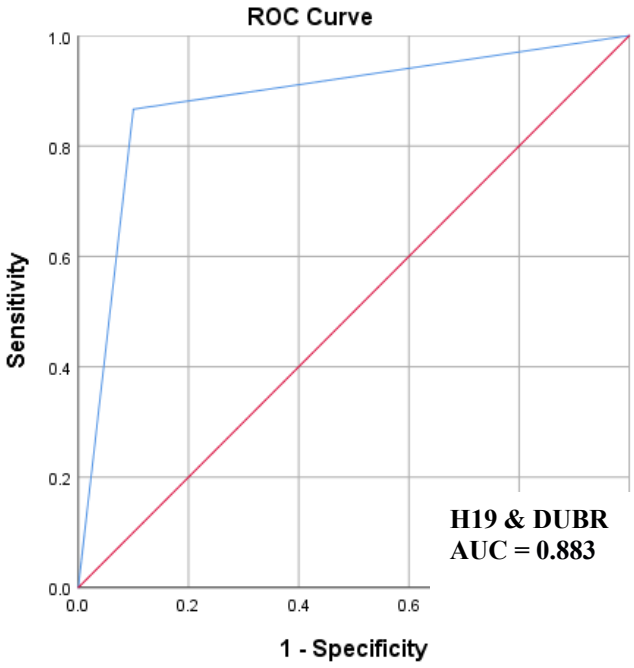
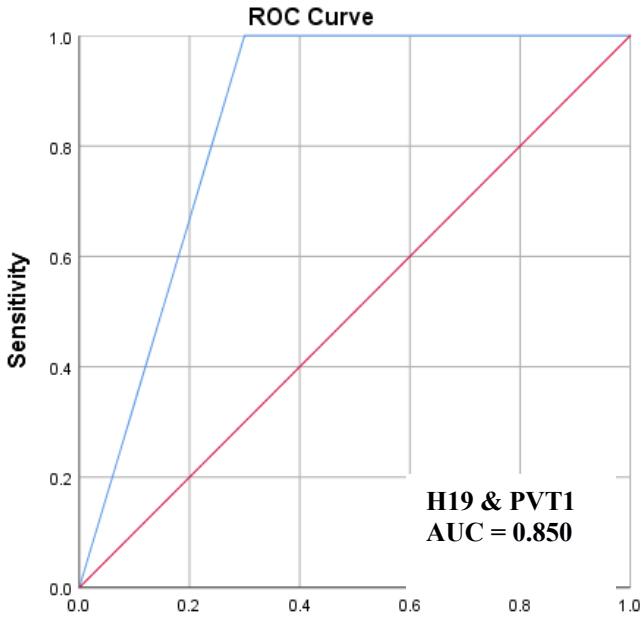


Figure 6. ROC curves and AUC for the diagnostic potential of the differentially expressed serum lncRNAs (H19, MEG3, DUBR and PVT1) in the differentiation between NASH patients and healthy individuals.









Figures 7. ROC curves and AUC for the diagnostic potential of the differentially expressed serum lncRNA (PVT1) in combined panel of NASH biomarkers.

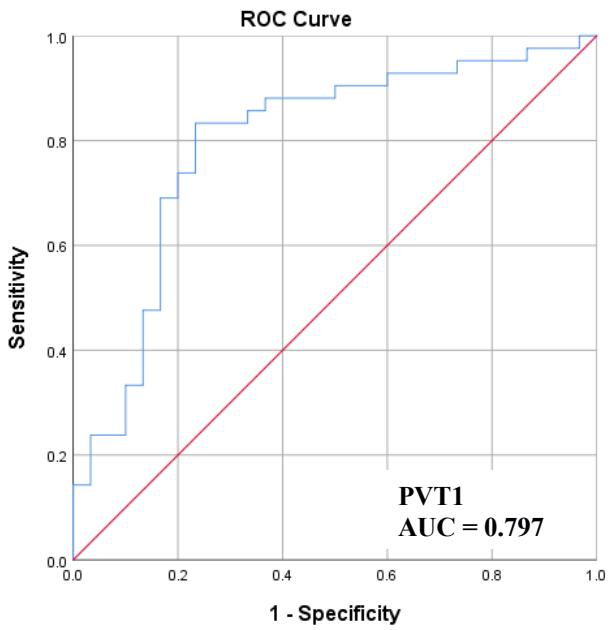


Figure 8. ROC curves and AUC for the diagnostic potential of the differentially expressed serum lncRNA (PVT1) in the differentiation between NAFL patients and NASH.

Table 15. ROC curve analysis of the investigated biomarkers in discriminating NAFL patients from healthy individuals

Target	AUC	SE	P-value	Cut-off	Sensitivity %	Specificity %	PPV	NPV	Accuracy	Chi- sq	P-value (2 sided)
NAFLD vs Control											
SNHG17	0.533	0.072	0.673 (N.S)								
H19	0.776	0.061	<0.001	3.05	73.81	100	100	64.52	82.26	29.524	<0.001
MEG3	0.668	0.067	0.034	1.60	61.90	75	83.87	48.39	66.13	7.381	0.007
DUBR	0.732	0.063	0.003	1.77	71.43	65	81.08	52	69.35	7.472	0.006
PVT1	0.990	0.009	<0.001	0.056	97.62	95	97.62	95	96.77	53.185	<0.001
Combined Panel (3 lncRNAs) (3/4) (H19 +MEG3+DUBR+PVT1)					73.81	100	100	64.52	82.26	29.524	<0.001
Combined panel (3 lncRNAs) (2/3) (H19 +DUBR+PVT1)					80.95	100	100	71.43	87.10	35.850	<0.001
Combined panel (2 lncRNAs) (1/2) (H19 + DUBR)					80.95	65	82.93	61.90	75.81	12.773	<0.001
Combined panel (2 lncRNAs) (1/2) (H19 + PVT1)					100	95	97.67	100	98.39	57.530	<0.001

Statistical significance is considered as *P*-value ≤ 0.05. ROC analysis was done on four statistically significantly candidate lncRNAs.

Table 16. ROC curve analysis of the investigated biomarkers in discriminating NASH patients from healthy individuals

Target	AUC	SE	P-value	Cut-off	Sensitivity %	Specificity %	PPV	NPV	Accuracy	Chi- sq	P-value (2 sided)
NASH vs Control											
SNHG17	0.595	0.081	0.259 (N.S)								
H19	0.880	0.057	<0.001	3.04	86.67	100	100	83.33	92	36.111	<0.001
MEG3	0.662	0.079	0.055	1.60	66.67	75	80	60	70	8.333	0.004
DUBR	0.739	0.074	0.004	2.27	70	90	91.30	66.67	78	17.391	<0.001
PVT1	0.964	0.022	<0.001	0.20	100	70	83.33	100	88	29.167	<0.001
Combined Panel (3 lncRNAs) (3/4) (H19 +MEG3+DUBR+PVT1)					73.33	95	95.65	70.37	82	22.558	<0.001
Combined panel (3 lncRNAs) (2/3) (H19 +DUBR+PVT1)					88.67	95	96.30	82.61	90	32.220	<0.001
Combined panel (2 lncRNAs) (1/2) (H19 + DUBR)					86.67	90	92.86	81.82	88	28.626	<0.001
Combined panel (2 lncRNAs) (1/2) (H19 + PVT1)					100	70	83.33	100	88	29.167	<0.001

Statistical significance is considered as *P-value* ≤ 0.05. ROC analysis was done on four statistically significantly candidate lncRNAs.

Table 17. ROC curve analysis of the investigated biomarkers in discriminating NAFLD patients from NASH

Target	AUC	SE	P-value	Cut-off	Sensitivity %	Specificity %	PPV	NPV	Accuracy	Chi- sq	P-value (2 sided)
NAFLD vs NASH											
PVT1	0.797	0.056	0.000	0.002	80	73.81	68.57	83.78	67.06	20.284	<0.001

Statistical significance is considered as *P-value* ≤ 0.05. ROC analysis was done on four statistically significantly candidate lncRNAs.

Table 18. Correlation between the target lncRNAs in the study groups

			RQ SNHG17	RQ H19	RQ MEG3	RQ DUBR	RQ PVT1
Spearman's rho	RQ SNHG17	Correlation Coefficient	1.000	.641**	.724**	.665**	.309**
		Sig. (2-tailed)	.	<0.001	<0.001	<0.001	.003
	RQ H19	Correlation Coefficient	.641**	1.000	.766**	.779**	.058
		Sig. (2-tailed)	<0.001	.	<0.001	<0.001	.580
	RQ MEG3	Correlation Coefficient	.724**	.766**	1.000	.786**	.140
		Sig. (2-tailed)	<0.001	<0.001	.	<0.001	.184
	RQ DUBR	Correlation Coefficient	.665**	.779**	.786**	1.000	.139
		Sig. (2-tailed)	<0.001	<0.001	<0.001	.	.187
	RQ PVT1	Correlation Coefficient	.309**	.058	.140	.139	1.000
		Sig. (2-tailed)	.003	.580	.184	.187	.

**. Correlation is significant at the 0.01 level (2-tailed).

CHAPTER 4: DISCUSSION

Non-alcoholic fatty liver disease (NAFLD) is becoming more common over the world. According to current estimates, 25% of the adult population, or one billion people globally, is impacted. Fatty liver appears to peak in males between 40 and 50 years of age and in females between 60 and 69 years of age, with prevalence declining somewhat in later (>70 years) cohorts. In addition, certain risk factors for the development of NAFLD, including hypertension, diabetes, hyperlipidemia, and obesity, are more common in older people. Diagnosis and management techniques for older persons can be difficult, and healthcare providers must take into account certain age-specific characteristics (Alqahtani & Schattenberg, 2021). The poor prognosis of NAFLD was the main motive for the researchers to determine new biomarkers than can aid in NAFLD early detection. Thus, this research study aimed to specify lncRNA panel to serve as a non-invasive biomarker for the prediction of NAFLD.

In this study, the patient's inclusion criteria relied on subjects shall undergo trans-abdominal ultrasonography performed by a single radiologist for evidence of fatty liver disease. The severity of steatosis will be recorded as mild, moderate, or severe fatty liver according to the findings of the bright liver, hepato-renal echo contrast, the blurring of vessels, and deep attenuation of ultrasound signal (Sharda et al., 2015). The choice of the healthy control group was sex-matched with NAFLD, as NAFLD is a female predominant disease (Giannitrapani et al., 2006). The results obtained from our bioinformatics analysis highlighted five lncRNAs found in microarray data from the Gene Expression Omnibus (GEO) database and literature. H19, MEG3, and PVT1 were upregulated, while SNHG17 and DUBR were downregulated in humans and mice. Previous research studies reported upregulation of H19 in mice which were comparable to our bioinformatics analysis results (H. Wang et al., 2020). Upregulation of MEG3 in human fibrotic liver as well as CCl4-induced mice models (He et al., 2014) and also was shown to be upregulated in the liver of NASH cirrhosis and liver fibrosis patients (Kim et al., 2020; L. Zhang et al., 2017b).

Based on our bioinformatics analysis, four of the selected lncRNAs were previously reported to have a strong association with NAFLD or HCC (Table 8). PVT1 which is the only upregulated

lncRNA in both human and mice microarray data (GSE107231, GSE72756, GSE108228 and GSE94790) has been reported to be upregulated in patients with activated HSCs and fibrotic liver (R. Huang et al., 2019; Lu et al., 2017). H19 which is upregulated in both of the mice datasets (GSE108228 and GSE94790) is the first reported lncRNA to be associated with liver disease (Ariel et al., 1998) and was also shown to induce steatosis in hepatocytes (H. Wang et al., 2020). MEG3 is downregulated in the mice datasets (GSE108228 and GSE94790) and was shown to be also downregulated in CCl4-induced mice models, as well as human fibrotic liver (He et al., 2014). In contrast to this finding, MEG3 was shown to be upregulated in the liver of NASH cirrhosis and liver fibrosis patients (Kim et al., 2020; L. Zhang et al., 2017a).

Furthermore, PVT1 was upregulated in Humans and Mice, and shown to be upregulated in patients with activated HSCs and fibrotic liver (R. Huang et al., 2019; Khalifa et al., 2020a; Lu et al., 2017). Moreover, analysis of the downregulation mechanisms revealed that SNG17 and DUBR were downregulated in humans and mice.

We believe, based on our analysis and literature survey, that the following lncRNAs can form a candidate panel of biomarkers for NAFLD (Table 9). The lncRNAs presented are shared among all or some of the analyzed datasets (GSE107231, GSE72756, GSE108228 and GSE94790), and some are supported by the literature.

Previous studies were in agreement with our results. (Y. Wang, Hylemon, & Zhou, 2021) They identified that H19 plays a role in diet-induced hepatic steatosis, yet no study has discussed the impact of H19 on both NAFL and NASH diseases. Our results also identified MEG3 as a potential biomarker for NAFLD which were similar to previous studies, yet we are the first to report on human samples (P. Huang et al., 2019b). There was not enough evidence in the literature about PVT1 and DUBR impact on NAFLD.

The results of this study showed highly a significant increase in serum concentration of H19, MEG3, DUBR, and PVT1 in NAFL patients compared to healthy individuals (P <0.001, P=0.034, P= 0.003, P <0.001 respectively). However, the expression of H19, MEG3, DUBR, and PVT1 were significantly altered in NASH patients compared to healthy individuals (P <0.001, P=0.055, P=0.004, and P < 0.001 respectively)**.** PVT1 had diagnostic power in differentiating NAFL from healthy individuals and NASH from

healthy individuals with sensitivity and specificity (80% and 73.81 respectively). Moreover, the combined panel was successfully used to assess the accuracy of distinguishing between NAFLD and healthy individuals. Having a combined panel of the most significant lncRNAs (H19 and PVT1); in which if 2 lnRNAs tested positive, the whole panel is considered positive; increased the sensitivity, specificity and accuracy of detection to 100%, 95 %, and 98.39% respectively with high statistical significance (*P*-value < 0.001) between NAFL and healthy individuals. Furthermore combined panel of 3 lnRNAs (H19, DUBR, and PVT1) improved overall sensitivity, specificity, and accuracy of detection to 88.67%, 95%, and 90% respectively with high statistical significance (*P*-value < 0.001) in NASH patients compared to healthy subjects.

Furthermore, lncRNA plays a role in insulin resistance in diabetic patients (Rashidmayvan, Sahebi, & Majid Ghayour-Mobarhan, 2022). Since diabetes is a risk factor for NAFLD, it was crucial to assess the NAFLD patients in our study. Our results revealed no statistical significance due to the lack of enough clinical data. Yet, previous studies proved there is a correlation between diabetes and NAFLD (Targher, Marchesini, & Byrne, 2016)

Previous studies were conducted in Egypt, one pilot study pilot study concluded that HSPD1/MMP14/ITGB1/miR-6881-5P/Lnc-SPARCL1-1:2 panel expression has potential in the differentiation and diagnosis of NAFLD (Albadawy et al., 2021). Another study concluded that PVT1 can be used as a diagnostic biomarker to differentiate patients with late NAFLD stages (Rashad et al., 2022). This study is in agreement with our study that PVT1 can be used to differentiate between NAFL and NASH.

To the best of our knowledge, we believe that the choice of the candidate lncRNAs within the panel makes it the first research to report a multifunctional tool for NAFLD early detection. The panel is composed of H19, MEG3, DUBR, and PVT1could act as a marker for early NAFLD detection. The use of the lncRNAs combined panel will facilitate and improve NAFLD diagnosis more than the conventional single biomarker approach.

3. CHAPTER 5: CONCLUSION AND FUTURE PERSPECTIVES

3.1. Conclusion

Around 25% of NAFLD patients are diagnosed globally. Early diagnosis remains a challenge despite the current use of invasive and non-invasive biomarkers. Yet, due to the poor accuracy, reliability, and cost of the current biomarkers, they are limited in use in clinical practice. Consequently, the identification of a specific non-invasive biomarker would enable early diagnosis of NAFLD, decrease the risks of fibrosis and cirrhosis, and permit the non-invasive monitoring and better therapeutic options.

The choice of lncRNAs as a reliable biomarker relied on the evidence that lncRNAs are sensitive predictors to physiological and pathological features of NAFLD. In this study, the serum differential expression of four lncRNAs (H19, MEG3, DUBR, and PVT1) were significantly overexpressed in NAFL and NASH patients relative to healthy individuals. Using a combined panel of four lncRNAs improved the overall sensitivity and specificity of NAFL detection to 73.81% and 100% respectively, and NASH diagnosis to 73.33% and 95% respectively. Moreover, a combined panel of H19 and PVT1 possess potential diagnostic power for early detection of NAFL and NASH with accuracy 98.39%, and 88% respectively. Further investigation is required to assess the therapeutic effects of these lncRNAs in NAFLD. In conclusion: Using a combined lncRNAs panel (H19, MEG3, DUBR, and PVT1) could serve as a non-invasive biomarker for an early detection for NAFLD. PVT1 is the first biomarker to be introduced that can differentiate between NAFL and NASH. This non-invasive technique surpasses the gold standard method, liver biopsy that has always been used by clinicians to differnetited NAFL from NASH patients.

3.2. Future perspectives

LncRNAs are cell- type and tissue specific, and can be released into circulating blood where they are stable, their use as new biomarkers for a variety of human illnesses, such as NAFLD, might have a therapeutic impact in the future. Despite the reliable use of lncRNA in cancer clinical diagnosis, only been studied as biomarkers of liver disease in a limited way.

Data from animal models and human patients, while still limited, provide convincing evidence for the involvement of functionally important lncRNAs in liver disorders. Given the large number of lncRNAs that have yet to be identified, much more research is needed to understand the molecular mechanisms by which lncRNAs contribute to liver diseases, the hepatic cell types and time points in disease pathogenesis when lncRNAs are activated or repressed, and the importance of lncRNA expression and molecular function in hepatic physiology and pathology.

3.3. Study limitations

Study limitations include the sample size wasn't high due to the difficulty to find patients diagnosed with NAFLD. Most patients discover liver disease in the late stages after fibrosis or cirrhosis development. Another reason, the availability of patients' data wasn't easily accessible due to the COVID shutdown, meeting with patients or obtaining the missing data wasn't applicable. The majority of samples were females which can be avoided in further research to compare between males and females prevalence.

REFERENCES

Akshintala, D., Chugh, R., Amer, F., & Cusi, K. (2019). Nonalcoholic Fatty Liver Disease: The Overlooked Complication of Type 2 Diabetes. *Endotext*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK544043/

Albadawy, R., Agwa, S. H. A., Khairy, E., Saad, M., El Touchy, N., Othman, M., & Matboli, M. (2021). Clinical significance of hspd1/mmp14/itgb1/mir-6881-5p/lnc-sparcl1-1:2 rna panel in nafld/nash diagnosis: Egyptian pilot study. *Biomedicines*, *9*(9), 1248. https://doi.org/10.3390/BIOMEDICINES9091248/S1

Alessio, E., Bonadio, R. S., Buson, L., Chemello, F., & Cagnin, S. (2020). A Single Cell but Many Different Transcripts: A Journey into the World of Long Non-Coding RNAs. *International Journal of Molecular Sciences 2020, Vol. 21, Page 302*, *21*(1), 302. https://doi.org/10.3390/IJMS21010302

Alkassabany, Y. M., Farghaly, A. G., & El-Ghitany, E. M. (2014). Prevalence, risk factors, and predictors of nonalcoholic fatty liver disease among schoolchildren: A hospital-based study in Alexandria, Egypt. *Arab Journal of Gastroenterology*, *15*(2), 76–81. https://doi.org/10.1016/j.ajg.2014.05.002

Alqahtani, S. A., & Schattenberg, J. M. (2021). NAFLD in the Elderly. *Clinical Interventions in Aging*, *16*, 1633–1649. https://doi.org/10.2147/CIA.S295524

Andre Paquin. (n.d.). Nonalcoholic Fatty Liver Disease (NAFLD) — American Liver Foundation. Retrieved January 21, 2022, from Liver Disease News and Information website: https://liverfoundation.org/for-patients/about-the-liver/diseases-of-the-liver/non-alcoholic-fatty-liver-disease/

Anstee, Q. M. (2011). Animal models in nonalcoholic steatohepatitis research: utility and clinical translation. *Liver International*, *31*(4), 440–442. https://doi.org/10.1111/j.1478-3231.2011.02463.x

Anstee, Q. M., Targher, G., & Day, C. P. (2013). Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. *Nature Reviews. Gastroenterology & Hepatology*, *10*(6), 330–344. https://doi.org/10.1038/NRGASTRO.2013.41

Araújo, A. R., Rosso, N., Bedogni, G., Tiribelli, C., & Bellentani, S. (2018). Global epidemiology of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: What we need in the future. *Liver International*, *38*, 47–51. https://doi.org/10.1111/liv.13643

Arciello, M., Gori, M., & Balsano, C. (2013). Mitochondrial dysfunctions and altered metals homeostasis: New weapons to counteract HCV-related oxidative stress. *Oxidative Medicine and Cellular Longevity*. https://doi.org/10.1155/2013/971024

Ariel, I., Miao, H. Q., Ji, X. R., Schneider, T., Roll, D., De Groot, N., … Ayesh, S. (1998). Imprinted H19 oncofetal RNA is a candidate tumour marker for hepatocellular carcinoma. *Molecular Pathology : MP*, *51*(1), 21–25. https://doi.org/10.1136/MP.51.1.21

Baffy, G., Brunt, E. M., & Caldwell, S. H. (2012, June 1). Hepatocellular carcinoma in non-alcoholic fatty liver disease: An emerging menace. *Journal of Hepatology*, Vol. 56, pp. 1384–1391. https://doi.org/10.1016/j.jhep.2011.10.027

Baranova, A., & Younossi, Z. M. (2008). The future is around the corner: Noninvasive diagnosis of progressive nonalcoholic steatohepatitis. *Hepatology (Baltimore, Md.)*, *47*(2), 373–375. https://doi.org/10.1002/HEP.22140

Bian, E. B., Wang, Y. Y., Yang, Y., Wu, B. M., Xu, T., Meng, X. M., … Li, J. (2017). Hotair facilitates hepatic stellate cells activation and fibrogenesis in the liver. *Biochimica et Biophysica Acta. Molecular Basis of Disease*, *1863*(3), 674–686. https://doi.org/10.1016/J.BBADIS.2016.12.009

Bioinformatics & Evolutionary Genomics. (n.d.). Draw Venn Diagram. Retrieved January 24, 2022, from http://bioinformatics.psb.ugent.be/webtools/Venn/

Buddies, S. (n.d.). Science Buddies: Genomics %G~C Content Calculator. Retrieved January 24, 2022, from https://www.sciencebuddies.org/science-fair-projects/references/genomics-g-c-content-calculator

Chen, G., Yu, D., Nian, X., Liu, J., Koenig, R. J., Xu, B., & Sheng, L. (2016). LncRNA SRA promotes hepatic steatosis through repressing the expression of adipose triglyceride lipase (ATGL). *Scientific Reports*, *6*. https://doi.org/10.1038/SREP35531

Chen, L. L. (2016). The biogenesis and emerging roles of circular RNAs. *Nature Reviews. Molecular Cell Biology*, *17*(4), 205–211. https://doi.org/10.1038/NRM.2015.32

Chen, Y., Chen, X., Gao, J., Xu, C., Xu, P., Li, Y., … Yu, C. (2019). Long noncoding RNA FLRL2 alleviated nonalcoholic fatty liver disease through Arntl-Sirt1 pathway. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *33*(10), 11411–11419. https://doi.org/10.1096/FJ.201900643RRR

Cobbina, E., & Akhlaghi, F. (2017). Non-Alcoholic Fatty Liver Disease (NAFLD) - Pathogenesis, Classification, and Effect on Drug Metabolizing Enzymes and Transporters. *Drug Metabolism Reviews*, *49*(2), 197. https://doi.org/10.1080/03602532.2017.1293683

Dhanoa, J. K., Sethi, R. S., Verma, R., Arora, J. S., & Mukhopadhyay, C. S. (2018). Long non-coding RNA: its evolutionary relics and biological implications in mammals: a review. *Journal of Animal Science and Technology 2018 60:1*, *60*(1), 1–10. https://doi.org/10.1186/S40781-018-0183-7

Di Mauro, S., Scamporrino, A., Filippello, A., Di Pino, A., Scicali, R., Malaguarnera, R., … Piro, S. (2021). Clinical and Molecular Biomarkers for Diagnosis and Staging of NAFLD. *International Journal of Molecular Sciences*, *22*(21), 11905. https://doi.org/10.3390/IJMS222111905

Distefano, J. K., & Gerhard, G. S. (2022). Long Noncoding RNAs and Human Liver Disease. *Https://Doi.Org/10.1146/Annurev-Pathol-042320-115255*, *17*, 1–21. https://doi.org/10.1146/ANNUREV-PATHOL-042320-115255

Dongiovanni, P., Romeo, S., & Valenti, L. (2015). Genetic Factors in the Pathogenesis of Nonalcoholic Fatty Liver and Steatohepatitis. *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/460190

Dongiovanni, P., & Valenti, L. (2017, July 16). A nutrigenomic approach to non-alcoholic fatty liver disease. *International Journal of Molecular Sciences*, Vol. 18. https://doi.org/10.3390/ijms18071534

Dorairaj, V., Sulaiman, S. A., Abu, N., & Murad, N. A. A. (2021). Nonalcoholic Fatty Liver Disease (NAFLD): Pathogenesis and Noninvasive Diagnosis. *Biomedicines 2022, Vol. 10, Page 15*, *10*(1), 15. https://doi.org/10.3390/BIOMEDICINES10010015

Elzafir Elsheikh; Linda L Henry; Zobair M Younossi. (n.d.). Management of Patients With Nonalcoholic Fatty Liver Disease. Retrieved March 28, 2020, from https://www.medscape.com/viewarticle/814180

Estes, C., Razavi, H., Loomba, R., Younossi, Z., & Sanyal, A. J. (2018). Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology*, *67*(1), 123–133. https://doi.org/10.1002/hep.29466

Foulds, C. E., Treviño, L. S., York, B., & Walker, C. L. (2017). Endocrine-disrupting chemicals and fatty liver disease. *Nature Reviews. Endocrinology*, *13*(8), 445–457. https://doi.org/10.1038/NRENDO.2017.42

G Seyda Seydel , Ozlem Kucukoglu , Akif Altinbasv , O Oguz Demir , Sezai Yilmaz , Hikmet Akkiz , Emrah Otan , Jan-Peter Sowa, A. C. (n.d.). Economic growth leads to increase of obesity and associated hepatocellular carcinoma in developing countries - PubMed. Retrieved January 21, 2022, from https://pubmed.ncbi.nlm.nih.gov/27493104/

Gangqing Hu, K. Z. (n.d.). Long Non-Coding RNA - an overview | ScienceDirect Topics. Retrieved March 28, 2020, from https://www.sciencedirect.com/topics/neuroscience/long-non-coding-rna

Giannitrapani, L., Soresi, M., La Spada, E., Cervello, M., D'Alessandro, N., & Montalto, G. (2006). Sex hormones and risk of liver tumor. *Annals of the New York Academy of Sciences*, *1089*, 228–236. https://doi.org/10.1196/annals.1386.044

Gourvest, M., Brousset, P., & Bousquet, M. (2019). Long Noncoding RNAs in Acute Myeloid Leukemia: Functional Characterization and Clinical Relevance. *Cancers*, *11*(11). https://doi.org/10.3390/CANCERS11111638

Guo, J., Fang, W., Sun, L., Lu, Y., Dou, L., Huang, X., … Li, J. (2018). Ultraconserved element uc.372 drives hepatic lipid accumulation by suppressing miR-195/miR4668 maturation. *Nature Communications 2018 9:1*, *9*(1), 1–15. https://doi.org/10.1038/s41467-018-03072-8

Handbook, miRNeasy M. (2013). miRNeasy Mini Handbook - QIAGEN. Retrieved April 22, 2022, from Qiagen website: https://www.qiagen.com/us/resources/resourcedetail?id=da6c8d17-58c4-411c-a334-bc1754876db3&lang=en

Handbook, R. lncRNA P. A. (2014). RT2 lncRNA PCR Array Handbook - QIAGEN. Retrieved April 22, 2022, from Qiagen website: https://www.qiagen.com/us/resources/resourcedetail?id=d9e189b0-6379-40c8-9f4e-f1c86582dc3d&lang=en

Haque, T. R., & Barritt, A. S. (2016). Intestinal microbiota in liver disease. *Best Practice & Research. Clinical Gastroenterology*, *30*(1), 133–142. https://doi.org/10.1016/J.BPG.2016.02.004

He, Y., Wu, Y. ting, Huang, C., Meng, X. M., Ma, T. tao, Wu, B. M., … Li, J. (2014).

Inhibitory effects of long noncoding RNA MEG3 on hepatic stellate cells activation and liver fibrogenesis. *Biochimica et Biophysica Acta*, *1842*(11), 2204–2215. https://doi.org/10.1016/J.BBADIS.2014.08.015

Hemida, K., Haroun, H., Mahmoud, M., & Mohamed, I. E. (2021). Non-invasive Methods in Diagnosis of Nonalcoholic Fatty Liver Disease. *The Egyptian Journal of Hospital Medicine*, *83*, 1038–1042. Retrieved from http://creativecommons.org/licenses/by/4.0/

Huang, P., Huang, F. zhou, Liu, H. zheng, Zhang, T. yi, Yang, M. shi, & Sun, C. zheng. (2019a). LncRNA MEG3 functions as a ceRNA in regulating hepatic lipogenesis by competitively binding to miR-21 with LRP6. *Metabolism: Clinical and Experimental*, *94*, 1–8. https://doi.org/10.1016/J.METABOL.2019.01.018

Huang, P., Huang, F. zhou, Liu, H. zheng, Zhang, T. yi, Yang, M. shi, & Sun, C. zheng. (2019b). LncRNA MEG3 functions as a ceRNA in regulating hepatic lipogenesis by competitively binding to miR-21 with LRP6. *Metabolism: Clinical and Experimental*, *94*, 1–8. https://doi.org/10.1016/J.METABOL.2019.01.018

Huang, R., Duan, X., Fan, J., Li, G., & Wang, B. (2019). Role of noncoding RNA in development of nonalcoholic fatty liver disease. *BioMed Research International*, *2019*. https://doi.org/10.1155/2019/8690592

Hydes, T., Brown, E., Hamid, A., Bateman, A. C., & Cuthbertson, D. J. (2021). Current and Emerging Biomarkers and Imaging Modalities for Nonalcoholic Fatty Liver Disease: Clinical and Research Applications. *Clinical Therapeutics*, *43*(9), 1505–1522. https://doi.org/10.1016/J.CLINTHERA.2021.07.012

Ihunnah, C. A., Jiang, M., & Xie, W. (2011). Nuclear receptor PXR, transcriptional circuits and metabolic relevance. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1812*(8), 956–963. https://doi.org/10.1016/J.BBADIS.2011.01.014

Integrated DNA Technologies. (2022). OligoAnalyzer Tool - primer analysis | IDT. Retrieved January 24, 2022, from idtDNA website: https://www.idtdna.com/pages/tools/oligoanalyzer?returnurl=%2Fcalc%2Fanalyzer

Jun Guo, et al., 2018. (n.d.). Ultraconserved element uc.372 drives hepatic lipid accumulation by suppressing miR-195/miR4668 maturation. | Sigma-Aldrich. Retrieved April 21, 2022, from MilliporeSigma website: https://www.sigmaaldrich.com/US/en/tech-docs/paper/1458667

Kahali, B., Halligan, B., & Speliotes, E. K. (2015). Insights from Genome-Wide Association Analyses of Nonalcoholic Fatty Liver Disease. *Seminars in Liver Disease*, *35*(4), 375–391. https://doi.org/10.1055/S-0035-1567870

Kent, J. (n.d.). UCSC In-Silico PCR. Retrieved January 24, 2022, from In -Silico PCR website: https://genome.ucsc.edu/cgi-bin/hgPcr

Khalifa, O., Errafii, K., Al-Akl, N. S., & Arredouani, A. (2020a). Noncoding RNAs in Nonalcoholic Fatty Liver Disease: Potential Diagnosis and Prognosis Biomarkers. *Disease Markers*, *2020*, 8822859–8822859. https://doi.org/10.1155/2020/8822859

Khalifa, O., Errafii, K., Al-Akl, N. S., & Arredouani, A. (2020b). Noncoding RNAs in Nonalcoholic Fatty Liver Disease: Potential Diagnosis and Prognosis Biomarkers. *Disease*

Markers, *2020*. https://doi.org/10.1155/2020/8822859

Kim, Y. A., Park, K. K., & Lee, S. J. (2020). LncRNAs Act as a Link between Chronic Liver Disease and Hepatocellular Carcinoma. *International Journal of Molecular Sciences 2020, Vol. 21, Page 2883*, *21*(8), 2883. https://doi.org/10.3390/IJMS21082883

Lander, E. S. (2019). Primer3 Input. Retrieved January 24, 2022, from Primer3web website: https://bioinfo.ut.ee/primer3/

Li, C., Chen, J., Zhang, K., Feng, B., Wang, R., & Chen, L. (2015). Progress and Prospects of Long Noncoding RNAs (lncRNAs) in Hepatocellular Carcinoma. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *36*(2), 423–434. https://doi.org/10.1159/000430109

Li, W., Chen, X., Lin, M., & Huang, D. (2017). Up-regulated HOTAIR induced by fatty acids inhibits PTEN expression and increases triglycerides accumulation in HepG2 cells. *Food and Nutrition Research*, *61*. https://doi.org/10.1080/16546628.2017.1412794/SUPPL_FILE/ZFNR_A_1412794_SM6966.TIF

Lindsey S Treviño, T. A. K. (n.d.). Endocrine Disruptors and Developmental Origins of Nonalcoholic Fatty Liver Disease - PubMed. Retrieved January 21, 2022, from https://pubmed.ncbi.nlm.nih.gov/29126168/

Liu, J., Tang, T., Wang, G. D., & Liu, B. (2019). LncRNA-H19 promotes hepatic lipogenesis by directly regulating miR-130a/PPARγ axis in non-alcoholic fatty liver disease. *Bioscience Reports*, *39*(7). https://doi.org/10.1042/BSR20181722

Liu, Y., Ding, W., Yu, W., Zhang, Y., Ao, X., & Wang, J. (2021). Long non-coding RNAs: Biogenesis, functions, and clinical significance in gastric cancer. *Molecular Therapy Oncolytics*, *23*, 458. https://doi.org/10.1016/J.OMTO.2021.11.005

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., … Econs, M. J. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, *518*(7538), 197–206. https://doi.org/10.1038/NATURE14177

Lu, D., Luo, P., Wang, Q., Ye, Y., & Wang, B. (2017). lncRNA PVT1 in cancer: A review and meta-analysis. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, *474*, 1–7. https://doi.org/10.1016/J.CCA.2017.08.038

Mancina, R. M., Dongiovanni, P., Petta, S., Pingitore, P., Meroni, M., Rametta, R., … Romeo, S. (2016a). The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent. *Gastroenterology*, *150*(5), 1219-1230.e6. https://doi.org/10.1053/J.GASTRO.2016.01.032

Mancina, R. M., Dongiovanni, P., Petta, S., Pingitore, P., Meroni, M., Rametta, R., … Romeo, S. (2016b). The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent. *Gastroenterology*, *150*(5), 1219-1230.e6. https://doi.org/10.1053/J.GASTRO.2016.01.032

Mantovani, A., Byrne, C. D., Bonora, E., & Targher, G. (2018). Nonalcoholic Fatty Liver Disease and Risk of Incident Type 2 Diabetes: A Meta-analysis. *Diabetes Care*, *41*(2), 372–382. https://doi.org/10.2337/DC17-1902

Marchesini, G., Day, C. P., Dufour, J. F., Canbay, A., Nobili, V., Ratziu, V., … Mathus-Vliegen, L. (2016). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *Journal of Hepatology*, *64*(6), 1388–1402. https://doi.org/10.1016/J.JHEP.2015.11.004

Maria D. Paraskevopoulou et al. (2016). DIANA tools - LncBase Experimental v2. Retrieved January 24, 2022, from DIANA tools website: https://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experimental

Marrero, J. A., Fontana, R. J., Su, G. L., Conjeevaram, H. S., Emick, D. M., & Lok, A. S. (2002). NAFLD may be a common underlying liver disease in patients with hepatocellular carcinoma in the United States. *Hepatology*, *36*(6), 1349–1354. https://doi.org/10.1053/jhep.2002.36939

McGlynn, K. A., Petrick, J. L., & London, W. T. (2015, May 1). Global Epidemiology of Hepatocellular Carcinoma: An Emphasis on Demographic and Regional Variability. *Clinics in Liver Disease*, Vol. 19, pp. 223–238. https://doi.org/10.1016/j.cld.2015.01.001

Mitra, S., De, A., & Chowdhury, A. (2020). Epidemiology of non-alcoholic and alcoholic fatty liver diseases. *Translational Gastroenterology and Hepatology*, *5*. https://doi.org/10.21037/TGH.2019.09.08

Nishida, N., & Goel, A. (2011). Genetic and Epigenetic Signatures in Human Hepatocellular Carcinoma:A Systematic Review. *Current Genomics*, *12*(2), 130–137. https://doi.org/10.2174/138920211795564359

Palmentieri, B., de Sio, I., La Mura, V., Masarone, M., Vecchione, R., Bruno, S., … Persico, M. (2006). The role of bright liver echo pattern on ultrasound B-mode examination in the diagnosis of liver steatosis. *Digestive and Liver Disease : Official Journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, *38*(7), 485–489. https://doi.org/10.1016/J.DLD.2006.03.021

Pouwels, S., Sakran, N., Graham, Y., Leal, A., Pintar, T., Yang, W., … Ramnarain, D. (2022). Non-alcoholic fatty liver disease (NAFLD): a review of pathophysiology, clinical management and effects of weight loss. *BMC Endocrine Disorders*, *22*(1), 1–9. https://doi.org/10.1186/S12902-022-00980-1/FIGURES/1

Pu, J., Wei, H., Tan, C., Qin, B., Zhang, Y., Wang, A., & Wang, J. (2019). Long noncoding RNA SNHG14 facilitates hepatocellular carcinoma progression through regulating miR-4673/SOCS1. *American Journal of Translational Research*, *11*(9), 5897. Retrieved from /pmc/articles/PMC6789279/

Rashad, N. M., Khalil, U. A., Ahmed, S. M., Hussien, M. H., Sami, M. M., & Wadea, F. M. (2022). The Expression Level of Long Non-Coding RNA PVT1 as a Diagnostic Marker for Advanced Stages in Patients with Nonalcoholic Fatty Liver Disease. *The Egyptian Journal of Hospital Medicine*, *87*(1), 1825–1834. https://doi.org/10.21608/EJHM.2022.230277

Rashidmayvan, M., Sahebi, R., & Majid Ghayour-Mobarhan, ·. (2022). Long non-coding RNAs: a valuable biomarker for metabolic syndrome. *Molecular Genetics and Genomics 2022*, 1–15. https://doi.org/10.1007/S00438-022-01922-1

Ratziu, V., Bonyhay, L., Di Martino, V., Charlotte, F., Cavallaro, L., Sayegh-Tainturier, M. H., … Poynard, T. (2002). Survival, liver failure, and hepatocellular carcinoma in obesity-

related cryptogenic cirrhosis. *Hepatology*, *35*(6), 1485–1493. https://doi.org/10.1053/jhep.2002.33324

Rinella, M., & Charlton, M. (2016). The globalization of nonalcoholic fatty liver disease: Prevalence and impact on world health. *Hepatology (Baltimore, Md.)*, *64*(1), 19–22. https://doi.org/10.1002/HEP.28524

Saadeh, S., Younossi, Z. M., Remer, E. M., Gramlich, T., Ong, J. P., Hurley, M., … Sheridan, M. J. (2002). The utility of radiological imaging in nonalcoholic fatty liver disease. *Gastroenterology*, *123*(3), 745–750. https://doi.org/10.1053/GAST.2002.35354

Seko, Y., Yamaguchi, K., & Itoh, Y. (2018, April 1). The genetic backgrounds in nonalcoholic fatty liver disease. *Clinical Journal of Gastroenterology*, Vol. 11, pp. 97–102. https://doi.org/10.1007/s12328-018-0841-9

Sharda, M., Yagnik, D., Soni, A., & Nigam, H. (2015). Non-alcoholic Fatty Liver Disease (NAFLD) and its association with metabolic syndrome and cardiovascular diseases. In *International Multispecialty Journal of Health (IMJH)* (Vol. 1).

Sivell, C. (2019). Nonalcoholic Fatty Liver Disease: A Silent Epidemic. *Gastroenterology Nursing : The Official Journal of the Society of Gastroenterology Nurses and Associates*, *42*(5), 428–434. https://doi.org/10.1097/SGA.0000000000000443

Sookoian, S., & Pirola, C. J. (2011). Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. *Hepatology (Baltimore, Md.)*, *53*(6), 1883–1894. https://doi.org/10.1002/HEP.24283

Sookoian, S., Rohr, C., Salatino, A., Dopazo, H., Gianotti, T. F., Castaño, G. O., & Pirola, C. J. (2017a). Genetic variation in long noncoding RNAs and the risk of nonalcoholic fatty liver disease. *Oncotarget*, *8*(14), 22917–22926. https://doi.org/10.18632/ONCOTARGET.15286

Sookoian, S., Rohr, C., Salatino, A., Dopazo, H., Gianotti, T. F., Castaño, G. O., & Pirola, C. J. (2017b). Genetic variation in long noncoding RNAs and the risk of nonalcoholic fatty liver disease. *Oncotarget*, *8*(14), 22917–22926. https://doi.org/10.18632/oncotarget.15286

Sukowati, C. H. C., Cabral, L. K. D., Tiribelli, C., & Pascut, D. (2021). Circulating Long and Circular Noncoding RNA as Non-Invasive Diagnostic Tools of Hepatocellular Carcinoma. *Biomedicines 2021, Vol. 9, Page 90*, *9*(1), 90. https://doi.org/10.3390/BIOMEDICINES9010090

Sweeney, B. A., Tagmazian, A. A., Ribas, C. E., Finn, R. D., Bateman, A., & Petrov, A. I. (2020). Exploring Non-Coding RNAs in RNAcentral. *Current Protocols in Bioinformatics*, *71*(1). https://doi.org/10.1002/CPBI.104

Tang, S. H., Gao, J. H., Wen, S. L., Tong, H., Yan, Z. P., Liu, R., & Tang, C. W. (2017). Expression of cyclooxygenase-2 is correlated with lncRNA-COX-2 in cirrhotic mice induced by carbon tetrachloride. *Molecular Medicine Reports*, *15*(4), 1507–1512. https://doi.org/10.3892/MMR.2017.6161

Targher, G., Marchesini, G., & Byrne, C. D. (2016). Risk of type 2 diabetes in patients with non-alcoholic fatty liver disease: Causal association or epiphenomenon? *Diabetes and Metabolism*, *42*(3), 142–156. https://doi.org/10.1016/J.DIABET.2016.04.002

Thornton, B., & Basu, C. (2011). Real-time PCR (qPCR) primer design using free online software. *Biochemistry and Molecular Biology Education*, *39*(2), 145–154. https://doi.org/10.1002/BMB.20461

Tomah, S., Hamdy, O., Abuelmagd, M. M., Hassan, A. H., Alkhouri, N., Al-Badri, M. R., Gardner, H., Eldib, A. H., & Eid, E. A. (2021). Prevalence of and risk factors for non-alcoholic fatty liver disease (NAFLD) and fibrosis among young adults in Egypt. *BMJ open gastroenterology*, *8*(1), e000780. https://doi.org/10.1136/bmjgast-2021-000780

Turchinovich, A., Baranova, A., Drapkina, O., & Tonevitsky, A. (2018). Cell-free circulating nucleic acids as early biomarkers for NAFLD and NAFLD-associated disorders. *Frontiers in Physiology*, *9*(SEP), 1256. https://doi.org/10.3389/FPHYS.2018.01256/BIBTEX

VoPham, T. (2019). Environmental risk factors for liver cancer and nonalcoholic fatty liver disease. *Current Epidemiology Reports*, *6*(1), 50. https://doi.org/10.1007/S40471-019-0183-2

Wang, H., Cao, Y., Shu, L., Zhu, Y., Peng, Q., Ran, L., … Fan, J. (2020). Long non-coding RNA (lncRNA) H19 induces hepatic steatosis through activating MLXIPL and mTORC1 networks in hepatocytes. *Journal of Cellular and Molecular Medicine*, *24*(2), 1399–1412. https://doi.org/10.1111/JCMM.14818

Wang, J., Yang, W., Chen, Z., Chen, J., Meng, Y., Feng, B., … Yang, J. (2018). *Long Noncoding RNA lncSHGL Recruits hnRNPA1 to Suppress Hepatic Gluconeogenesis and Lipogenesis*. https://doi.org/10.2337/db17-0799

Wang, Y., Hylemon, P. B., & Zhou, H. (2021). Long Noncoding RNA H19: A Key Player in Liver Diseases. *Hepatology*, *74*(3), 1652–1659. https://doi.org/10.1002/HEP.31765

Wu, H., Yang, L., & Chen, L. L. (2017). The Diversity of Long Noncoding RNAs and Their Generation. *Trends in Genetics : TIG*, *33*(8), 540–552. https://doi.org/10.1016/J.TIG.2017.05.004

Younossi, Z. M., Koenig, A. B., Abdelatif, D., Fazel, Y., Henry, L., & Wymer, M. (2016a). Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology (Baltimore, Md.)*, *64*(1), 73–84. https://doi.org/10.1002/HEP.28431

Younossi, Z. M., Koenig, A. B., Abdelatif, D., Fazel, Y., Henry, L., & Wymer, M. (2016b). Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology*, *64*(1), 73–84. https://doi.org/10.1002/hep.28431

Yu, F., Chen, B., Dong, P., & Zheng, J. (2017). HOTAIR Epigenetically Modulates PTEN Expression via MicroRNA-29b: A Novel Mechanism in Regulation of Liver Fibrosis. *Molecular Therapy*, *25*(1), 205. https://doi.org/10.1016/J.YMTHE.2016.10.015

Yu, F., Jiang, Z., Chen, B., Dong, P., & Zheng, J. (2017). NEAT1 accelerates the progression of liver fibrosis via regulation of microRNA-122 and Kruppel-like factor 6. *Journal of Molecular Medicine (Berlin, Germany)*, *95*(11), 1191–1202. https://doi.org/10.1007/S00109-017-1586-5

Yu, F., Lu, Z., Cai, J., Huang, K., Chen, B., Li, G., … Zheng, J. (2015). MALAT1 functions as

a competing endogenous RNA to mediate Rac1 expression by sequestering miR-101b in liver fibrosis. *Cell Cycle (Georgetown, Tex.)*, *14*(24), 3885–3896. https://doi.org/10.1080/15384101.2015.1120917

Yu, F., Zheng, J., Mao, Y., Dong, P., Li, G., Lu, Z., … Fan, X. (2015). Long non-coding RNA APTR promotes the activation of hepatic stellate cells and the progression of liver fibrosis. *Biochemical and Biophysical Research Communications*, *463*(4), 679–685. https://doi.org/10.1016/J.BBRC.2015.05.124

Yuan, X., Wang, J., Tang, X., Li, Y., Xia, P., & Gao, X. (2015). Berberine ameliorates nonalcoholic fatty liver disease by a global modulation of hepatic mRNA and lncRNA expression profiles. *Journal of Translational Medicine*, *13*(1). https://doi.org/10.1186/S12967-015-0383-6

Zhang, L., Yang, Z., Trottier, J., Barbier, O., & Wang, L. (2017a). Long noncoding RNA MEG3 induces cholestatic liver injury by interaction with PTBP1 to facilitate shp mRNA decay. *Hepatology (Baltimore, Md.)*, *65*(2), 604–615. https://doi.org/10.1002/HEP.28882

Zhang, L., Yang, Z., Trottier, J., Barbier, O., & Wang, L. (2017b). Long noncoding RNA MEG3 induces cholestatic liver injury by interaction with PTBP1 to facilitate shp mRNA decay. *Hepatology*, *65*(2), 604–615. https://doi.org/10.1002/HEP.28882/SUPPINFO

Zhang, T. N., Wang, W., Huang, X. M., & Gao, S. Y. (2021). Non-Coding RNAs and Extracellular Vehicles: Their Role in the Pathogenesis of Gestational Diabetes Mellitus. *Frontiers in Endocrinology*, *12*. https://doi.org/10.3389/FENDO.2021.664287

Zhang, X., Hong, R., Chen, W., Xu, M., & Wang, L. (2019). The role of long noncoding RNA in major human disease. *Bioorganic Chemistry*, *92*. https://doi.org/10.1016/J.BIOORG.2019.103214

Zhou, J., Zhou, F., Wang, W., Zhang, X. J., Ji, Y. X., Zhang, P., … Li, H. (2020). Epidemiological Features of NAFLD From 1999 to 2018 in China. *Hepatology (Baltimore, Md.)*, *71*(5), 1851–1864. https://doi.org/10.1002/HEP.31150

APPENDIX 1

Informed consent form: English version

THE AMERICAN UNIVERSITY IN CAIRO
INSTITUTIONAL REVIEW BOARD

Documentation of Informed Consent for Participation in Research Study

Project Title: *Finding biomarkers from long non-coding RNAs in serum of patients with non-alcoholic fatty liver disease.*

Principal Investigator: *Nouran Yonis; Tel : +201094813939; email : nouranyonis@aucegypt.edu*

*You are being asked to participate in a research study. The purpose of the research is to identify and quantify differentially expressed long non-coding RNAs in samples of blood, which would help to detect non-alcoholic fatty liver disease (NAFLD) at an early stage, consequently circumventing serious complications such as liver cirrhosis and cancer. The findings may be published, or presented. The expected duration of your participation is 15 minutes.

* The procedures of the research will be as follows : A group of NAFLD patients will be recruited for participation in the study (group A). patients with NAFLD. An additional group of healthy volunteers will also be included (group B). For all participants, a trained nurse with all the listed high precautions will take a blood sample from you. The nurse would be following all the safety and hygienic practices and would place the blood sample in special glass closed container and the sample will be subjected to further analysis.

* There will be certain discomforts associated with this research: A mild stinging sensation might result from the injection needle with minor possibility of bruises at the injection site after collecting blood samples. Also, you might experience a drop in blood pressure. In such case or in case of any harm affects you due to participation in the research; you will be subjected to urgent medical care.

*There *will be* benefits to you from this research. This study aims to define a novel modality for an early diagnosis of NAFLD, which is rapidly rising above other etiologies as the instigating factor in the development of hepatocellular carcinoma. Moreover, by 2030, NAFLD is projected to be the leading cause of liver transplantation. Given the overwhelming proportion of the Egyptian population with NAFLD (silent majority).

Although a new diagnostic test is not guaranteed, your participation will certainly help with advancing research in this area, especially in Egypt, where the condition is highly epidemiological and undersurveilled. Moreover, you will have access to an advanced mode of screening that could potentially serve as a primary factor in predicting disease prognosis and play a key role in determining the course of treatment to be followed. Healthy participants would benefit in learning of how their profiles compare to the diseased state and would be getting free medical assessment, in the course of the recruitment process.

*The information you provide for purposes of this research *is confidential*.

**Questions about the research, my rights, or research-related injuries should be directed to Nouran Yonis* at *+20 109 48 13939* and *Dr. Anwar Abdelnaser* at *+20 100 981 3624*.

*Participation in this study is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled. You may discontinue participation at any time without penalty or the loss of benefits to which you are otherwise entitled.

Signature _____

Printed Name _____

Date _____

APPENDIX 2

Informed consent form: Arabic version

الجامعة الأمريكية بالقاهرة

استمارة موافقة مسبقة للمشاركة في دراسة بحثية

عنوان البحث : العثور على المؤشرات الحيوية من الأحماض النووية الريبية الطويلة غير المشفرة (lncRNAs) المصاحبة لمرض الكبد الدهني غير الكحولي في الدم الخاص بالمصابين.

الباحث الرئيسي: نوران يونس، باحث مساعد بالجامعة الأمريكية بالقاهرة
البريد الالكتروني: nouranyonis@aucegypt.edu
الهاتف: 3939 481 109 20+

انت مدعو للمشاركة فى دراسة بحثية تهدف للتشخيص المبكر لمرض الكبد الدهني غير الكحولي عن طريق ايجاد طرق جديدة، دقيقة، غير مكلفة وغير مؤلمة مما قد يتيح التدخل الطبي المبكر، قبل حدوث مضاعفات تصل إلى تليف الكبد وسرطان الكبد.

هدف الدراسة هو الكشف عن وقياس مستوى الاحماض النووية الريبية الطويلة غير المشفِّرة (lncRNAs) المصاحبة لمرض الكبد الدهني غيرالكحولى في الدم مما قد يساهم في التوصل لطريقة للتشخيص المبكر للمرض وتفادى مضاعفاته.

نتائج البحث ستنشر فى دوريه متخصصة أو مؤتمر علمي أو ربما كليهما.

المدة المتوقعة للمشاركة فى هذا البحث هي خمسة عشر دقيقة.

اجراءات الدراسة تشتمل على جمع عينات دم من مرضى متطوعين للمشاركة في البحث (مجموعة أ)، فئة مصابة بمرض الكبد الدهني غير الكحولي بالإضافة لجمع عينات دم من متطوعين أصحاء (مجموعة ب). سيتم سحب العينات بواسطة ممرضة متمرسة، ثم استخدام تلك العينات لاستخلاص الحمض النووي الريبي الطويل غير المشفِّر وقياس مستوى وجوده في كل عينة.

المخاطر المتوقعة من المشاركة في هذه الدراسة هي احتمال ظهور تجمع دموي بسيط ومؤقت من مكان اخذ عينة الدم، وربما تشعر بهبوط نسبيا جراء سحب الدم. عند حدوث أي أضرار بسبب المشاركة بالبحث فسيتم إعطاؤك الرعاية الطبية العاجلة اللازمة.

الاستفادة المتوقعة تهدف هذه الدراسة إلى تحديد طريقة جديدة للتشخيص المبكر لـمرض الكبد الدهني الغير كحولي ، والذي يعلو بسرعة فوق العوامل الأخرى المسببة لسرطان الكبد. علاوة على ذلك، بحلول العام 2030، من المتوقع أن يكون الكبد الدهني الغير كحولي هو السبب الرئيسي في زراعة الكبد مما يجعل تشخيصه المبكر في غاية الأهمية نظرا لنسبته المرتفعة بين المصريين والتي تمثل الأغلبية الصامتة.

على الرغم من أن نتائج الدراسة غير مضمونة، مشاركتك ستساعد بالتأكيد في تطوير الأبحاث في هذا المجال، خاصةً في مصر، نظرا لوبائية المرض الشديدة. بالإضافة إلى ذلك، ستتمكن من الحصول على نمط متقدم من الفحص يمكن أن يكون عاملًا أساسيًا في التنبؤ بتشخيص المرض ويساهم بدورًا رئيسيًا في تحديد مسار العلاج الذي يجب اتباعه. سوف يستفيد المشاركون الأصحاء في مقارنة نتائجهم بالحالة المرضية والحصول على تقييم طبي مجاني.

السرية واحترام الخصوصية: المعلومات التي ستدلى بها في هذا البحث سوف تكون سرية.

أي أسئلة متعلقة بهذه الدراسة أو حقوق المشاركين فيها أوعند حدوث أى اصابات ناتجة عن هذه المشاركة يرجى التوجه بالاتصال ب (د. أنور عبد الناصر – تليفون: 01009813624) او نوران يونس -

ان المشاركة فى هذه الدراسة ماهى الا عمل تطوعى, حيث أن الامتناع عن المشاركة لايتضمن أى عقوبات أو فقدان أى مزايا تحق لك. ويمكنك أيضا التوقف عن المشاركة فى أى وقت من دون عقوبة أو فقدان لهذه المزايا.

الامضاء: .....

اسم المشارك : .....

التاريخ : ...../...../.....

APPENDIX 3

AUC IRB approval form

CASE #2020-2021-018

THE AMERICAN UNIVERSITY IN CAIRO
INSTITUTIONAL REVIEW BOARD

To: Nouran Samy Ahamed

Cc: Sherihan Hassan

From: Atta Gebril, Chair of the IRB

Date: Oct. 26, 2020

Re: IRB approval

This is to inform you that I reviewed your revised research proposal entitled **" Finding biomarkers from long-non coding RNAs in serum of patients with non-alcoholic fatty liver disease"** and determined that it required consultation with the IRB under the "expedited" category. As you are aware, the members of the IRB suggested certain revisions to the original proposal, but your new version addresses these concerns successfully. The revised proposal used appropriate procedures to minimize risks to human subjects and that adequate provision was made for confidentiality and data anonymity of participants in any published record. I believe you will also make adequate provision for obtaining informed consent of the participants.

This approval letter was issued under the assumption that you have not started data collection for your research project. Any data collected before receiving this letter could not be used since this is a violation of the IRB policy.

Please note that IRB approval does not automatically ensure approval by CAPMAS, an Egyptian government agency responsible for approving some types of off-campus research. CAPMAS issues are handled at AUC by the office of the University Counsellor, Dr. Ashraf Hatem. The IRB is not in a position to offer any opinion on CAPMAS issues, and takes no responsibility for obtaining CAPMAS approval.

This approval is valid for only one year. In case you have not finished data collection within a year, you need to apply for an extension.

Thank you and good luck.

Atta egebril

Dr. Atta Gebril
IRB chair, The American University in Cairo
2046 HUSS Building
T: 02-26151919
Email: agebril@aucegypt.edu

Institutional Review Board
The American University in Cairo
AUC Avenue, P.O. Box 74
New Cairo 11835, Egypt.
tel 20.2.2615.1000
fax 20.2.27957565
Email: aucirb@aucegypt.edu

Supplementary (Bioinformatics analysis)

	lncRNA	Sequence	Th
1			
2	LINC01198	CUCUCGCCUGCCUUCGGCCCUUAGGCUCCGGGAGAUU	
3	LINC00607	>URS0000126579 Homo sapiens lncRNA	
4	MAFTRR	>URS00003CD6E9 Homo sapiens lncRNA	
5	PWRN1	>URS000075E11D Homo sapiens lncRNA	
6	LINC00907	>URS000075B00D Homo sapiens lncRNA	
7	DUXAP8	>URS00008B9D07 Homo sapiens lncRNA	
8	LINC00924	>URS0000759C1D Homo sapiens lncRNA	
9	KCNMA1-AS1	>URS0000759FA9 Homo sapiens lncRNA	
10	LINC01255	>URS000075D667 Homo sapiens lncRNA	
11	PPFIA2-AS1	>URS000075A7BF Homo sapiens lncRNA	
12	LINC00635	>URS0000759AAC Homo sapiens lncRNA	
13	DCTN1-AS1	>URS0000D77F4C Homo sapiens lncRNA	
14	FZD10-AS1	>URS000075C477 Homo sapiens lncRNA	
15	AL591501.1	>URS0000EF542F Homo sapiens lncRNA	
16	CA3-AS1	>URS000075D522 Homo sapiens lncRNA	
17	SPRY4-AS1	>URS000009C154 Homo sapiens lncRNA	
18	LINC00299	>URS000075F09A Homo sapiens lncRNA	
19	TEX41	>URS000075B9B4 Homo sapiens lncRNA	
20	SOX2-OT	>URS000075D6CA Homo sapiens lncRNA	
21	DPP10-AS2	>URS000018337C Homo sapiens lncRNA	
22	LINC02332	>URS000076FB38 Homo sapiens lncRNA	
23	AC116565.1	>URS0000EED225 Homo sapiens lncRNA	
24	AC024267.6	>URS00000C649C Homo sapiens lncRNA	
25	AC055854.1	>URS00008BF90F Homo sapiens lncRNA	
26	SLIT3-AS1	>URS000075C609 Homo sapiens lncRNA	
27	AC239809.3	>URS0000EED181 Homo sapiens lncRNA	
28	GM32017		
29	GM17494	>URS000078758E Mus musculus lncRNA	
30	SNHG18	>URS000075CB75 Homo sapiens lncRNA	
31	5430405H02RIK	>URS000075AF84 Mus musculus lncRNA	
32	BACH2OS	>URS000075DD42 Mus musculus lncRNA	
33	H19	>URS0000812128 Homo sapiens lncRNA	
34	GM4651	>URS0000CC73D9 Mus musculus lncRNA	
35	9330158H04RIK	>URS00008C4225 Mus musculus lncRNA	
36	9330159M07RIK	>URS000075B3CB Mus musculus lncRNA	
37	SNHG3	>URS000075EC7A Homo sapiens lncRNA	
38	D630024D03RIK	>URS000075C60F Mus musculus lncRNA	
39	A530083I20RIK	>URS00009B4709 Mus musculus lncRNA	
40	GM16028	>URS0000784FF6 Mus musculus lncRNA	
41	4930405D11RIK	>URS0001BBE385 Mus musculus lncRNA	
42	2810410L24RIK	>URS000075CD69 Mus musculus lncRNA	
43	RIAN	>URS00008BC75E Homo sapiens lncRNA	
44	GM17206	>URS00007758F2 Mus musculus lncRNA	
45	E330020D12RIK	>URS000075A8DA Mus musculus lncRNA	
46	E330012B07RIK	>URS000075DEF3 Mus musculus lncRNA	
47	AC160336.1	>URS0000CC4B4C Mus musculus lncRNA	
48	SNHG17	>URS0000CCDFFF Homo sapiens lncRNA	
49	2610016A17RIK	>URS000076F4AA Mus musculus lncRNA	
50	BC048559	>URS0000A888D0 Mus musculus lncRNA	
51	GM50166	>URS0000E7CFCD Mus musculus lncRNA	
52	MEG3	>URS0000759EA9 Homo sapiens lncRNA	
53	GM17028	>URS000077CF25 Mus musculus lncRNA	
54	GM15651	>URS00009AF8BE Mus musculus lncRNA	
55	4930448E22RIK	>URS0000A7AFA6 Mus musculus lncRNA	
56	4930467D21RIK	>URS0000759C45 Mus musculus lncRNA	
57	4933413J09RIK	>URS000075C903 Mus musculus lncRNA	
58	A530058N18RIK	>URS000075EDF9 Mus musculus lncRNA	
59	GM2200	>URS000066ABDC Mus musculus scaRNA	
60	SNHG14	>URS0000ABD7E9 Homo sapiens lncRNA	
61	GM17167		
62	GM10874	>URS0000811992 Mus musculus lncRNA	
63	PVT1	>URS00008E3A67 Homo sapiens lncRNA	
64	Wider Range	>URS0000D59BEE Homo sapiens lncRNA	
65	DUBR	>URS000075AD4F Homo sapiens lncRNA	
66	SNHG17	>URS0000EEE0FC Homo sapiens lncRNA	
67	SNHG17	>URS0000CCDFFF Homo sapiens lncRNA	
68	SNHG14	>URS0000ABD7E9 Homo sapiens lncRNA	
69	LncARSR	>URS0000811F6D Homo sapiens lncRNA	
70	NEAT1	>URS000075DAEC Homo sapiens lncRNA	
71	MALAT-1	>URS00008C0A2F Homo sapiens lncRNA	
72	lnRNA-CoX2		
73	HOTAIR	>URS000075C808 Homo sapiens lncRNA	
74	APTR	>URS000075BCAA Homo sapiens lncRNA	
75	B4GALT1-AS1	>URS000075AF86 Homo sapiens lncRNA	
76	FLRL2		
77	uc.372	>hg38:chr14:35573866-35574077	
78	uc.372	AGCCAGATCTAATACTAAGCTC	
79			