Medical University of South Carolina

MEDICA

Spring 5-20-2023

Identifying the Temporal N-linked Glycosylation Changes During Liver Disease Progression: from Liver Injury to End-stage Liver Disease

Shaaron Ochoa-Rios
Medical University of South Carolina

Recommended Citation

Identifying the Temporal N-linked Glycosylation Changes During Liver Disease

Progression: from Liver Injury to End-stage Liver Disease

Shaaron Ochoa-Rios

A dissertation submitted to the faculty of the Medical University of South Carolina in

partial fulfillment of the requirements for the degree of Doctor of Philosophy in the

College of Graduate Studies.

Department of Cell and Molecular Pharmacology and Experimental Therapeutics

2023

Approved by:

Chairman, Advisory Committee _____

Anand Mehta

_____

Peggy Angel

_____

Haizhen (Jen) Wang

_____

Michael C. Ostrowski

_____

Suzanne Craig

Table of Contents

ABSTRACT

SHAARON OCHOA-RIOS. Identifying the Temporal N-linked Glycosylation Changes During Liver Disease Progression: from Liver Injury to End-stage Liver Disease. (Under the direction of ANAND MEHTA)

The high mortality rates of liver diseases and primary liver cancers can be attributed to the lack of screening and diagnostic strategies currently available for early detection. Non-alcoholic fatty liver disease (NAFLD) is an early stage of liver disease known to progress to a variety of pre-malignant and malignant conditions, like advanced fibrosis, cirrhosis, hepatocellular carcinoma (HCC), and cholangiocarcinoma (CCA). Based on the wide variety of diseases that NAFLD can progress to, strategies to understand and detect the progression of NAFLD are of great value. Core fucosylation of N-linked glycans has been demonstrated to be useful for the clinical diagnosis of HCC, through the use of the serum biomarker AFP-L3. However, the role of core fucosylation, catalyzed by Fucosyltransferase (*Fut8*) in HCC or any liver diseases is still not fully understood. Here, we utilized human samples and an *in vivo* approach to characterize the origin, temporal changes, and biomarker use of modifications in N-liked glycosylation by MALDI-IMS (Matrix-assisted laser desorption/ionization imaging mass spectrometry) in NAFLD and its progressive form. Spatial N-glycan analysis in NAFLD mouse and human liver biopsies revealed that fucosylated N-linked glycan modifications correlate with areas of fibrosis. Next, we use an *in vivo* liver disease induction time point study to elucidate that bisected fucosylated N-glycan modifications can be observed even before histopathological alterations are significant and are consistently altered from fatty liver disease, up to stages with liver dysplasia. For CCA, we took a biomarker discovery

approach in tissue and serum and identified that bisected fucosylated structures distinguished CCA patients from those with any other type of liver disease and normal tissue, better than the gold-standard serum biomarker for CCA, CA-19-9. Finally, we generated the first liver specific *Fut8* mouse model to further study the impact of core fucosylation in these liver diseases. Overall, the studies presented in this dissertation, elucidate the value of N-glycosylation for biomarker strategies for early detection and different stages of progressive liver disease. Importantly, these studies set the field for a mechanistic approach to one of the most characterized N-glycan modifications in liver cancer by the generation of a tissue-specific mouse model.

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

valuable knowledge for the completion of my projects but also made very special friendships.

To all my MUSC peers that throughout the years have become my closest friends, thank you for your encouragement, for your patience, and for being there during the ups and downs. To Casey, will always remember our chats, coffee/Halo pastry breaks, and our late shifts in lab during COVID. Thank you for always being so supportive, sweet, and knowledgeable. To Denys, thank you for your beautiful friendship, I'm so glad I got to do this process with you. Thank you for always being there for office talks, experiments, science talks, conferences, and crazy adventures/trips. And of course thanks for the best cakes, cupcakes, and cookies. I'm so glad I moved offices and got to meet two of my best friends. To Harmin, my birthday twin, thank you for all your support, for pushing me to learn about N-glycans, for your patience, and for all the special moments we had. To Alyson, thank you for the long chats sweet friend, siempre estare disponible cuando necesites practicar tu espanol. To Taylor, thank you for always making me smile, being supportive, and for having the best random facts. To Jordan, you always had the best stories and made conversations fun, I'm going to miss the coffee visits to the office.

To all of the past and present members in the Drake lab: Calvin, Colin, Grace, Kim, Elizabeth, and Helena. And in the Angel lab: Jacklyn, Jadey, Stephen, and Harrison. Thank you for welcoming me to the Angel office, could have not asked for better people to share an office with.

I have great gratitude for the support from my mentors and teachers at Fayetteville State University. I would not have persuaded a graduate degree in research if it was not for my experiences at FSU. To Dr. James Raynor, many thanks for introducing research to me, for your encouragement to join the RISE program, and for

supporting me to attend summer programs and doing intramural research allowing me to find my passion for scientific research. To Dr. Shirley Chao, thank you for being my first research mentor, for your patience in letting me learn what research meant, and for all the fun memories working with cockroaches.

Outside of the amazing support I had at MUSC and FSU, I could have not completed any of my academic journeys without the amazing support from my family. To my husband, Vairon, thank you for pushing me when I was doubting myself, for your understanding, for listening when I had to complain or celebrate about the lab, for listening to my science conversations, and even for your advice in mouse experiments. You made this journey much easier and more special. To my parents, mis papitos I'm immensely grateful for all of your support, and efforts you guys have made that have allowed me to complete this stage. Gracias por ensenarnos que con dedicacion y constancia podemos cumplir nuestros suenos. Hoy les doy mas que gracias por la decision mas importante que hemos tomado como familia, gracias a esa decision hoy podemos celebrar una de nuestras metas al venir a este pais. To my sister, mi hermanita, thank you for always believing in me, for all your support, for reminding me that I need to take breaks, and for pushing me to be the best of myself; eres mi ejemplo a seguir. To all my family in the US and Colombia, each of you has always encouraged me to follow my dreams, gracias por siempre estar orgullosos de mi, los amo familia.

Chapter 1: Introduction

1.1 Chronic liver diseases

The liver is the largest solid organ in the body with an array of functions: bile production, drug metabolism, immunity, digestion detoxification, and vitamin storage, to name a few (1). Due to the many functions, the liver is responsible for, maintaining a healthy liver is imperative. However, 1.5 billion cases of chronic liver diseases (CLD) have been reported worldwide with 1.32 million deaths in 2017 (2,3). A variety of etiologies can induce CLD, alcohol abuse, unhealthy lifestyle, metabolic disorders, genetic factors, or infections. Patients with CLDs do not present many symptoms, leading to disease progression, and diagnosis at late stages. CLD is defined as a constant liver injury over time affecting liver function for a period longer than six months (4). During CLDs the liver will suffer from inflammation, destruction, and regeneration leading to fibrosis and cirrhosis (4). CLD can be considered an umbrella term for a variety of liver diseases, all known to progress to cirrhosis but with different etiologies and even different disease progression characteristics. Due to the variety of risk factors that can lead to CLDs, it has been challenging to select a specific population to screen. Overall, there is an urgent need to understand more about their disease progression and identify biomarkers for early diagnosis.

Here, we will focus on two main CLDs that are risk factors for two different primary liver cancers: Non-alcoholic fatty liver disease for hepatocellular carcinoma or cholangiocarcinoma and primary sclerosing cholangitis for cholangiocarcinoma.

1.1.2 Non-alcoholic fatty liver disease (NAFLD)

Non-Alcoholic Fatty Liver Disease (NAFLD) is the most common type of CLD in the world, its prevalence increase has been linked to the alarmingly high rates of obesity globally, in addition to genetic variants (5–7). NAFLD includes a full spectrum of the

disease from non-alcoholic fatty liver (NAFL) to non-alcoholic steatohepatitis (NASH).

NAFL is considered a benign condition, where there is more the 5% of hepatocyte

steatosis (or fat accumulation) while the progressive form of the disease, NASH is

characterized by steatosis in addition to inflammation, hepatocyte ballooning, and

fibrosis (8–11). NASH itself can progress to cirrhosis, Hepatocellular Carcinoma (HCC),

Cholangiocarcinoma (CCA), and even death. Specifically, NASH-induced HCC is

predicted to become the number one cause of liver transplant in the United States, and it

has become a public health concern due to the lack of treatment (9,12). Both NAFL and

NASH are independent of alcohol or drugs, leaving the induction of these diseases

mainly in metabolic disorders or genetic risk factors (5,8) **(Figure 1)**.
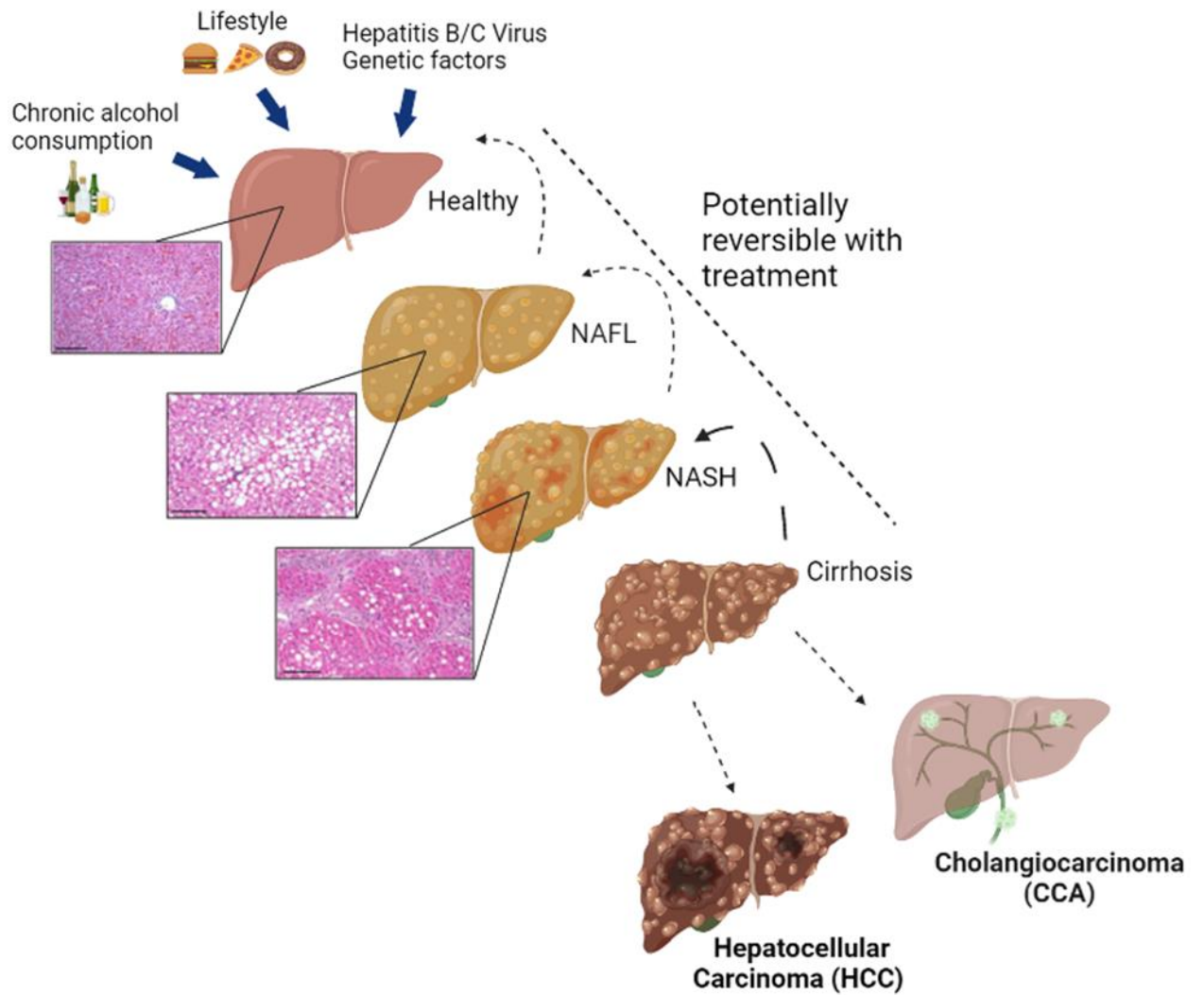
Figure 1. Schematic of the progression of NAFLD (non-alcoholic fatty liver disease) with representative Hematoxylin and Eosin (H&E) images of the histological alterations in each stage (left). NASH (non-alcoholic steatohepatitis).

1.1.2.1 Incidence and risk factors

While the incidence of NAFLD in the general population is 25%, this percentage can increase up to 90% in extremely obese populations (13). A geographical variation has been identified in NAFLD where black ethnicity has a protective factor with the lowest NAFLD prevalence of 14% while South America has the highest prevalence of 31%. Similarly in the United States, the prevalence of hepatic steatosis is higher in Hispanics (22.0%-45.0%) while non-Hispanic blacks have the lowest prevalence (10%-24%) (5,9,14). Although only a small percentage of NAFLD patients will have major complications from CLD, the overall number of patients with end-stage liver disease caused by NAFLD is dramatically increasing, demonstrated by the increase in the percentage of liver transplants caused by NASH (15–17). One of the reasons for the low percentage of liver-related complications is that cardiovascular disease is the most common type of death (40%) in NAFLD patients, thus most patients will die of cardiovascular complications before liver complications can progress further (9,10,18).

Another major factor that should be taken into consideration in NAFLD are sex differences. NAFLD and its etiologies are more prevalent in men and pre-menopausal women. However, after menopause, the rate of NAFLD in women significantly increases, this has been associated with a protective mechanism related to estrogen against metabolic syndrome progression. While it is known that sex can influence prevalence, risk factors, and clinical outcomes in NAFLD; sex differences have not been addressed as well as in other diseases; studies and clinical guidelines should include sex differences, menopause status, and age to address current sex bias in the field (19).

Risk factors for NAFLD include genetic (ie., variants *PNPLA3, TM6SF2*) and environmental factors (diets high in saturated fat and fructose and a sedentary lifestyle).

These factors can result in clinical features like obesity, type 2 diabetes mellitus, dyslipidemia, and hypertension (5,9). The idea that NAFLD is linked to a genetic factor is determined by the knowledge that the disease tends to be more prone in patients with a family history of NAFLD and as previously mentioned, the prevalence is based on geographic regions. Patatin-like Phospholipase domain-containing 3 (*PNPLA3*) is the most characterized variant in NAFLD. *PNPLA3* is highly associated with hepatic steatosis specifically in Hispanics (5). Transmembrane 6 superfamily member 2 (*TM6SF2*) is another missense variant associated with increased fat content and elevated serum levels of liver enzymes (5,14,20,21). Further studies continue to determine the role and mechanism of these variants in NAFLD.

Overall, studies have elucidated many genetic variations in NAFLD, but these variants only partially explain chronic liver disease outcomes according to ethnic differences, suggesting that most likely these genetic risk factors are influenced by environmental risk factors like obesity (5).

1.1.2.2 Screening and diagnosis

Most NAFLD patients do not present any significant symptoms, even when there is disease progression, this affects severely the early detection and diagnosis of NAFL/NASH (22). Serum-based- noninvasive biomarkers are considered the first tool to screen for NAFLD. For hepatic steatosis identification, fatty liver index (FLI), visceral adiposity index, NAFLD Liver Fat Score, and hepatic steatosis index (HIS) are some of the panels used (23,24). However, their sensitivity is less than 80% for detecting greater than 5% steatosis (the minimum percentage for NAFL diagnosis) (9). In addition, some of these panels consider a Body Mass Index (BMI) greater than 30 to be positive for NAFLD, which in some cases is not accurate as some NAFL/NASH patients are not

obese (25). Biomarker panels for NASH have also been implemented but also with low success, these panels consider elevated transaminases as the only decisive factors for NASH disease, and it is known that some NASH patients might have normal levels of these liver enzymes even with aggressive forms of the disease (9). Finally, biomarkers for fibrosis seem to be helpful mostly for advanced fibrosis but are not reliable to detect early fibrosis (9)(26).

The gold standard for NAFLD diagnosis is liver biopsies and while they provide a complete histological assessment of all the characteristics of NAFLD, the current scoring systems to provide a definitive diagnosis have several limitations, in addition, liver biopsies can have sample variability according to what tissue is selected for analysis (8,27).

Imaging modalities also present disadvantages for NAFLD diagnosis, including the lack of correlations between the level of fat accumulation that is observed during imaging with that observed in biopsies, they are also unable to distinguish NAFL from NASH and there are difficulties to get accurate readings in highly obese patients (8,9,26).

Overall, screening strategies for NAFLD are limited and the diagnostic strategies have serious issues that inhibit their use for accurate detection and treatment of NAFLD. To reduce the high rates of end-stage liver disease caused by NAFLD, more strategies for early disease detection are urgently needed.

1.1.3 Primary Sclerosing Cholangitis (PSC)

Primary Sclerosing Cholangitis (PSC) is a chronic and progressive biliary disease, characterized by severe inflammation, and fibrosis, with damage to the intrahepatic and extrahepatic bile ducts (28). PSCs progression seems to be consistent

from bile ducts stricture, biliary cirrhosis, portal hypertension, and progression to liver

failure (29,30). 10%-20% of PSC patients will progress to cholangiocarcinoma (CCA)

and PSC patients are considered the surveillance population for CCA diagnosis (31–33).

1.1.3.1 Incidence and risk factors

The incidence rate for PSC ranges from 0 to 1.3 per 1000,000 inhabitants/year,

and most patients diagnosed with PSC are between 30 and 40 years old (31). Males

have a two-fold increase in being diagnosed with PSC compared to females (34,35).

The precise etiologies for PSC are still not well understood. However, different

studies have reported that a combination of exposure to environmental factors and

genetic predisposition is required for disease development (32,36). Siblings of patients

with PSC have a 9-39-fold increase in PSC risk, confirming genetic predispositions. In

addition, a genome-wide study reported a strong association with specific human

leukocyte antigens (HLA), where pSC is associated with HLA class 1, 2, and 3 regions)

(33). PSC can also be categorized to be a progressive form of bowel disease, as 60%-

80% of patients with PSC have inflammatory bowel disease (IBD) (33,37).

1.1.3.2 Screening and diagnosis

About 50% of PSC patients are asymptomatic and most cases are diagnosed

accidentally from other non-related exams (34,35). Symptoms can vary from right upper

abdominal pain to fatigue, jaundice, and some weight loss (38). A cholestatic pattern in

liver biochemical tests with an elevation of the serum alkaline phosphatase is commonly

seen, and elevation of other transaminases is also common in PSC patients (33).

PSC diagnosis is based on the following criteria: 1. Increase serum alkaline

phosphatase levels for more than six months. 2. Cholangiographic findings of bile duct

structures by magnetic resonance cholangiopancreatography (MRCP), and 3. Exclusion

of causes of secondary sclerosing cholangitis (33).

Overall, screening and diagnostic strategies, are limited, and the criteria

previously mentioned do not apply to all patients. Strategies that can identify PSC

patients before disease progression are urgently needed.

1.2 Primary Liver Cancers

1.2.1 Hepatocellular Carcinoma

Hepatocellular Carcinoma (HCC) is the third leading cause of cancer-related

deaths worldwide and accounts for at least 90% of liver cancer cases (39,40). HCC risk

factors can vary based on geographical locations where Hepatitis B virus (HBV) is

reported mainly in Africa and Asia while Hepatitis C virus (HCV), excessive alcohol

consumption, and metabolic diseases like diabetes and nonalcoholic fatty liver disease

(NAFLD) have a higher incidence in Western countries (41). Sex differences also account

for the incidence, etiology, and treatments of HCC. The ratio between the prevalence of

HCC in men to women is between 2:1 and 4:1, validating a male-bias mechanism in this

cancer. Due to the significant incidence and mortality in males, studies have focused to

understand the sex-shared and sex-specific dysregulation of gene expression in HCC

(42). However, similar to the case in NAFLD, older females have a higher prevalence in

HCC and are most likely overlooked by the HCC surveillance guidelines, suggesting the

need for gender-specific screening based on specific clinical factors including age and

hormonal factors (43).

Early detection and treatment strategies for HCC are urgently needed, as the 5-

year survival rate for HCC remains less than 15% and new HCC cases continue to rise

with more than 40,000 cases reported in 2017 (40,44,45). One of the biggest limitations to improving patient prognosis and survival is the high level of heterogeneity in HCC, including diversity in genetic aberrations and cell morphology, all of which are known to affect the type of treatment and treatment response (39)(44). Treatment options include tumor resection, liver transplant, and chemotherapy, all present with specificity limitations, limited availability, and treatment resistance (46,47). Overall, there is a crucial need for biomarkers that can detect the early stages of the disease and therapeutic strategies that can address HCC heterogeneity in each specific patient.

In this dissertation, HCC is studied as an end-stage disease, where the main goal was to study the liver disease progression up to cancer instead of characterizing the type of cancer itself. Previous work by our group has focused on molecular signature characterization specific to HCC serum and tissue (48).

1.2.2 Cholangiocarcinoma

Cholangiocarcinoma (CCA), also known as bile duct cancer is considered a rare, but fatal malignancy where cancer cells form in the bile ducts of the liver (23,49,50). CCA is heterogeneous cancer that can be classified into three primary subtypes based on the tumor origin location: 1) intrahepatic CCA (iCCA), 2) perihilar CCA (pCCA), and 3) distal CCA (dCCA); pCCA and dCCA have been considered as extrahepatic CCA (eCCA) (23,50,51) **(Figure 2)**. Epithelial cells lining the bile duct, cholangiocytes are thought to be responsible for the differentiation and epithelial cell malignancy that occurs along the biliary tree and/or within the hepatic parenchyma (this is one of the factors that will determine the subtype classification) (28,50).

second-order bile ducts

Intrahepatic (iCCA)
Proximal to the second-order bile ducts within the liver parenchyma

Perihilar (pCCA)
Between the second-order bile ducts and insertion of the cystic bile duct

Distal (dCCA)
confined to the common bile duct below the cystic duct insertion

Extrahepatic (eCCA)

Cystic duct insertion

common bile duct
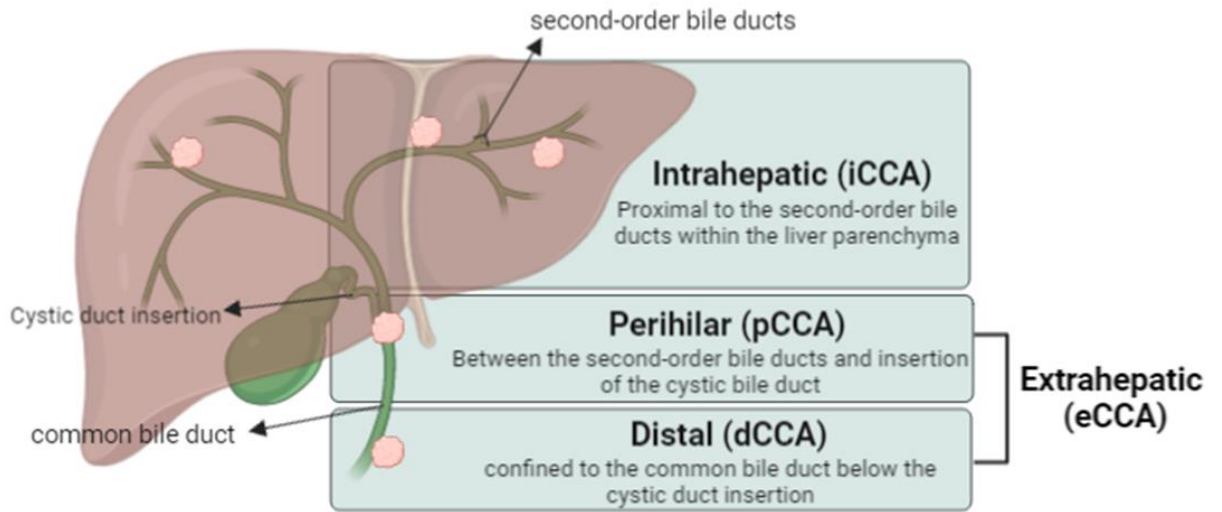
Figure 2. Schematic of the subtypes of Cholangiocarcinoma (CCA) based on the anatomical location of the tumor. CCA can be classified as intrahepatic or extrahepatic (including perihilar and distal) CCA.

1.2.2.1 Incidence and Risk Factors

CCA is the second most common type of liver cancer after HCC, about 10-15% of all primary liver cancers are diagnosed as CCA (52–54). iCCA and pCCA combined account for more than 90% of all CCAs worldwide (50). Incidence for CCA is specifically high in regions like South Korea, China, and Thailand (49).

Risk factors for iCCA include cirrhosis, chronic pancreatitis, NAFLD, diabetes and obesity, congenital hepatic fibrosis, and alcohol consumption. While risk factors for extrahepatic CCA (pCCA and dCCA) include primary sclerosing cholangitis (PSC), liver fluke infection (priority in South-eastern Asia), diabetes and obesity, chronic pancreatitis, and gout. The prevalence of risk factors like alcohol consumption, and NAFLD is increasing, consequentially increasing the incidence of CCA (23,31,38,55).

1.2.2.2 Screening and diagnosis

iCCA patients are often asymptomatic in the early stages but can develop symptoms like abdominal pain and in some cases, jaundice as the disease progresses to more advanced stages. While patients with extrahepatic CCA can present with some jaundice due to the biliary obstruction that characterizes pCCA and dCCA (31,50,51). For all CCA subtypes, Carbohydrate antigen 19-9 (CA19-9) is the serum biomarker for screening (49). However, the specificity for this biomarker is remarkably poor, due to being elevated in different conditions, including pancreatic diseases (28,56,57).

Conventional imaging modalities like MRI, CT, and ultrasonography are used for CCA diagnosis. Some work best than others according to the CCA subtype but overall, there are still many disadvantages in differentiating subtypes, and detecting tumors due to their location within the biliary tree. Histological analysis and pathologic confirmation of a biopsy are required to make a definitive diagnosis for all CCA subtypes (50,51).

1.3 N-linked Glycosylation

Glycosylation is one of the most common post-translational modifications (PTMs), where a sugar molecule is covalently attached to other macromolecules, i.e., proteins or lipids. Protein glycosylation occurs in over 50% of all human proteins at the cell surface and/or secreted proteins (58–60). The attachment of a sugar residue to the protein has key biological roles like protein folding, stability, cell-cell communication, signaling, and immune responses (61,62). Glycan structures, unlike proteins, are not encoded by a DNA template, instead, glycans can have an extensive array of possible combinations, making their characterization complex (63–65). However, the field of glycosylation has been considered a promising field of study for clinical use, specifically N-linked glycosylation (65–67). This class of glycosylation is determined by the addition of oligosaccharide moieties or glycans at the Asparagine (Asn) residue of the protein and occurs at the consensus sequence Asn–X–Ser/Thr, where X is any amino acid except proline (68). While changes in N-linked glycosylation are normal to maintain homeostasis in the cell, alterations in the expression of specific N-glycan structure modifications have been observed during disease, including cancer; this valuable fact has been exploited for the use of these modifications as promising biomarkers.

The studies presented in this dissertation focus on the characterization of N-linked glycosylation for the elucidation of specific modifications in the progression of liver disease to end-stage liver disease.

1.3.1   Processing and maturation of N-glycans

The N-glycan synthesis starts at the cytoplasmic side of the Endoplasmic Reticulum (ER) with a pre-formed oligosaccharide precursor of 14 sugars Glc3Man9(GlcNAc) (69). A 5-sugar core of the 14 residues is conserved on all secretory

and membrane proteins, including 2 N-acetyl glucosamine (GlcNAc) and 3 mannose (Man) residues. The 14 residues pre-formed precursor is then transferred to a lipid-like molecule called dolichol phosphate, with GlcNAc being the first residue added to the dolichol phosphate, followed by additional sugar residues added by a set of reactions in the ER lumen. Once all the residues of the precursor are added to the dolichol phosphate, the oligosaccharide is flipped to the ER lumen and transferred to the Asn-X-Thr/Ser residue of the nascent proteins synthesized by the ribosome anchored at the cytoplasmic side of the ER. Protein-bound N-glycans are remodeled in the ER and Golgi by reactions catalyzed by glycosidases and glycosyltransferases. Glycosidases will hydrolyze the glycosidic bonds in the glycan structure causing the removal of sugar moieties while glycosyltransferases will build linear or branched N-glycans by adding specific monosaccharide moieties to the glycoprotein. Activated forms of monosaccharides are necessary for glycan synthesis, these are transferred from nucleotide sugars like UDP-GlcNAc, UDP- GalNAc, GDP-Man, GDP-Fuc, and CMP-Sia to acceptors like proteins or growing glycan structures. The glucose transporter (GLUT) family is one of the types of sugar transporters that will transport specifically activated sugars from the cytoplasm where most monosaccharide activation occurs to the Golgi where glycosylation maturation continues, to finally add the activated monosaccharides to the glycan (70). Calnexin and Calreticulin at the ER will ensure that only properly folded and assembled proteins will be transported from the ER to the Golgi to finally be released to the cell surface or their final cellular destination (70) **(Figure 3).** Overall, N-glycan synthesis is an established mechanism that remodels N-glycan structures added to the synthesized proteins on the Asn residue.

The final N-glycome profile on a glycoprotein is modified by glycosyltransferases, glycosidases, and transporters responsible for the biosynthesis of N-glycan structures which can be affected based on the physiological and biochemical changes occurring inside the cell.
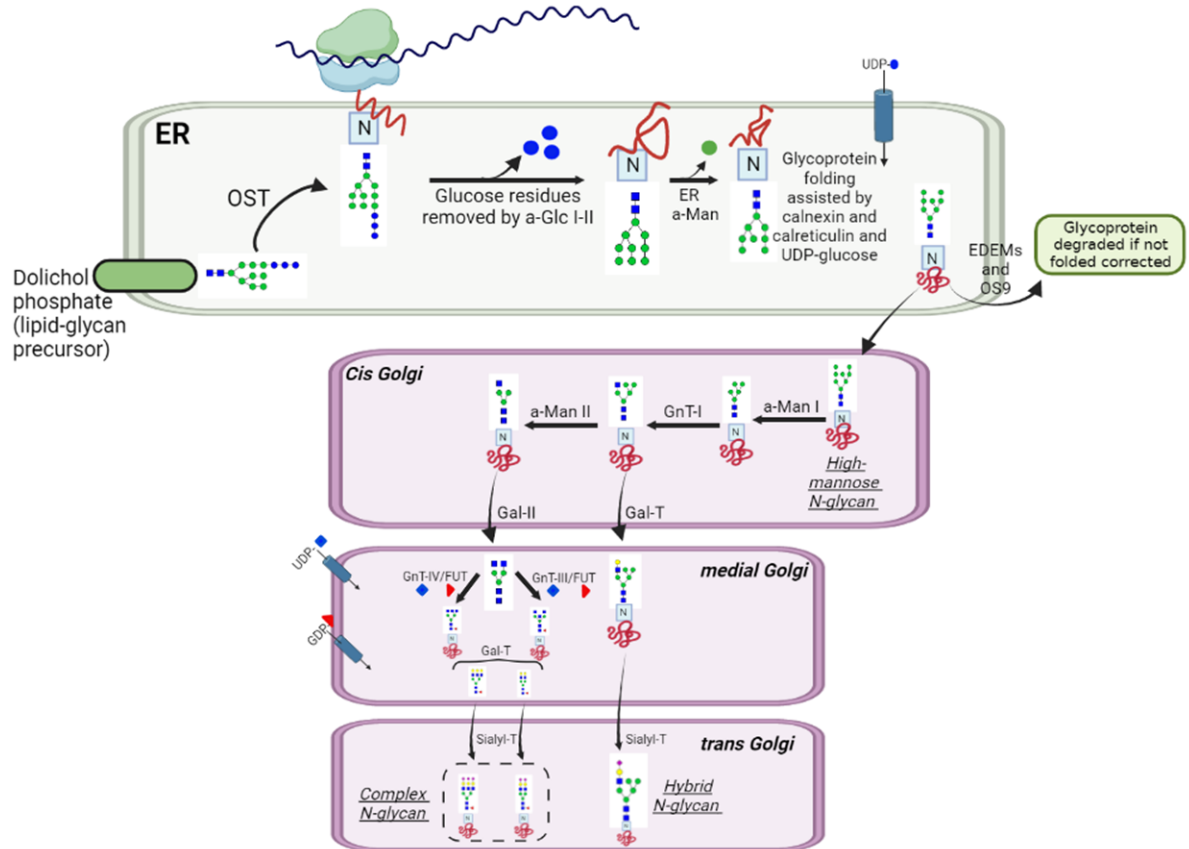
Figure 3. Schematic of the biosynthesis of N-glycans. The N-glycan synthesis starts at the Endoplasmic reticulum (ER) and continues into the Cis, medial, and trans-Golgi. OST (oligosaccharyltransferase), a-Glc I-III (a-glucosidases), ER a-Man (ER a-mannosidase) are all enzymes that help in the modification of the N-glycan structure. EDEMs (ER degradation-enhancing α-mannosidase I–like proteins) and OS9 (lectin) will collaborate for mannose trimming and degradation of the misfolded glycoprotein. Sugar donors activated form UDP-glucose, UDP-GlcNAc, and GDP-Fuc will assist with the nucleotide transport into the ER and/or Golgi. Glycosidases and transferases: GlcNAc-transferase I–IV (GnT-I–IV), β1,4 galactosyltransferases (Gal-T), α2,3 and a2,6 sialyltransferase, and Fucosyltransferases (FUTs) will further modify the N-glycans structure throughout the Golgi finalizing with the addition of terminal sialic acids.

16

1.3.2 N-glycan structure composition

During N-glycan synthesis and processing in the Endoplasmic Reticulum (ER) and Golgi, glycan structures are modified by the addition and removal of sugar residues. Eukaryotic N-glycans conserve the same core sequence at the Asparagine residue: Man3GlcNAc2, that is, 3 mannose residues attached to the core N-glycan composed of 2 GlcNAc residues attached by N-glycosidic bonds. From these core structures, N-glycans can continue to be modified into more complex, and matured N-glycans. These have been divided into three main categories: 1) Oligomannose (referred here as high mannose), which is the extension of only mannose residues from the core of the N-glycan. 2) Complex (including bi-antennary, tri-antennary, and bisecting), where the antennae after the mannose residue are initiated by a GlcNAc residue. 3) Hybrid where a mannose residue extends from the Man α1-6 arm of the core, and one of Man α1-3 arm. In addition to these three types of N-glycans, there are specific monosaccharides that can decorate and further modify any of these types of N-glycans: fucose (or fucosylation) or sialic acid (sialyation) in humans (NeuAc) or animals (NeuGc) **(Figure 4)**.
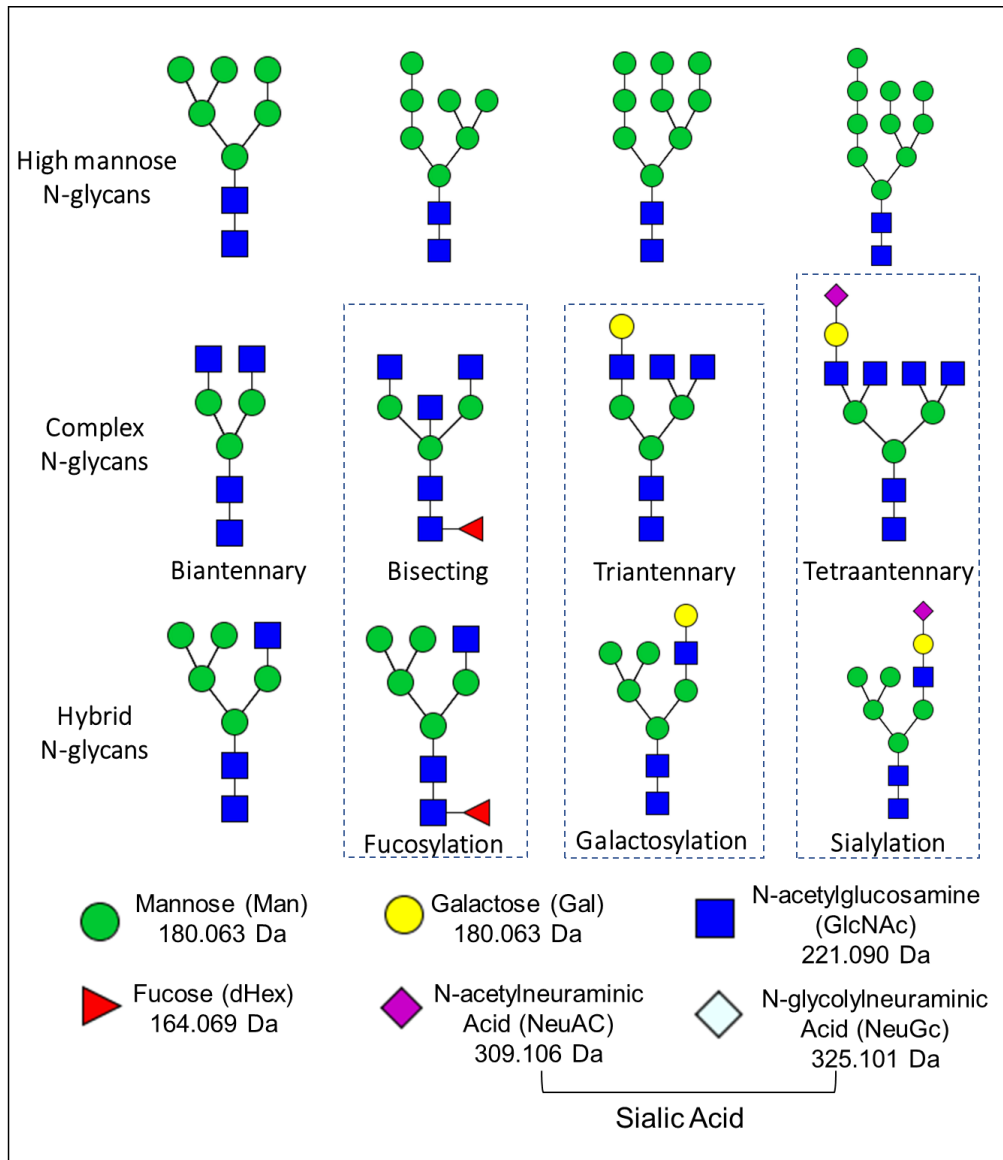
Figure 4. Representative N-glycan structures classes and compositions. The core of the N-glycan structure is characterized by two GlcNAc residues attached to three Man residues. The core structure will then be modified into a high mannose, complex or hybrid N-glycan. Complex N-glycans can be classified based on the number of antennas as biantennary, bisecting, tri-antennary, or tetra-antenary. Complex and hybrid N-glycans can be further modified by the addition of fucose, galactose, or sialic acid residues. NeuGcs sialic acids are only present in mouse-origin samples.

1.4 Glycomics and glycomic analysis

Glycomics is a comprehensive study that defines the structure of a glycan and its functions (71). Cell surface glycan structures can be attached to glycoproteins, glycolipids, and glycosaminoglycans and are produced by a cell type, tissue, biofluid, or organism (72). The glycome can be affected by numerous factors: the transcriptome, proteome, and/or environmental factors (65). Specifically, structure alterations in N-linked glycans (those attached to the asparagine residue of the glycoprotein) have been widely reported in response to many diseases, including cancer. Importantly, changes to the state of the cell during different cellular processes can be reflected in the N-glycome profile alterations at the cell surface; the analysis of these N-glycan structures (released from their protein carriers) alterations offers great clinical use for biomarker, diagnostic, and treatment strategies (73,74).

Methodologies to analyze N-glycosylation will vary according to the type of N-glycan analysis anticipated, but common techniques for the analysis of biological samples include the use of (1) lectins, carbohydrate-binding proteins that recognize specific residues of the N-glycan structure, these are used in immunohistochemical (IHC) or immunofluorescent (IF) assays; (2) High-performance liquid chromatography (HPLC), size-exclusion chromatography; and (3) Liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS) (70)(75). While each of these strategies has been valuable for N-glycan analysis, they also have limitations in specificity and extensive preparation protocols. In addition, the complexity and non-linearity of N-glycan structures have added challenges when characterizing N-glycan structures and compositions (58).

1.5 N-glycan alterations in liver disease

To address the lack of biomarkers for liver diseases (NAFL/NASH, HCC, and CCA), and improve the rates of early detection, the field of N-glycobiology has identified specific gene alterations as potential biomarkers. N-glycan modifications like fucosylation and branching have been reported as cancer-associated alterations in different liver diseases, and some are currently used as serum biomarkers (48,76–81).

1.5.1   Fucosylation

Fucosyltransferases are a subfamily of the large family of glycosyltransferases formed by 11 types of enzymes (FUT1-11) (82,83) **(Figure 5)**. Each FUT transfers an L-fucose moiety from the GDP fucose donor to the acceptor substrate (i.e., glycoproteins, glycolipids, or proteoglycans) with different linkages (84). The glycan modifications catalyzed by fucosyltransferases are the addition of fucose residues, which is known as fucosylation (82). Fucose alterations have been widely linked to different liver diseases and are recognized to be clinically relevant alterations for biomarker use in HCC (68,76). Serum glycoproteins develop abnormal glycosylation patterns during disease and because most of them originate from the liver, they offer the advantage to be used as biomarkers that can improve detection strategies at early stages and thus improve treatment options (58,85).

Proteins like fucosylated (Fc) Fc-GP73d, Fc-AFP, Fc-IgG were found to be hyperfucosylated in samples from patients with Hepatitis B-virus induced HCC (86). Elevated levels of the oncofetal glycoprotein, alpha-Fetoprotein (AFP) have been used as a biomarker to monitor liver disease progression but have been demonstrated to have low specificity (87). Instead, core fucosylated AFP-L3 was found to be specifically expressed in the serum of HCC patients, and glycoform L3 of the AFP protein also

20

demonstrated to have the highest level of core fucosylation compared to other AFP glycoforms. AFP-L3 has been approved by the US Food and Drug Administration (FDA) as a biomarker for the early detection of HCC (88).

Fucosylation has also been reported to be elevated in NASH patients when compared to a cohort of NAFL patients, specifically a tri-antennary, sialylated, fucosylated N-glycan structure on glycoprotein alpha-1-antitrypsin. The mechanism for this alteration was due to an inflammatory response based on the positive significant correlation of FUT6 with iL-6 expression in this cohort (89). Similarly, glycoprotein haptoglobin was also identified to be highly fucosylated in NASH patients relative to patients with simple steatosis. In a biomarker panel of 3 N-glycans proposed to distinguish between conditions, 2 N-glycans were fucosylated. The correlation discussed for this difference in fucosylation was based on a higher level of lobular inflammation and fibrosis in the NASH cohort (90).

Studies in CCA have also elucidated fucosylation-related alterations. A cohort that included samples from extrahepatic CCA, benign bile tract disease, and healthy control samples revealed that out of 13 N-glycans structures altered, one specific bi-fucosylated tri-antennary N-glycan had the best performance to distinguish extrahepatic CCA from the other groups (91). Fucosylated glycoprotein Fetuin-A was reported to have the ability to differentiate CCA from those with PSC and was proposed to have value in the surveillance of those at risk of developing CCA (92).

Overall, altered fucosylation in different liver glycoproteins or fucosylated N-glycan structures is one of the most common modifications reported in many liver diseases. Further studies would be needed to determine the type of fucosylation, the

type of patients this modification would be specific to, and the usefulness of this modification in a clinical setting.
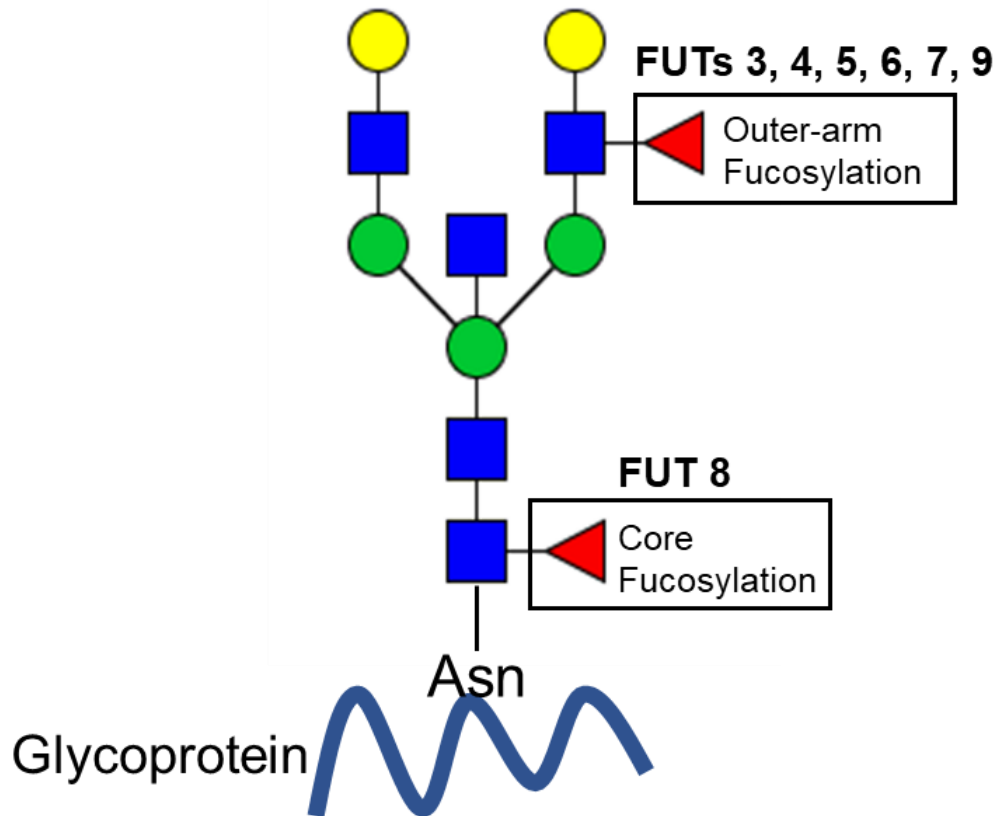
Figure 5. Schematic of the respective fucosyltransferases (FUTs) that catalyze the addition of the fucose residue, resulting in outer-arm fucosylation or core fucosylation.

1.5.1.1 Core Fucosylation

α1,6-Fucosyltransferse 8 (FUT8) is one of the most studied FUT members and is considered a disease-associate alteration in the liver. FUT8 is the only fucosyltransferase capable of adding a fucose product from GDP-fucose to the innermost GlcNAC residue at the 1,6 linkage of the glycoprotein, forming a core fucosylated glycan (70,93). (figure).

FUT8 upregulation has been reported in HCC tumor tissue, metastatic cell lines, and cells with proliferative, migratory, and invasive characteristics (82)(94). *Fut8* whole-body deletion in a mouse model treated with liver-specific carcinogen diethylnitrosamine (DEN) demonstrated more than 50% reduction of liver tumor incidence and significantly reduced tumor growth in xenografts compared to controls, suggesting that *Fut8* has an oncogenic role in liver disease progression (94). However, *Fut8* was demonstrated to be required to maintain liver regeneration abilities and hepatocyte proliferation after 70% partial hepatectomy (PH) (93). Similarly, in vitro studies using shRNA against *Fut8* reported a decrease in cells with migratory and proliferative capabilities (82). One of the mechanisms of how Fut8 can affect cells functions including EMT genes is based on the ligand-receptor binding affinity between Epithelial Growth Factor (EGF) and its receptor Epithelial Growth Factor Receptor (EGFR) consequentially affecting the phosphorylation of downstream signaling cascades including C-met, ERK, and AKT levels (94,95).

Core fucosylation alterations have also been elucidated in CCA serum. A study analyzed the N-glycan structural changes of CCA serum when compared to healthy controls, a core fucosylated tri-antennary glycan was significantly altered, while high mannose N-glycans were downregulated in the CCA cohort (80).

Human livers diagnosed with NASH were reported to have elevated *FUT8* gene expression while genes with roles in N-glycan quality control and precursor synthesis were downregulated in NASH livers compared to healthy or fatty livers (96). This study concluded that NASH could obstruct N-glycosylation processing resulting in increased FUT8 expression. Similarly, in a serum-based study looking at NASH, NAFL, and healthy controls it was found that the N-glycan structure that could differentiate NASH from the groups was a core fucosylated bi-antennary N-glycan (97). These studies highlight the association between core fucosylation and early liver damage.

1.5.2 Branching

N-acetylglucosaminyltransferases (MGATs) are a family of five specific glycosyltransferase glycoproteins (MGAT 1-5) (70) **(Figure 6).** Members of this family catalyze the branching of structures seen in all types of N-glycans (70). Each adds GlcNAc monosaccharides at different linkages during the initial steps of N-glycan processing to the final branching on the mature glycoprotein. MGAT1 and MGAT2 add the first GlcNAc moieties in the Golgi apparatus in the initial steps of N-glycan processing and can be further elongated by the addition of other moieties. However, MGAT3 adds a bisecting GlcNAc that cannot be elongated. Finally, the third and fourth branches are synthesized by MGAT4 and MGAT5 (68,70). Further elongation from GlcNAc branching by the addition of their sugar moieties results in hybrid N-glycans. However, branching varies depending on glycosyltransferase expression, localization of branching enzymes, the level of donor substrates, and tissue-specific roles of glycosyltransferases (88). Dysregulation of the branched N-glycan processing, MGATs, and glycan structure upregulations has been linked to a variety of cancers (98).
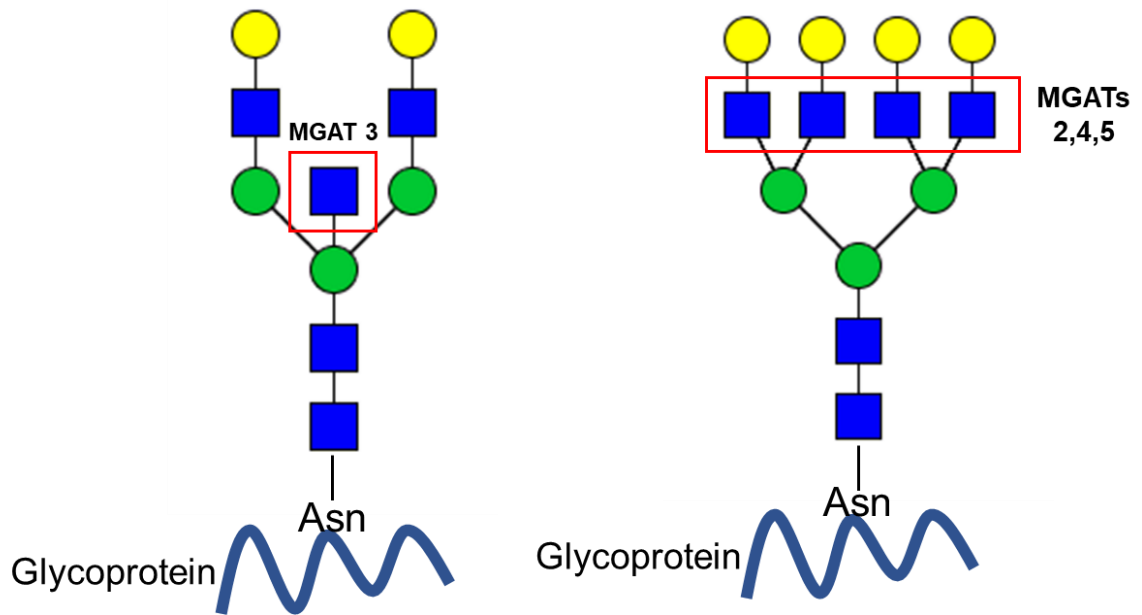
Figure 6. Schematic of the respective N-acetylglucosaminyltransferases (MGATs) that catalyze the addition of GlcNAc residues at different linkages of the N-glycan, resulting in a bisecting N-glycan (left) or a tri- or tetraantenary N-glycan (right).

1.5.2.1 Bisecting

N-acetylglucosaminyltransferase III (MGAT3) catalyzes the addition of a GlcNAc residue in β 1-4 linkage to the β- linked mannose already attached to the core of the N-glycan (70). The addition of this GlcNAc residue on complex and hybrid N-glycans is labeled as a bisecting N-glycan (70). The presence of MGAT3 inhibits the addition of more GlcNAc residues by other glycosyltransferases that are responsible for extensive branching such as MGAT5, it is hypothesized that MGAT3 competes with other enzymes to achieve this linkage (99,100).

Normal liver and hepatocytes express minimal MAGT3. However, MGAT3 expression in cancer varies and its role in cancer has been classified to be dependent on the tissue context. In the liver, inhibition of MGAT3 in a tumor-induced mouse model reduces liver tumors and progression (101,102). These results were explored further and the mechanism for liver tumor reduction accounted for a reduction in cell proliferation instead of cell apoptosis (103). A different study analyzed the effects of MGAT3 deletion in a DEN-treated mouse model and in addition to tumor suppression as previously reported, it was also found that haptoglobin was heavily glycosylated in the liver tumors of experimental mice (104).

However, inhibition of MGAT3 in a polyoma middle T (PyMT)-induced breast cancer model increased tumor growth and metastasis (105). Similarly, overexpressing MGAT3 demonstrated suppression in lung cancer metastasis (106). In a cell-based study, MGAT3 expression on EGFR was determined to inhibit EGFR/ERK signaling and consequentially inhibit malignant phenotypes of breast cancer (100). Some of these functions can be explained by the mechanism between MGAT3 and E-cadherin or other cell adhesion molecules. E-cadherin can upregulate MGAT3 levels and the

overexpression of MGAT3 retains E-cadherin at the cell surface, both mechanisms which can promote cell adhesion (107,108).

Overall, the role of MGAT3 HCC models, further studies are needed to establish the role of MGAT3, specifically in early and end-stage liver diseases.

1.5.2.2 Extensive branching

N-Acetylglucosaminyltransferase V (Mgat5 or GnT-V) catalyzes the final branching step on the glycoprotein by the addition of a GlcNAc residue at the β-1-6-linked branch from the α-1-3-linked mannose (70). Many studies have linked upregulated *Mgat5* gene expression and its glycan products β1-6 branching expression with metastasis, tumor invasion, and tumor growth in different HCC experimental models, and is considered as a disease signature for HCC (88,109–111). Studies in vitro further explored the association between Mgat5 expression and activation of signaling pathways commonly involved in cancer. PI3K/AKT signaling pathway was reported to be activated in HCC cell lines that had a high expression of CD147/basigin and GnT-V (112). CD147/basigin is a tumor-associated glycoprotein, heavily glycosylated with β1,6 GlcNAc branching modifications, in this study, pathway activation was linked to growth factor TGF-B (112). Other in vivo and in vitro studies have also linked Mgat5 upregulation with aberrant branching on EGFR and activation of EMT-related genes (113,114). Mgat5 analysis in tissue has also revealed similar findings; a study conducted in China with a cohort of 300 surgical specimens of patients undergoing hepatectomy revealed higher mgat5 expression in tumor tissue relative to non-tumor tissues and was associated with higher tumor classification, vascular invasion, and poor survival (111). This study classified Magt5 as an independent prognostic factor for survival and recurrence in HCC patients. Matrix-Assisted Laser Desorption Ionization (MALDI) N-glycan imaging mass spectrometry elucidated in situ

elevated expression of complex branching specific to HCC tissue compared to normal or cirrhotic tissue (76). This technology showed a direct association of branched glycans with tumor phenotypes at a late stage of the disease.

The role of Mgat5 has been studied more extensively in chronic liver injury, but there is some evidence of the disease mechanism in metabolic diseases like NAFL/NASH. NASH inhibits N-linked glycosylation processing and results in altered Mgat5 gene expression in human livers diagnosed with NASH compared to fatty or normal livers (96). However, Mgat5 overexpression in a mouse model fed a diet high in fat and cholesterol (HFHC) to induce NASH revealed a different response. Experimental mice developed less serious hepatic damage compared to control mice fed the HFHC. Specifically, a significant reduction in liver fibrosis was observed in experimental mice compared to control, this defense mechanism to inhibit NASH was explained by dysfunction in Hepatic Stellate Cells (HSCs) to express collagen when inducing steatohepatitis (115). The defense mechanism can be associated with Mgat5's roles in metabolism and its ability to effectively use and metabolize the excess of nutrients during high-caloric diets (116).

1.6 MALDI-IMS overview

Matrix-Assisted Laser Desorption Ionization (MALDI) is a soft ionization technique developed in the 1980s and is based on the conversion of analytes into ions by the addition or loss of protons (117). This is dependent on the use of a matrix, an energy-absorbent, organic compound that is mixed, spotted, or sprayed with the sample of choice. Upon laser irradiation, the matrix absorbs energy and transfers a portion to the analyte resulting in ionization. The sample now embedded in the matrix can be ionized with a laser beam, where the matrix will absorb most of the energy and be released from the surface (117,118). The matrix will transfer some of this energy to biomolecules in the

28

sample resulting in their ionization and desorption and creating singly protonated ions with minimal disruption to the sample of interest. These protonated ions are then separated based on their mass-to-charge (m/z) ratio. The correct selection of MALDI matrices is the critical factor to obtain better sensitivity, resolution, moderate matrix background, and signal-to-noise (S/N) ratio (119,120). Some of the matrices commonly used according to the biomolecule (proteins, oligosaccharides, lipids, or peptides) of interest include α-cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxy benzoic acid (DHB), and 1,5-diaminonaphthalene (DAN).

Ionized ions are then detected and measured by a mass analyzer, where ions will be applied kinetic energy by the electrical field and ions will travel through a drift tube and get separated by their different m/z values, heavier molecules will have a higher travel time relative to smaller molecules. Different types of mass analyzers include Time Of Flight (TOF) mass analyzers, Quadrupole (QTOF) mass analyzers, and ion trap mass analyzers. The ideal mass analyzer is selected according to the mass range, cost, speed, resolution, and sensitivity needed for the study. TOF mass analyzers are the most common type for biological applications including N-glycan analysis (117).

Finally, a detector at the end of the drift tube will measure each ion's flight time, and this information is transferred to a software package for analysis. The information generated from MALDI-MS is displayed in a vertical graph where m/z ratios are on the x-axis and the relative intensity is presented on the y-axis of the graph. Each bar in the graph will represent the m/z ratio that was elucidated based on their travel time and the length of the bar will represent the relative intensity of the ion. m/z ratio values can then be identified based on libraries generated from previous identifications (75,117,118,121).

While coupling MALDI-TOF and MS has provided an ample amount of information specified in its ability to analyze complex mixtures and generate understandable spectra, this still requires lengthy sample purifications and derivatization protocols (122). Instead, MALDI-TOF MS was coupled with imaging mass spectrometry (IMS) in 1997, and provided spatial information on the ions of interest within a biological tissue sample (118,123). One of the most valuable advantages of MALDI-IMS is that it can provide in-situ distribution of molecular information of the analyte of interest on the surface of the tissue section, this is of great value for clinical translational studies, where any disease human samples can be analyzed for the identification of biomarkers, diagnostic, and treatment strategies. MALDI-IMS works independently of antibodies or other labeling techniques and provides hundreds of analytes that can be identified from molecular imaging of post-translational modifications, lipids, and/or metabolites, all at a high resolution from a single tissue section and experiment (75).

Overall, the principle of MALDI-IMS is based on the coupling of three main techniques, (1) MALDI, (2) MS, and (3) IMS, where their combination provides a wide variety of advantages with clinical value, sample preparation, identification of analyte, specificity, and spatial analysis/identification of different biomolecules while also circumventing the disadvantages of their individual use.

1.6.1 MALDI-IMS for N-glycan analysis

MALDI-IMS has been incorporated for the quantitative and qualitative analysis of in situ of N-linked glycosylation of human and mouse model tissues and biofluids (77,124–126). MALDI-IMS N-glycan analysis works based on the same principle as previously mentioned, with the incorporation of exoglycosidase enzymes during sample preparation for the release of N-glycans from glycoproteins allowing their instrument

detection (127). One of the most common exoglycosidases used for N-glycan release is the bacterial enzyme peptide-N-glycosidase F (PNGase F), this enzyme will remove oligomannose, hybrid, and complex N-glycans attached to the Asparagine (Asn) residue. Since this enzyme functions as an amidase, where releasing the N-glycan attached to the nitrogen of Asn and converting Asn to aspartic acid (Asp) will provide the opportunity to elucidate sites of glycosylation based on the amino acid sequence before and after the treatment with the enzyme (70). A second commonly used enzyme is Endoglycosidase F3 (Endo F3), characterized by specificity in releasing bi and tri-antennary core fucosylated N-glycans between the two GlcNAc residues within the core of the N-glycans (128). The use of Endo F3 allows for the identification of only core fucosylated N-glycans, catalyzed by the enzyme previously mentioned, FUT8. Mass spectrums generated from MALDI-IMS will consist of m/z ratio values that are representative of specific N-glycan structures with their respective relative abundance. Proposed N-glycan structures are identified based on the mass of monosaccharide residues of the N-glycan structure and according to the possible combinations according to the N-glycan biosynthetic pathway. Generated data can be analyzed using a software package where mass spectra and corresponding 2D heatmap images are included for simultaneous analysis. m/z values that correspond to specific N-glycan structures can be selected for further analysis and their relative intensity can be visualized using a color gradient scale in the corresponding 2D heatmap images. In addition, MALDI-IMS provides valuable clinical translational information, where 2D heatmap images that correlate to specific N-linked glycoforms can be overlayed with different immunohistochemical stains to elucidate the N-glycan alterations with histopathological alterations in the tissue (Figure 7). Spatialomics N-glycan characterization tissue

imaging overlayed with histological disease-related characteristics can provide

information on possible targets for disease biomarkers and treatment (124)(129).

The use of MALDI-IMS for the study and characterization of N-linked

glycosylations provides information about the localization, distribution, structure, and

relative abundance of hundreds of N-glycoforms within the preparation of a single tissue

section.
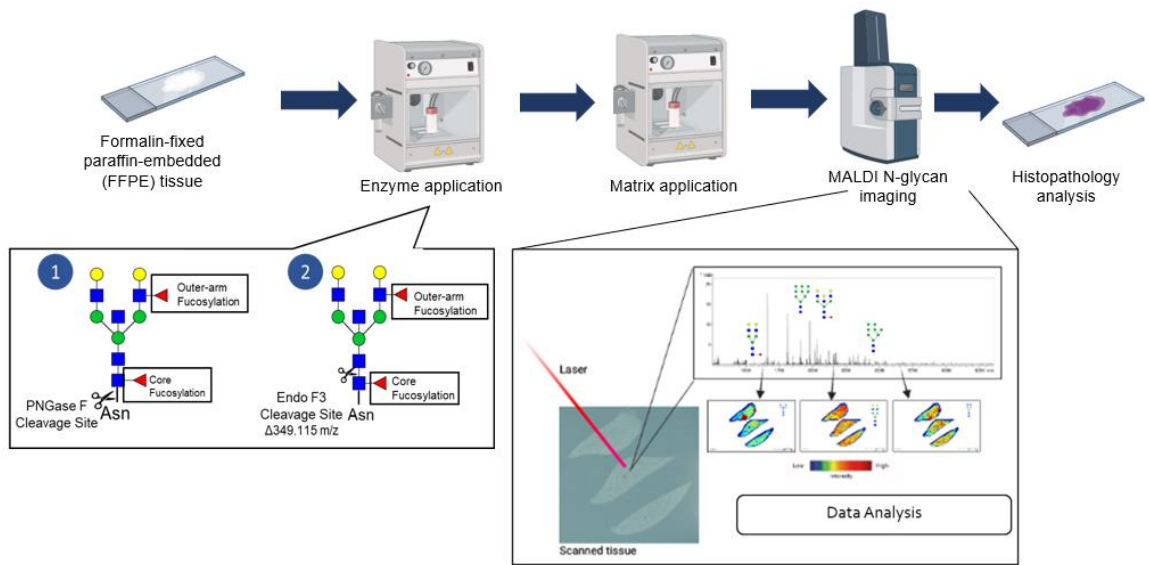
Figure 7. MALDI-IMS sample preparation workflow for paraffin-embedded tissue samples using exoglycosidases PNGaseF and/or EndoF3, followed by matrix application and finally imaging samples. Tissue sections can be stained for histopathological analysis after the MALDI-IMS protocol.

1.6.2 Mouse to human N-glycan studies relevance

Mouse models for biomedical research, specifically for liver-related studies, are the preferred species due to their anatomical, physiological, and genetic similarities to humans (130,131). The use of animal models has allowed researchers to induce a wide variety of human diseases that can mimic many of the characteristics observed in humans. Other advantages of using animal models include access to an unlimited number of samples, control of experimental conditions, fast breeding, and small size (130). Mouse and human livers have very similar functions like nutrient processing, protein production, and detoxification, but there are some specific differences like location, number of lobes, and generation of fibrous connective tissue, and while these differences are taken into consideration in experiment planning and data analysis, these are considered small differences in quantitative research relative to the many other similarities between species (132,133).

The applicability of N-linked glycosylation studies in mice to processes observed in humans has been reported in different cell types and tissues, confirming the relevance of using animal models to elucidate N-glycan-related modifications in humans (134–136). Similarly, other studies in different transgenic mouse models have reported important findings in the function of individual N-glycans, classes of N-glycans, and N-glycan-related genes in different cells, tissue, and organ functions (101,105,113,115,137–141). N-glycan structures, N-glycome alterations, and N-glycan-related mechanisms are known to be conserved from mouse to human (64,142). However, there is an important exception in the type of sialic acids that are expressed in mice and humans. Humans have a unique inability to form NeuGc sialic acid, due to a germline inactivating mutation in the enzyme that converts NeuAc to NeuGc (CMP-

NeuAc hydrolase), this results in a difference in the type of sialic acid in mouse (NeuGc) and humans (NeuAc) (134,143).

Overall, the use of mouse models for N-glycan-related studies can be considered ideal models to provide great advancements in comparative research by elucidating mechanisms, target genes, and other modifications that could be of interest to human disease clinical strategies.

1.7 Broad Overview

In conclusion, CLDs can include a wide variety of diseases that are also considered risk factors for the progression of primary liver cancers. Many patients with CLDs do not present any symptoms and are diagnosed at the late stages of the disease. In addition, to the silent progression of early liver diseases, current screening strategies to identify early diseases are also lacking. These limitations result in the diagnosis of liver disease at very later stages like cirrhosis and cancer. The ideal solution would be to identify biomarkers with high sensitivity and specificity to be used to screen a population of patients with risk factors for the respective CLD.

N-liked glycosylation is a common post-translational modification that can get altered by different internal and external factors, these alterations at the cell surface are also known to expose the status of the cell based on any molecular changes. The idea that N-glycan structure alterations are occurring at the cell surface involves a wide variety of N-glycan-related enzymes like glycosidases, glycosyltransferases, nucleotide donors, and transporters. Many of these enzymes have proved to be needed for development, cell homeostasis, proliferation, tissue regeneration, etc. While the role of some of these genes has also been studied in a liver context, the impact of many of these N-glycan modifications is still not understood in liver disease and much less in the

progression from CLDs to primary liver cancers. This research gap is in part due to (1) limitations in experimental models, current models are not able to mimic specific stages of liver disease and reliable transgenic models for gene-targeted investigations are limited. (2) identification of the best N-glycan gene to target and study, N-glycan studies have reported alterations in different N-glycan genes, and many have been labeled as disease-related modifications, but from this many modifications is challenging to identify a single modification that could be inducing the disease. (3) common N-glycan techniques, require extensive sample preparation, and while many of these techniques have provided important information for the field, there is now a need to expand this analysis and identify a more complete N-glycan profile that can also be relevant for clinical use.

MALDI-IMS for N-glycan analysis offers a valuable advantage not only in the identification of a complete N-glycan structures profile but also in the ability to determine the origin of these modifications within the tissue. MALDI-IMS requires minimal sample preparation that does not require homogenization of the sample allowing for the maintenance of morphological information. Representative images of each N-glycan structure modification can be overlayed to histological information (i.e, fibrotic, inflammation, tumor areas) and elucidate the spatial localization of modifications in tissue.

The ideal solution for the lack of biomarkers for progressive liver diseases and to address the high rates of end-stage liver disease is the elucidation of biomarkers that can identify early liver diseases. Currently, N-linked glycosylation modifications are used in the clinic for liver cancers and have been demonstrated to be a promising field for

biomarker discovery. However, further research is needed to identify N-glycan genes of interest that could be having a role in the different stages of liver disease progression.

To address, the research gaps previously mentioned, we hypothesized that by leveraging MALDI-IMS N-glycan analysis in a controlled experimental mouse model for the different liver stages, we could identify specific N-glycan modifications according to the liver disease to be targeted for further elucidation of their liver-specific roles. The use of MALDI-IMS would allow for the in-situ analysis, which is important due to the many cellular processes taking place during liver disease progression, the identification of the specificity of an N-glycan to a cellular process and/or cell type would provide vital information for deciding the best modification to target.

Diet and carcinogen-inducible mouse models were chosen with specific time points that would recapitulate different stages of liver disease progression. N-glycan analysis by MALDI-IMS coupled with histological information by immunohistochemistry and/or pathological annotations would provide specific modifications that could be involved in the progression of the liver disease, Finally, the N-glycan-related gene responsible for this modification would be targeted in an in vivo model for further elucidation of their liver-specific role. In general, the characterization of N-glycan modifications in liver disease progression, from CLD to liver cancer, and the generation of an in vivo model for mechanistic investigation are described in this dissertation.

Chapter 2: Hypothesis

2.1 Introduction

The current prevalence of chronic liver disease (CLD), including all etiologies and all disease stages, is estimated to be 1.5 billion worldwide, with non-alcoholic fatty liver disease (NAFLD) accounting for 59% of these cases, driven by high obesity rates (2). NAFLD includes a spectrum of different liver complications, such as steatosis, inflammation, and fibrosis; all of which are known to be risk factors for the progression to end-stage-liver diseases often in the form of Hepatocellular Carcinoma (HCC) and Cholangiocarcinoma (CCA).

N-linked glycosylation is one of the most common post-translational modifications (PTM) where N-glycan structures are added to the cell surface or secreted proteins at the Asparagine (Asn) residue. N-liked glycan modifications, specifically core fucosylation, have provided great value for the elucidation of specific alterations that occur in cancer. In addition alterations in glycosylation have played roles as biomarkers for liver cancers, AFP-L3 for HCC, and CA19-9 for CCA.

Fucosyltransferase 8 (FUT8), which catalyzes core fucosylation, has been linked to increased fibrosis and has been classified as an HCC N-glycan signature (58,76,87,144). However, few studies address the role these N-glycan modifications play during disease progression to liver cancer, especially during NAFLD up to liver cancer formation. A major reason for this is the lack of controlled models where we can track and recapitulate different stages of liver disease while also identifying spatially, the origin of the N-glycan modifications through the progression of the disease.

To address current limitations, we propose to study all stages of liver disease progression (steatosis, inflammation, fibrosis, end-stage disease, and liver cancer) and identify N-glycosylation signatures and their origin. We will use human samples and

established diet and carcinogen-induced mouse model systems coupled with in situ N-glycan imaging methods to characterize N-glycan alterations at each stage of the disease. N-linked glycosylation is known to be a dynamic process and is modified in response to the state of the cell affected by environmental insults (i.e., diet, metabolic syndromes), similarly, these same insults are risk factors for NAFLD. Thus, we hypothesized that N-linked glycosylation is altered during early liver damage and throughout the progressive disease. The proposed research will (1) identify specific N-glycan modifications that could be useful for biomarkers strategies, (2) precisely define localized spatiotemporal N-glycosylation alterations, with a focus on core fucosylation, as correlated to pathologically defined regions, and (3) elucidate N-glycan modifications that should be targeted in vivo for mechanistic studies. Overall, the proposed study will provide fundamental information on liver disease progression in the context of N-linked glycosylation for biomarkers and therapeutic strategies.

2.2 Specific Aims

<u>Specific aim 1</u> will characterize the N-glycan profile in Nonalcoholic Steatohepatitis (NASH) and determine the relevance between N-glycan changes in mouse models and human disease. Diet-induced mouse models and human NASH liver biopsies will be analyzed for N-glycosylation alterations. Similarly, the N-glycan origin within the tissue will be correlated to histopathological changes. The completion of this aim will result in a full N-glycan characterization of NASH taking into consideration the histological changes in mice and humans. This will represent the first such study of this kind.

<u>Specific aim 2</u> will characterize the temporal changes in N-glycosylation from early liver disease to liver cancer. Established NASH and HCC-induced mouse models will be used

to induce different stages of liver disease by selecting specific time points, which will allow for the examination of liver disease progression. CCA will also be addressed in this aim by the use of human tissue and serum since the induction of CCA in a mouse model has been challenging. Another important goal of this aim is to correlate the histopathological changes to N-glycosylation alterations in each stage of disease, including fibrosis, inflammation, and tumor development. Our proposal will be able to correlate previously established N-glycan modifications to temporal liver disease changes from a normal liver up to liver fibrosis and/or HCC/CCA. Completion of this aim will provide a complete characterization of all stages of liver disease for the discovery of N-glycan signatures as biomarkers.

Specific aim 3 will develop a liver-specific in vivo approach to study N-linked core fucosylation. *Fut8* will be targeted for conditional deletion in the liver, using a Cre/Lox system under the Albumin promoter. This aim focuses on the generation of the first and only hepatocyte-specific deleted *Fut8* mouse model to elucidate the role of *Fut8* in liver diseases. Completion of this aim will be a critical advancement in the fields of molecular biology, glycobiology, and pharmacology to provide important knowledge on the value of FUT8 to be used as a biomarker or target in different types of cancers.

2.3 Goals and impact

The overall goal of this study is to identify and elucidate the role of N-glycan signatures, specifically core fucosylation during different stages of NAFL/NASH and HCC progression. Current models to study N-glycosylation changes in the liver rely mostly on human serum and HCC tissue samples which are often limited and do not allow to answer specific questions. The proposed study uses an *in vivo* approach to address unanswered

biological questions and generates an *in vivo* model to study the importance of core fucosylation/FUT8 in the liver. It is noted that the floxed FUT8 mouse could also be used to generate cell-specific deletions that may be of great value in other cancers, outside of the liver (145–147). That is because alterations in core fucosylation have been observed in many other cancers but have not been examined using a murine knock-out model. A long-term goal is to solve the lack of early diagnostic and treatment strategies for HCC by exploring FUT8 as a clinically relevant strategy.

2.4 Research Strategy

2.4.1 Significance

Chronic Liver Diseases (CLD) accounted for 1.32 million deaths in 2017, affecting 440,000 females and 883,000 males worldwide (3). Non-alcoholic Steatohepatitis (NASH) is the most common etiology for CLD, followed by hepatitis B and C virus (HBV and HCV), and chronic alcohol consumption. Very common pre-existing conditions like obesity or Type II diabetes are major risk factors for NASH. Non-Alcoholic Fatty Liver (NAFL) is the first stage of the disease characterized by hepatic steatosis followed by the progressive form, NASH characterized by hepatic inflammation, and fibrosis (148,149). Without intervention, in some patients, NASH can progress to cirrhosis and potentially Hepatocellularlar Carcinoma (HCC) or Cholangiocarcinoma (CCA) (49,148,150). Indeed, the high rate of NASH patients is increasing the incidence rate for primary liver cancer, consequently increasing liver transplant cases and liver disease-related deaths (33,149). To improve current treatment and diagnostic strategies, a greater molecular understanding is needed in the formation and progression of progressive liver diseases. Due to the limitations in the treatment strategies for advanced-liver diseases, the best approach for improving outcomes in those with liver disease is early disease detection.

However, currently, there are no direct noninvasive markers for NASH at any of the stages of the disease.

The field of N-linked glycosylation has offered to be a promising strategy since immune responses, cell proliferation, and cell signaling activation are related to N-glycan alterations (70,72,74,98). N-linked glycosylation has been demonstrated to be helpful in blood-based biomarkers for HCC and CCA (79,91,151–153) and we have previously shown that alteration in N-linked glycosylation can be observed in patients with NASH. However, most NASH N-glycan-related studies, including our own (77,89,97) do not take into consideration the different disease characteristics and instead address NASH as a complete disease. Additionally, the progression from NASH to primary liver cancer is not addressed. Therefore, this proposal is significant in that it will be the first study to couple N-glycan in situ imaging with the induction of all stages of liver disease in a controlled manner to characterize N-glycan alterations at each stage and correlate this to the disease-related characteristics within the tissue. This characterization will identify N-glycan signatures that could be used as biomarkers for the detection of a specific stage of liver disease and address the lack of strategies for early detection.

2.4.2 Approach

Specific aim 1: Characterize the N-glycan profile in human and mouse Nonalcoholic Steatohepatitis (NASH)

We and others have established that during liver diseases like fibrosis, CCA, and HCC, N-glycan structures are significantly altered (48,76,80,91,154). In addition, we have demonstrated that these N-glycan modifications are specific to histopathological areas of tissue, such as fucosylation and extensive branching in tumor vs nontumor areas in HCC (48,76). Here, we aim to address the research gap in the knowledge of N-

glycan alterations in early liver diseases, like NASH. As mentioned before, NAFLD includes liver events like steatosis, inflammation, and fibrosis, we propose the use of a NAFL/NASH diet-induced mouse model with human NASH samples to identify N-glycan modifications at each of these events by a novel mass spectrometry technique, matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI). The use of MALDI-MSI will allow studying in-situ N-glycan modifications in liver tissues, this is accomplished by the use of enzymes to release N-glycans followed by matrix application allowing the N-glycans to maintain their localization to then be ionized and measured by mass spectrometry. Once imaging is complete, the same tissue slide will be processed for Hematoxylin & Eosin (H&E) staining to overlay the 2D heatmap images produced by the analysis and correlate glycan changes to histopathological changes. The completion of this aim will 1. Elucidate N-glycan alterations in NASH before there is evidence of cancer. 2. Addresses the clinical relevance of using NAFLD mouse models by using the use of mouse and human tissues.

Specific aim 2: Characterize N-glycosylation changes in liver disease progression: from early liver disease to primary liver cancers.

Specific aim 2.1: Characterize the temporal changes in N-glycosylation in NASH and HCC-induced mouse models.

N-glycosylation modifications like fucosylation, sialylation, and branching have been characterized in different types of liver diseases (79,90,97,153,155). The correlation of these modifications with different characteristics of the disease has provided insight into biomarker discovery. However, the specific timing of when these modifications occur during liver disease progression is still unknown. In general, current

reports lack information on the N-glycan modifications in the progression from a healthy liver to CLD to end-stage liver diseases. Due to these limitations, we propose the use of NASH and/or HCC mouse models, where stages of liver damage are possible by the collection of samples at different time points (from normal liver to severe liver damage/HCC). We propose to use paraffin-embedded liver tissues to do a full N-glycan analysis profile throughout the tissue and specific to in situ pathological changes by MALDI-MSI. For this aim, we propose to use wild-type male C57BL/6J mice fed a Western Diet (WD, 40kcal% fat and 20kcal% fructose) to induce NASH and Diethylnitrosamine (DEN) hepatocarcinogen treatment 50mg/kg to induce HCC. Two different models will be induced 1. NAFL/NASH model (Timepoints: day 0, month 2, month 4, month 6, and month 9) 2. NASH-induced HCC model (Timepoints: day 0, month 4, and month 11.5). Timepoints are hypothesized to represent a healthy liver, the beginning of liver damage, moderate liver damage, and severe liver damage. The completion of this sub-aim will 1. Elucidate N-glycan signatures with each stage of liver damage: steatosis, inflammation, fibrosis, and tumor formation within the same study, 2. Identify the first N-glycan alterations responsive to early liver damage, and 3. Determine if the modifications observed in cancer can be detected before cancer develops.

Specific aim 2.2: Identify promising biomarker strategies based on N-glycosylation changes in Cholangiocarcinoma tissue and serum.

NASH has been identified as a risk factor for the development of cholangiocarcinoma (CCA), specifically intrahepatic cholangiocarcinoma (iCCA) (23,156,157). Since NASH is on the rise, most likely leading to an increase in CCA cases, biomarker strategies for CCA are urgently needed. Cancer-associated glycoprotein CA19-9 is the established serum-based biomarker for CCA, but lacks

specificity and is altered in other types of cancers and bile duct diseases. Here, we propose the use of N-glycan structures on these glycoproteins, instead of the use of the glycoproteins themselves for biomarker discovery. Due to the limitations of inducing iCCA in a mouse model, we will use human CCA tissue and serum for N-glycan analysis and biomarker identification. Tissue samples from patients with iCCA, HCC, other liver conditions, and healthy will be analyzed for N-glycan imaging by MALDI-IMS. Simultaneously, a serum sample set including the same groups will be analyzed for the N-glycan modifications identified in the serum using a serum-based N-glycan imaging method. N-glycan information identified in tissue and serum will be analyzed in combination with clinical information to obtain the best biomarker strategy that can distinguish patients with CCA from patients with other types of liver diseases. The completion of this sub-aim will (1) elucidate the biomarker clinical value of N-glycan alterations identified in tissue and their correlation to what is observed in serum, (2) identify specific N-glycan modifications to be used as promising biomarker strategies for the detection of iCCA.

Overall, based on the identification of specific N-glycan structures, aim 2 will provide information on glycosyltransferases responsive to the first liver insults and identify potential targets for further investigations on their role in the initiation and progression of liver disease.

Specific aim 3: Generation of an in vivo approach specific to the liver to study N-linked core fucosylation.

As previously mentioned, fucosylation, specifically core fucosylation, catalyzed by the enzyme Fut8 is one of the most common N-glycan alterations observed in liver

diseases (82,153,158,159). Previous work demonstrated a reduction in liver tumor incidence, formation, and decreased cell proliferation in a Diethylnitrosamine (DEN) treated whole body *Fut8* knockout mice (94). Based on these results, *Fut8* was proposed as a prognostic marker and therapeutic target for HCC. This study initially established the involvement of FUT8 in HCC (94). However, mice with *Fut8* deletions suffered significant phenotypic alterations including embryonic lethality, limiting the identification of FUT8 liver-specific functions (160). This model has also never been made available to the wider scientific community and has remained in the country in which it originated, Japan. We propose the generation of a hepatocyte-specific *Fut8* knockout mouse model by the LoxP/Cre strategy. To generate a conditional *Fut8* knockout, exon 9 will be flanked by two LoxP sites, and an Albumin Cre-mediated excision will result in a hepatocyte *Fut8* deletion. Mice will be genotyped by PCR using established primers to confirm a hepatocyte-specific deletion after Albumin-Cre breeding (FUT8$^{-/-}$). Following the completion of aim 1, we expect to have obtained a liver-specific conditional deleted *Fut8* model. Since *Fut8* expression is low under normal conditions in the liver, we expect no breeding issues or lethal phenotypes in the *Fut8$^{-/-}$* model. Future studies using this model would be used to determine if alterations in FUT8 and core fucosylation previously reported are a byproduct or driver of liver damage when used in a liver disease-related model.

2.4.3 Conclusions

We believe that the successful completion of these three aims will allow us to have a deeper understanding of the N-linked glycosylation changes that occur in the progression of liver disease. Including a correlation of these alterations to histopathological changes in liver tissues will provide greater insight into the role they

play in the pathology. Based on the proposed aims, we aim to address a lack of information in areas that are vital to help the field advance in terms of biomarker identification for early and advanced liver disease and also the role these glycans play in disease progression. The completion of this proposal will open opportunities for the field to study N-glycosylation and cancer taking a mechanistic approach and advancing toward clinically relevant strategies. A very important advantage of the generation of a *Fut8*$^{f/f}$ mouse model in Aim 3, is that this mouse model could be used for any desired tissue deletion with the use of the Cre-excision model to study different types of cancers in which *Fut8* has been associated, including pancreas, breast, and colon.

Chapter 3: Experimental mouse models to investigate liver diseases

3.1 Liver disease induction

Due to the overwhelming high rates of chronic liver disease diagnoses worldwide and the lack of predictive and diagnostic tools for early detection of liver disease, the use of animal models has become a crucial tool to gain more insight into the mechanisms of disease, elucidate possible therapeutic targets and/or to test novel drugs. The wide variety of options to induce liver injury provides the ability to investigate the pathophysiology of a specific type of liver damage based on the goal of the study. Current experimental animal models are capable of imitating diseases like alcoholic and non-alcoholic steatohepatitis, advanced chronic liver disease, hepatocellular carcinoma (HCC), and cholangiocarcinoma (CCA) (161–172). This chapter will focus on the different types of mouse models available based on liver disease, focusing on the liver diseases mentioned throughout this dissertation.

3.1.1 Mouse models for Non-alcoholic Fatty Liver Disease (NAFLD)

The experimental model selected to study NAFLD will depend on the stage of the disease of interest: that is early stages like fatty liver disease (NAFL), the progressive form of the disease nonalcoholic steatohepatitis (NASH), or a full spectrum of the disease (NAFLD). It is important that the chosen method can induce similar risk factors to what is observed in humans (i.e., obesity, type 2 diabetes, metabolic disorders) and a similar histological profile according to the type of liver injury including steatosis, inflammation, and fibrosis.

3.1.1.1 Diet inducible models

One of the most common methods to induce any of the stages of disease mentioned above is overnutrition. Different combinations of oral administration of diets with high caloric contents have proven to induce most of the stages of the disease.

Dietary models focus on Ad libitum feeding of a high-fat diet (HFD), choline or

methionine-deficient diet (CDD or MDD), fructose, and/or a combination of diets. Table 1

provides an overview of NAFLD-diet-induced models commonly used and the level of

the disease each can induce.

Normal diets available at mouse facilities for colony maintenance have a calorie

content of 10kcal% of fat. To induce very early stages of fatty liver disease, an HFD is

used, characterized by more than 60 kcal% of fat contents. Feeding this diet for at least

20 weeks has been demonstrated to induce mild steatosis with minimal inflammation

and fibrosis. Similarly, a diet high in fructose might induce some hepatocyte ballooning

and mild fibrosis but no steatohepatitis. HFD and high fructose diets by themselves are

considered to not be appropriate models of NASH(162).

To induce a more advanced stage of NAFLD and most likely achieve a full

spectrum of the disease, diet combinations have been demonstrated to be the best

models. A western diet (WD) or a "fast-food diet" characterized by 40% energy fat, 12%

saturated fat, and 2% cholesterol in addition to fructose is a successful dietary model for

NAFLD(168,173–178). In humans, a western diet tries to recapitulate highly processed

diets, fast foods, sugar-sweetened beverages, and a lack of fruit and vegetables. Similar

to the effects observed in humans, a WD is known to induce obesity, insulin resistance,

high levels of steatosis, inflammation, and fibrosis by 25 weeks of feeding. One limitation

of this diet is that it does not completely recapitulate the diverse fat sources and

micronutrients that are normally included in a human WD diet. However, a WD is a well-

established method capable of inducing NAFLD clinically and histologically.

A second diet combination accepted to induce NAFLD is choline and methionine-

deficient diets (CDD or MDD). Since choline is required for the export of triglycerides via

very low-density lipoprotein (VLDL) packaging, deficiency causes steatosis. Similarly, methionine is an essential amino acid and an intermediate in the synthesis of S-adenosylmethionine and glutathione, deficiency causes oxidative stress leading to inflammation and fibrosis (179). The combination of CDD/MDD has been demonstrated to induce high levels of steatosis, inflammation, and fibrosis by 10 weeks of feeding. One limitation of these diets is that mice do not develop obesity or insulin resistance and mice can suffer from Cachexia within 8 weeks of feeding, which is not a relatable phenotype observed in humans (168,177).

Another method commonly used for mice to ingest high caloric contents is by adding cholesterol, fructose, or trans-fat to their water in addition to the chosen high-caloric diets. Different percentages of these additives are used and trying to replicate what is equivalent in humans has been a challenge when using this method (168).

Overall, the idea of the success of diet combinations to induce a full pathophysiological profile including fibrosis is based on nutritional intervention in addition to a proinflammatory hit.

While some dietary models can be strain-dependent, most studies that are diet based only have been done using C57BL/6J mice, as this strain is known to be susceptible to diet-induced obesity and type 2 diabetes (161). Sex differences have also been established in primary NAFLD between males and females. Human NAFLD prevalence is known to be higher in men than women, specifically premenopausal women. This can be explained by an estrogen protective mechanism that protects the liver from fibrosis in NAFLD. Studies in postmenopausal women or those with estrogen deficiency have been associated with severe liver fibrosis (180). Similar sex differences have been observed in mouse models in different mouse strains and under different diet

52

combinations, correspondingly, male mice are preferred to be used when the main goal of the study is to induce the full spectrum of NAFLD (181). However, the idea that the full spectrum of the disease can be recapitulated only in male mice should also be considered a limitation of the model since further research is still needed to understand why there is a higher prevalence of NAFLD and risk of advanced disease in post-menopausal women (13,182). Overall, female mice should still be incorporated in NAFLD studies to address hormonal and genetic intrinsic differences, and for these in vivo studies to be relevant in pre-clinical studies.

3.1.1.2 Chemical inducible models

Based on the idea that the progression of NAFLD is driven by multiple environmental factors throughout life that could be inducing liver disease-a "multiple-hit" theory, this is recapitulated in a mouse model by the combination of inducing obesity by diet in combination with a chemical agent. Table 1 provides an overview of NAFLD-hepatotoxin-induced models commonly used and the level of the disease each can induce. Carbon Tetrachloride (CCL_4) is a potent hepatotoxin agent known to induce oxidative stress to the liver, and lipid peroxidation leading to inflammation and resulting in a necrotic response with steatosis (183). CCl_4 administration in multiple doses by peritoneal injection by itself induces fibrosis but no obesity and is not considered a NALFD model, instead if used in combination with a high caloric intake has been demonstrated to induce chronic oxidative stress leading to inflammation, hepatocyte ballooning, apoptosis, and finally stage 3 fibrosis at 12 weeks (184). Treatment with CCl_4 is considered a good model to study liver fibrosis with characteristics that resemble human liver fibrosis including liver regeneration, fiber formation, and collagen type 1α1 expression (185,186). A model that can induce a high level of fibrosis is valuable

because one of the main challenges with other mouse models is the inability to recapitulate the same level of fibrosis as what is observed in humans. Discontinuation of CCl_4 doses will regress fibrosis while long-term use can worsen histological profiles of the disease (165,184). Overall, CCL_4 works as a potentiator and accelerator of the effects of diet leading to the full progression of liver disease, including liver cancer. One of the disadvantages of this model is the lack of research on the effects and role of CCl_4-derived metabolites in terms of human NASH pathogenesis (168).

The use of Streptozotocin (STZ), a commonly used Type 2 diabetes model, in combination with an HFD has been useful as a NAFLD model. STZ is a glucosamine-nitrosourea derived from *Streptomyces achromogenes* that has toxicity specifically to pancreatic B cells (187). Administration of this agent intraperitoneal or subcutaneously after birth leads to chemical inflammation and the destruction of pancreatic islets. Disease-related factors can worsen with time, at 4 weeks there is simple steatosis and by 8-12 weeks progressive pericellular fibrosis has been reported (166).

3.1.1.3 Genetically inducible models

Mouse models with mutations in appetite-regulating genes like leptin-deficient (ob/ob), deficient leptin signaling (db/db), and Apolipoprotein deficient (ApoE$^{-/-}$) are some of the most used models for early liver disease and NAFLD. These models will develop steatosis under a normal diet, while high-caloric diets will lead to NASH. One of the main differences and advantages of these genetic models relative to dietary models is their ability to develop a higher level of obesity and their propensity to metabolic complications similar to what is observed in humans. However, each model has their limitations: ob/ob model will only develop steatosis even in an HFD for up to 52 weeks, db/db model requires an MCD to promote steatohepatitis with fibrosis, and ApoE-/-

model does not always reproduce pathobiology of NASH(162). Table 1 provides an

overview of NAFLD-genetically modified-induced models commonly used and the level

of the disease each can induce. Overall, choosing the correct genetic model will depend

on the NASH profile desired including metabolic implications.

Table 1. Overview of mouse models for Nonalcoholic fatty liver disease (NAFLD).
This table was adapted from reviews (162,168,183)

Models for NALFD		Duration/treatment	Obesity	Steatosis	Inflammation	fibrosis	NASH	Comments
Diet models	HFD	At least 20 weeks	Y	mild	N	Very low	N	NASH after 36-50 weeks of exposure
	High fructose	16-24 weeks (normally used in combination)	Y	mild	N	N	N	plasma triglyceride and glucose levels
	WD	25 weeks	Y	high	high	high	Y	A greater amount of cholesterol relative to human consumption
	CDD/MDD	4-10 weeks advanced disease	N	high	high	mild	Y	No metabolic features were seen in human NAFLD
	HFD + fructose	16-30 weeks	Y	high	mild	mild	Y	Fructose added to drinking water
	HFD + cholesterol	16-20 weeks	Y	high	mild	mild	Y	Cholesterol content is not physiologically relevant to humans
Chemical models	HFD + CCl₄	12-52 weeks HFD in diet and agent by i.p or inhalation	Y	high	high	high	Y	Derived metabolites have unknown roles in NASH
	HFD + STZ	8-12 weeks STZ treatment to neonatal mice	Y	high	high	high	Y	Do not mirror the etiology and natural history of NASH. Steatosis and inflammation decrease over time
Genetic models	Leptin (ob/ob⁻ᐟ)	4-12 weeks	Y	high	N	N	N	Needs a second stimulus (chemical or diet) for the development of NASH
	Leptin receptor gene (db/db⁻ᐟ)	4-12 weeks	Y	high	N	N	N	Needs a second stimulus (chemical or diet) for the development of NASH
	Apolipoprotein (ApoE ⁻ᐟ)	12 weeks	Y	mild	N	N	N	When combined with HFD will develop NASH
	DIAMOND mice + HFD	16 weeks (steatohepatitis)-36 weeks (fibrosis) / Inbred isogenic strain: 60% C57Bl6/J 40% S1929S1/svlm	Y	high	high	high	Y	Only if combined with HFD and glucose/fructose in drinking water will develop NASH

3.1.2 Mouse models for primary liver cancers

Progression from NAFLD to advanced chronic liver diseases like cirrhosis are some of the main risk factors for the development of disease to liver cancer. Hepatocellular Carcinoma (HCC) and Cholangiocarcinoma (CCA) are considered the main primary liver cancers. An ideal experimental mouse model for induction of HCC/CCA should be able to mimic a full immune cell response, and chemokine signaling, and recapitulate genetic mutations. Another important factor is that mouse models should develop tumors rapidly while also maintaining the histological features observed in the respective cancer.

3.1.2.1 Chemical inducible models

One of the most common strategies for liver cancer induction is the use of hepatic substances, these allow for the characterization of the full liver disease progression, that is from initiation, promotion, and progression(188). Table 2 provides an overview of liver cancer-hepatotoxin-induced models commonly used and the level of the disease each can induce

3.1.2.1.1 Diethylnitrosamine (DEN)

Hepatic carcinogen Diethylnitrosamine (DEN) is a well-established method for inducing liver cancer in mice. The irreversible effects of DEN can vary according to the dosing, age of injections, administration frequency, administration route, and combinations with diet or other chemical agents.

DEN is metabolically activated in hepatocytes by enzymes of the cytochrome P450 family leading to different DNA modifications and generation of reactive oxygen species (ROS) to finally induce oxidative stress and DNA damage (188,189). This process leads to the alteration in the expression of genes involved in stress responses

that will stimulate hepatocellular purinergic receptors. Finally, the increased proliferation of hepatocytes with DNA damage mutations will predispose the liver to HCC (190).

Different studies have administered DEN intraperitoneal injections at 2 weeks of age at a concentration of 25mg/kg body weight and have reported HCC by 8-10 months (188). A combination of DEN and a high-caloric diet like an HFD can promote tumor development by inducing cell damage and oxidative stress and activating pro-fibrogenic signaling and tumor-promoting cytokines(188). This will provide the advantage of a major progression of disease while also allowing to study of HCC in the background of NAFLD which is known to be one of the major inducers of HCC incidence in humans. Tumor incidence when using DEN doses ranging from 25mg/kg up to 100mg/kg in combination with HFD at 2 weeks up to 8 weeks of age can vary from 3 months up to 10 months (188). The idea of administering DEN treatment at juvenile age like 2 weeks old is due to the high hepatocyte proliferation rate at this stage leading to a rapid increase of affected cells (188).

One of the main differences between DEN-induced tumors in mice and humans is that the genetic pathway and transcriptomic profile are different from what is seen in HCC humans. Mouse HCC tumors that are induced by a single dose of DEN administrations are known to be mutated in 80% in *H-ras* or *B-raf* and 20% in *Egfr*. While transcriptomic analysis of human HCC tumors has identified expression signatures in HBV, TP53 signaling, and WNT or AKT pathway activation (191). However, *H-ras* and *B-raf* have been associated with metastasis and poor prognosis in human HCC. A very important similarity of DEN-induced HCC tumors to human tumors is their histological comparison in their heterogeneous microenvironment, low and high-grade dysplastic nodules, and HCCs identification grade 1 to 2 (188).

58

DEN-induced HCC can also be dependent on mouse strains, FVB or C3H/HeJ are known to develop HCC 2-3 months earlier after DEN administration, relative to C57BL/6, BALB/c, or 129 (192). Like HCC prevalence in humans, DEN-induced HCC is gender-dependent, about 30% of females treated with DEN for long periods will develop HCC, while 100% of males will develop HCC. The mechanism that explains this protective mechanism has accounted for hormonal changes in females where androgens will stimulate interleukin-6 (IL-6) production in Kupfer cells while this mechanism is inhibited by estrogens. Induction of pro-tumorigenic transcription factors NF-kB and STAT-3 are known to result from an increase in IL-6 in male mice (193–195). Maintaining a consistent mouse strain, age, and sex are important factors to avoid experimental variability that would not be related to the experimental goals.

Models to induce HCC are more established relative to CCA, DEN in combinations with Thioacetamide (TAA) and bile duct ligation have been used to induce CCA. However, these models have demonstrated a higher incidence of HCC, limiting the study of CCA subtypes, similarly, bile duct ligation is considered a demanding procedure and not considered a well-established method (196).

3.1.2.1.2 Phenobarbital (PB)

PB is commonly used for the initiation of liver carcinogenesis in combination with DEN treatment. The mechanism of PB to induce tumor incidence is by increasing the expression of cytochrome p450 complexes resulting in the activation of mutagenic metabolites (197). PB is more commonly used in combination with DEN, with the idea to support two-step liver carcinogenesis. This combination has also been helpful when DEN is administered in adult mice, where hepatocyte proliferation is infrequent, and PB can reinforce toxicity and induce tumor incidence.

59

3.1.2.2 Genetically inducible models

Some of the most commonly used strategies include genetically engineered mouse (GEM) models: with constitutive gene expression systems and/or conditional and inducible gene expression systems. GEM allows the investigation of the role of a specific gene, alterations in a signaling pathway related to apoptosis, proliferation, and angiogenesis during the progression of cancer, and even allow for elucidation of the target gene in development. These systems are based on modifications of the expression of oncogenes, and tumor suppressor genes. Constitutive transgenic mouse models where the mutation is present throughout the body were one of the first strategies to evaluate liver cancer. However, this model can suffer from embryonic death, is not specific to the tissue of interest, and fail to mimic the process of tumorigenesis. While the use of conditional expression systems can circumvent most of those disadvantages and is preferably used in liver cancer models (198). Table 2 provides an overview of liver cancer-GEM-induced models commonly used and the level of the disease each can induce.

3.1.2.3 Implantation models

Implantation models are one of the most widely used models to induce liver cancer formation, this method is based on the implantation of HCC/CCA cell lines or tumor tissue fragments. The recipient mice could be a syngeneic tumor model (implanted sample originated from the same strain as the recipient mice), or more preferably a xenograft model (recipient mice are immunodeficient). Immunodeficient mice have a deficiency of T lymphocytes and impaired T and B cell function, this model has been the preferred strategy because it allows the use of any origin cell line and tumor tissue (i.e., HCC human cell lines, carcinogen-induced tumors, or tumors from

GEM models). However, this model is not useful when the interest is in an immune response profile.

Since immunotherapy has demonstrated some success in HCC patients, a more reliable method to understand immune profile responses is allograft models. This model is not the same as the syngeneic model, instead, murine cell lines or tumors are implanted into immunocompetent mice. A second model that allows the study of a full immune response in tumor development is humanized mouse models. In these models, human immune cells (haematopoietic stem cells and precursor cells isolated from fetal cord blood) are engrafted into immunocompromised mice to induce a human immune system and are then injected with HCC tumor cells. Humanized models allow to study of immunotherapy-related responses but they are difficult to establish and have a risk in the engraftment of the human immune system (170). Table 2 provides an overview of liver cancer-implantation-induced models commonly used and the level of the disease each can induce.

Implantation models can be divided between ectopic (subcutaneously injection in any site) or orthotopic (injection into the same site from where the implanted material was derived) sites. Ectopic injections are inexpensive, have rapid tumor formation, and can easily measure tumor size non-invasively but are not useful for microenvironmental studies, while orthotopic injections are better for metastatic studies and provide tumor microenvironment information, these are also preferred for liver cancer studies(172). Overall, the use of implantation models varies depending on the origin of the graft and host and the anatomical location of the implantation (198).

Table 2. Overview of mouse models for primary liver cancers. This table was adapted from reviews (163,170,172,196).

Models for primary liver cancer		Duration /treatment /strain/sex	Type of liver cancer	Comments
Chemical models	**Diethylnitrosamine (DEN)**	12-15 days of age i.p. (25 mg/kg) C57Bl/6 male	More than 80% HCC incidence in males in 8-10 months	A higher concentration of DEN is required in older mice (4-5 weeks of age)
	DEN + HFD	2 weeks of age i.p. (25 mg/kg) HFD chow pellets C57Bl/6 male	HCC incidence in males in 8-10 months	Mice develop NAFLD-associated HCC
	DEN + Phenobarbital (PB)	DEN: 4 weeks of age 100mg/kg (i.p.) PB: daily dose 0.07% (in drinking water)	HCC incidence in 7 months	Mixed CCA and HCC incidence in some animals
	TAA	300 mg/L in drinking water Sprague-Dawley rats	iCCA incidence in 16–22 weeks	Model limited primarily to rats
	Dimethylnitrosamine (DMN)	Induced liver fluke infection 0.0025% DMN in drinking water Syrian Golden Hamsters	100% CCA incidence in 15 weeks	Also induces gastrointestinal tumors, skin, lung, and hematopoietic tumors
Genetic models	**c-myc model**	Tissue-specific Alb-Cre C57BL/6 X CBA/J males	65% HCC incidence at 20 months	Mutations are present during embryogenesis, a time-controlled deletion would be preferred to avoid activation of compensatory pathways.
	TGF-α	MT promoter CD1 strain males	50% HCC incidence in less than 12 months	
	E2F-1	Tissue-specific Alb-Cre C57BL/6 X CBA/J	33%-60% HCC incidence at 12 months	
	***Smad4* and *Pten* model**	Tissue-specific Alb-Cre *Smad4* and *Pten* embryonically lethal	CCA tumor development by 4-7 months of age	Mice also develop HCC tumors. Tumors develop in the absence of inflammation
	***Kras-Pten* model**	*Kras* activation and *Pten* deletion with Alb-Cre	Well-differentiated iCCA by 46 days	The model mimics symptoms of human iCCA (jaundice and weight loss) Lethal by 46 days of age
Implantation models	**Xenograft models**	**HCC/CCA cells** are injected directly into the liver (orthotopic)	HCC/CCA: High tumor incidence, reproducibility, and modest cost	The host needs to be immunocompromised. The model cannot mimic an immune response
		PDX model (human tumors)	iCCA: Similar morphology, histology, and immunohistochemical profile as in primary patient tumor	The host needs to be immunocompromised. PDX models can be expensive, low-rate efficacy, and time-consuming
		CCA organoids Primary CCA grown in 3D in vitro culture- "tumoroid"	High engraftment rate with high tumor incidence. Similar histological profile of the primary CCA they were derived from	Models do not mimic inflammatory and stromal microenvironments during in vitro or in vivo growth

3.1.3 Conclusions

Based on the previous studies presented here, there is a wide variety of models that can be used for studies related to NASH/NAFLD and primary liver cancers. For the studies presented in this dissertation, the main goal was to use NASH and HCC mouse models that would allow the investigation of the progression of liver diseases, specifically NAFL, NAFLD, and NAFLD-induced HCC. For studies related to only fatty liver disease, a High Fat Diet (HFD) was selected, this diet was used to investigate very early liver damage, like steatosis. For studies, where the goal was to observe a more advanced disease like NASH/NAFLD, a high caloric-combination diet, western diet (WD) was chosen to induce a full spectrum of disease. As mentioned before, this diet induces metabolic alterations and all histological changes that are commonly seen in human NASH/NAFLD. For studies related to NAFLD-induced HCC, hepatotoxin Diethlynitrosamine (DEN) was used with a combination of a western diet. This model has been used to induce hepatic tumors in at least 80% of mice, in addition to the profile of NAFLD (steatosis, inflammation, and fibrosis). Different treatment time points were assigned according to the level of disease desired, based on previous literature, we hypothesized the induction of fatty liver, NASH, NAFLD, chronic liver disease, and HCC. Since previous studies have reported that sex hormones, menopausal status, and reproductive information are influencers in the effects caused by diet and carcinogens (19,194,199–201). In this initial study, we focused on male mice to induce effects that were only dependent on treatments and analyze molecular changes at each stage induced by these factors only. However, further studies should include female mice to evaluate their protective mechanism in all stages of liver disease. Finally, we used C57BL/6 mice instead of genetic models to induce liver diseases, since C57BL/6 mice

are an established strain and to avoid other effects that could affect our analysis due to

an unknown relationship between gene deletion and N-linked glycosylation.

Chapter 4: N-Linked Glycosylation characterization in Nonalcoholic

Steatohepatitis in Mouse and Human

Imaging Mass Spectrometry Reveals Alterations in N-Linked Glycosylation That Are Associated With Histopathological Changes in Nonalcoholic Steatohepatitis in Mouse and Human

Shaaron Ochoa-Rios, Ian P. O'Connor, Lindsey N. Kent, Julian M. Clouse, Yannis Hadjiyannis, Christopher Koivisto, Thierry Pecot, Peggi M. Angel, Richard R. Drake, Gustavo Leone, Anand S. Mehta, and Don C. Rockey

Corresponding Authors: Shaaron Ochoa-Rios ochoaari@musc.edu

Anand Mehta: mehtaa@musc.edu

4.1 Introduction

Non-Alcoholic Fatty Liver Disease (NALFD) affects about 25% of people in the United States, as the liver manifestation of the metabolic syndrome, is expected to continue rising along with obesity(9). NAFLD is defined as an excess of fat stored in the liver, (>5%) without any non-metabolic causes, e.g., alcohol or viral hepatitis(14). While NAFLD is most often considered a common and benign condition, about a third of patients with NAFLD progress to NASH, a silent liver disease distinguished from NALFD by the presence of inflammation, hepatocyte injury, and fibrosis (167,202) which has become the number one cause of liver transplant in women in the United states and the second leading cause in men(203). About 10-15% of NASH patients progress to cirrhosis and/or HCC (204). Currently, there are limitations for NAFLD/NASH diagnosis as liver biopsy continues to be the standard to distinguish fatty liver disease from advanced stages of disease like NASH (97,203). Changes in the N-glycan profile of a glycoprotein can occur in response to genetic and environmental factors including NAFLD/NASH(205). Glycosylation is one of the most common post-translational modifications, with more than 50% of proteins being glycosylated (58,206). Many of these glycoproteins are known to have functions in the regulation of signaling pathways, protein folding, and immune responses (70,207). The addition of oligosaccharide moieties or glycans at the Asparagine (Asn) residue of the cell surface or secreted protein is known as N-linked glycosylation (68,70,208). Terminal N-glycan modifications like increased fucosylation, branching, and sialyation are considered disease-signatures and have been linked to gastrointestinal and liver disease including liver fibrosis, liver cirrhosis, NAFLD/NASH, gastric cancer, colorectal cancer and HCC (88,205,208,209). Similarly, glycosyltransferases and glycosidases that add and remove specific

monosaccharides, respectively have also been identified to be upregulated in various diseases (68,76,88,144,209).

Alterations in N-glycosylation with disease progression have been exploited in the clinic with the use of glycosylated Alpha-Fetoprotein (AFP) in serum as an HCC biomarker and have demonstrated an immense potential to be used as diagnostic and therapeutic biomarkers for other liver diseases including metabolic disease like NAFLD/NASH. Similarly, research has focused on using the serum to identify N-glycan modifications or glycoproteins with clinical relevance to differentiate between stages of simple steatosis and NASH (89,90,97,204,210). However, the origin of these N-glycan modifications within liver tissue is poorly understood; NASH is a progressive disease known to start with liver steatosis and progress to severe liver damage caused by advancing stages of fibrosis (7,18). The ability to identify the histopathological origin of N-glycan modifications would be of great value for biomarkers development.

To address the lack of information on in-situ N-glycosylation modifications, we utilized the advantages of matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS) to quantitively and qualitative study of the localized N-glycan changes on NAFLD/NASH mouse models and NASH human paraffin-embedded liver tissue.

In our NAFLD/NASH mouse model we identified a complete N-glycan profile where the level of liver disease was driven by diet; since the daily excess of calories in processed foods, especially those high in fructose can induce severe liver injury resulting in NAFLD/NASH (22,177,211–214). Moreover, we validated our findings through the analysis of human liver biopsies of patients with NAFLD/NASH and report a

correlation of fucosylated glycan structures with the level of fibrosis based on the N-glycan expression and fibrotic histopathological tissue. Here, we characterize the N-glycome alterations and their origin in NAFLD and NASH tissue. The N-glycan modifications and their correlation to histopathological profiles of the disease reported here could contribute to non-invasive serum biomarker development in early liver diseases like NAFLD and NASH based on fibrosis.

4.2 Methods

4.2.1 Diet and mouse model

All animals received humane care according to the criteria outlined in the Guide for the Care and Use of Laboratory Animals. Animal studies were approved by the university's Institutional Animal Care and Use Committee (IACUC). Mice were housed and maintained under normal husbandry conditions Experimental mice (10th generation FVB/NT) were fed standard chow until weaning (21 days of age). Male mice were randomly selected at weaning and placed in each group for respective diets. Three types of diets were fed ad libitum: Low-Fat diet with 10 kcal% of fat, high-fat diet with 60 kcal% of fat (Research Diets), and western diet with 40 kcal% of fat, 40 kcal% of fructose, and 2% cholesterol (Research Diets-D09100301). Mice were sacrificed and tissue was harvested at 18 months of age. Dissected livers at necropsy were frozen in liquid nitrogen and/or frozen in dry ice-cooled heptane or fixed in 10% formalin for histology. Abnormalities were based on observations upon euthanasia and tissue collection, no scoring system was used for this analysis. Tissue abnormalities detected by observation and measurement were quantified to determine the level of damage induced by the respective diets in whole liver tissue and other major tissues. A tissue abnormality was considered when the tissue had discoloration, lesions, lipidosis, or increased in size.

4.2   Histology and scoring

Formalin-fixed paraffin-embedded (FFPE) mouse tissues were processed and
stained for Hematoxylin and Eosin (H&E) (Cancer Diagnostics) and Masson's Trichrome
stain (Polysciences Kit 25088) using standard protocols. Heptane-frozen liver tissues
were used for Oil Red O staining (ORO, VWR kit 95027-208) to quantify lipids. Blinded
histopathological analysis by an independent pathologist in H&E staining was done
following the scoring systems explained in Table 3. For immunohistochemistry, an
antibody against CD3 (Abcam 16669) was used following the manufacturer's
instructions. Quantifications were done using the Vectra Polaris Imaging system. For
Picrosirius red (PSR) staining, polarized view images were taken with a 20x objective
and quantified using the image analysis Nikon software. Mouse liver tissues used for
mass spectrometry studies were first processed for glycan analysis protocol followed by
H&E staining. Human NASH biopsies used for steatosis analysis were given a fibrosis
score of 0 by an independent pathologist and considered low steatosis if no lipid droplets
were observed in the whole tissue scan and considered high steatosis if lipid droplets
were observed. Images, where whole tissues were analyzed with specific stains, were
scanned using Hamamatsu Scanner and exported from Hamamatsu NDP software.

4.2.3 RNA Extraction and qPCR

RNA was extracted from snap-frozen liver tissues from an 18-month-old cohort
for qPCR using TRIzol reagent protocols. RNA was separated from the solution after
chloroform treatment, next RNA was precipitated with isopropyl alcohol, and finally,
isolated RNA was washed in 75% ethanol and the RNA pellet was resuspended in
Hypure water. cDNA was synthesized by reverse transcription using random hexamer
primers and qScript cDNA SuperMix. qPCR reaction was performed on a QuantStudio™

5 Real-Time PCR System (Thermo Fisher) using SYBR-green Supermix (Bio-Rad). Reactions were normalized to *Gapdh* using the $\Delta\Delta Ct$ method. Primer sequences:

Fut8   *CAGGGGATTGGCGTGAAAAAG*

     CGTGATGGAGTTGACAACCATAG

Gapdh   *CGGTGTGAACGGATTTGGC*

     TTTGATGTTAGTGGGGTCTCGC

4.2.4 Human biopsies

The study involving human samples was approved by MUSC's Institutional Review Board and met all guidelines set forth by the 1975 Declaration of Helsinki for good clinical practice. A convenience sample of 51 NASH liver FFPE biopsy specimens that represented the full range of fibrosis stages was analyzed. Biopsies were assessed for fibrosis by an independent pathologist using the NASH Clinical Research Network (CRN) scoring system. Samples were divided by fibrosis scores (FS): 8 samples with FS 0, 8 samples with FS 1, 9 samples with FS 2, 15 samples with FS 3, and 11 with FS4. Fibrosis was also assessed using histomorphometric analysis of the collagen area as described (215). In brief, PSR staining was performed on formalin fix liver biopsies and collagen content was quantified by the area percent In Image J (National Institutes of Health) (216).

4.2.5 Enzymes and Reagents

Trifluoroacetic acid, Harris-modified hematoxylin, and α-cyano- 4-hydroxycinnamic acid (CHCA) were obtained from Sigma- Aldrich. HPLC grade methanol, ethanol, aceto- nitrile, xylene, hydrogen peroxide, and water were obtained from Fisher Scientific. Recombinant peptide N-glycosidase F (PNGase F PRIME™) and

endoglycosidase F3 (Endo F3 Prime™) were provided by N-Zyme Scientifics (Doylestown, PA). Both enzymes were expressed and purified as previously described (128,217).

4.2.6 Tissue Preparation for MALDI-IMS

FFPE liver tissues from mice and human livers (biopsy specimens) were processed using standardized glycan imaging workflows of matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-IMS) as described (124). In brief, after sample dewaxing and antigen retrieval steps, samples were treated for enzymatic release of N-glycans by treatment with PNGase F PRIME™ using an M5 TM Sprayer (HTX Technologies LLC). Next, tissue slides were incubated at 37°C for 2 hours in prewarmed humidity chambers followed by desiccation before α-cyano-4-hydroxycinnamic acid (α-CHCA) matrix application. For analysis of core fucosylated glycans, Endo F3 Prime™ was applied first and analyzed as previously described (128,218). In brief, the slides were placed in 100% ethanol for removal of the matrix and then placed in a series of dilutions of ethanol (95 and 70%). Next, the slides were placed in a high-pH cleaning solution (10 mM Tris, pH 8.98), HPLC grade water, then a low-pH cleaning solution (citraconic buffer, pH 3), and then HPLC grade water again. The slides were then desiccated and dried. Following the cleaning, the tissues were then prepped for PNGase F PRIME™ application by following the same tissue preparation and glycan imaging protocol as previously described; however, the dewaxing and antigen retrieval steps were omitted, beginning with enzyme application on the tissue with Endo F3 Prime™. Finally, samples were imaged using: (1) A timsTOF Flex trapped MALDI-QTOF mass spectrometer (Bruker Daltonics) (m/z 500–4000) operating in a positive mode for mouse liver tissues and human NASH biopsies with Endo F3 Prime™

treatment; images were collected at 100 µm raster with 200 laser shots per pixel. Images

and (2) A RapifleX TissueTyper MALDI TOF mass spectrometer (Bruker Daltonics) (m/z

600–3,500) operating in a positive ion mode for human NASH biopsies with PNGase F

PRIME™ treatment only; images were collected at 50 µm raster with 200 laser shots per

pixel.

4.2.7 Data Processing and Analysis

Data analysis was done in SCiLS Lab 2020 imaging software (Bruker) for

analysis of the mass range m/z 500 to 4000. SCiLS-generated N-glycan spectra were

normalized to the total ion count (ICR Noise Reduction Threshold = 0.95), which were

then matched within ±5 ppm against an in-house database of known N-glycans

generated using GlycoWorkbench and GlycoMod for annotation (63), or to MS/MS data

previously done by our group (219). Maximum mean values for each m/z value were

used to examine the quantitative and spatial expression of glycan changes based on the

type of diet and disease stage. All data obtained by MALDI-IMS were correlated to FS

previously established by pathology. In our glycan analysis by MALDI-IMS, structural

assignments of glycan structures were made based on mass values and rules of glycan

synthesis. Specifically, a chemical composition of four or more N-acetyl hexosamine

(HexNAc) was assumed to be tri-antennary glycans while more than four N-acetyl

hexosamine (HexNAc) with two N-acetylgalactosamine (GalNAc) were considered

bisecting glycans.

4.2.8 Experimental Design and Statistical Rationale

A total of 40 mice were used in the NAFLD/NASH mouse model, mice were

divided as follows: LFD: 16 mice, HFD: 14 mice, and WD: 9 mice. To determine the

significance of liver abnormities compared to no abnormalities, a Chi-Square test was

used. Mouse liver tissue stainings (ORO, and PSR) and immunohistochemistry (CD3) used 3-8 mice per diet. Pathologist scoring assessment (lipid accumulation, inflammation, and fibrosis) used: LFD: 16 mice, HFD: 12 mice, and WD: 8 mice. A Mann-Whitney test was applied to determine if there were significant changes between diets. All human NASH biopsies were from patients treated at South Carolina hospitals during a time range of 1-3 years. Our research laboratory obtained all samples at the same time and processed them for MALDI-IMS experiments within 2 months of receiving them. Samples were deidentified and an identification number was given to each sample based on the order they were registered in our laboratory database. In addition, all files were loaded into a single analysis file and normalized to the total ion count of all runs independent of different acquisition times between samples.

For MALDI-IMS N-glycan experiments, mouse n=3-5 and human n=51. SCiLS Lab was used to generate maximum mean values for all tissues after normalizing glycan intensity values to total ion current, which were subsequently exported to statistical software for analysis. SCiLS analysis was performed by groups/treatments: Mouse tissue samples, human biopsies (PNGase F PRIME™), and human biopsies (EndoF3 + PNGaseF). For Endo F3 Prime™ analysis, the maximum mean value was normalized to the natural logarithm of a number (LN). A total number of 11 NASH biopsies were used: 3 with an FS0, 5 with an FS2, and 6 with an FS4. All of the maximum mean values here were based on the intensity of the whole tissue instead of specific regions or colocalization analysis. The fibrosis scores correspond to the entire tissue slide based on the scoring system done by an independent pathologist before any of the MALDI-IMS experiments were done. To determine statistically significant changes in glycan expression for all analysis presented here, maximum mean values were evaluated using

a Mann-Whitney test assuming nonparametric distribution where a p-value less than 0.05 was considered statistically significant.

4.3 Results

4.3.1 Mouse NASH phenotypes

To examine NALFD and NASH *in vivo*, mice were fed three different diets including the following (162,173,177,213,214,220): (1) a low-fat diet (LFD) with 10 kcal% of fat, a high-fat diet (HFD) with 60 kcal% of fat, or a Western diet (WD) with 40 kcal% of fat, 40 kcal% of fructose, and 2% cholesterol; these diets were started at weaning and mouse livers were harvested at 18 months of age. Diets high in fat are known to induce a NAFLD phenotype, while high-fat diets combined with fructose have been reported to additionally induce inflammation and fibrosis consistent with NASH (167,183,212). Mice fed a WD had a higher number of liver abnormalities (*defined in Methods*) than mice fed a HFD or LFD, while mice fed a HFD had more abnormalities in other tissues (i.e., enlarged spleen, eyes, and pancreas masses) (Figure 8A and 8B). Several age-related abnormalities were also observed in mice fed a LFD (such as enlarged bladder and corneal opacity) (Figure 8B). In addition, mice fed a HFD gained more weight over the 18-month period compared to WD and LFD (Figure 8C). When liver weight was normalized to total body weight (221), the liver/body weight ratio was significantly increased in mice fed a WD compared to the control diet and a HFD (Figure 8D).

Lipid accumulation, inflammation, and fibrosis were measured by specific staining of liver tissue and blinded qualitative scoring using the University of Pittsburgh Medical Center (UPMC) histological system for components of NAFLD activity score (NAS) and fibrosis staging (222) (Table 3).
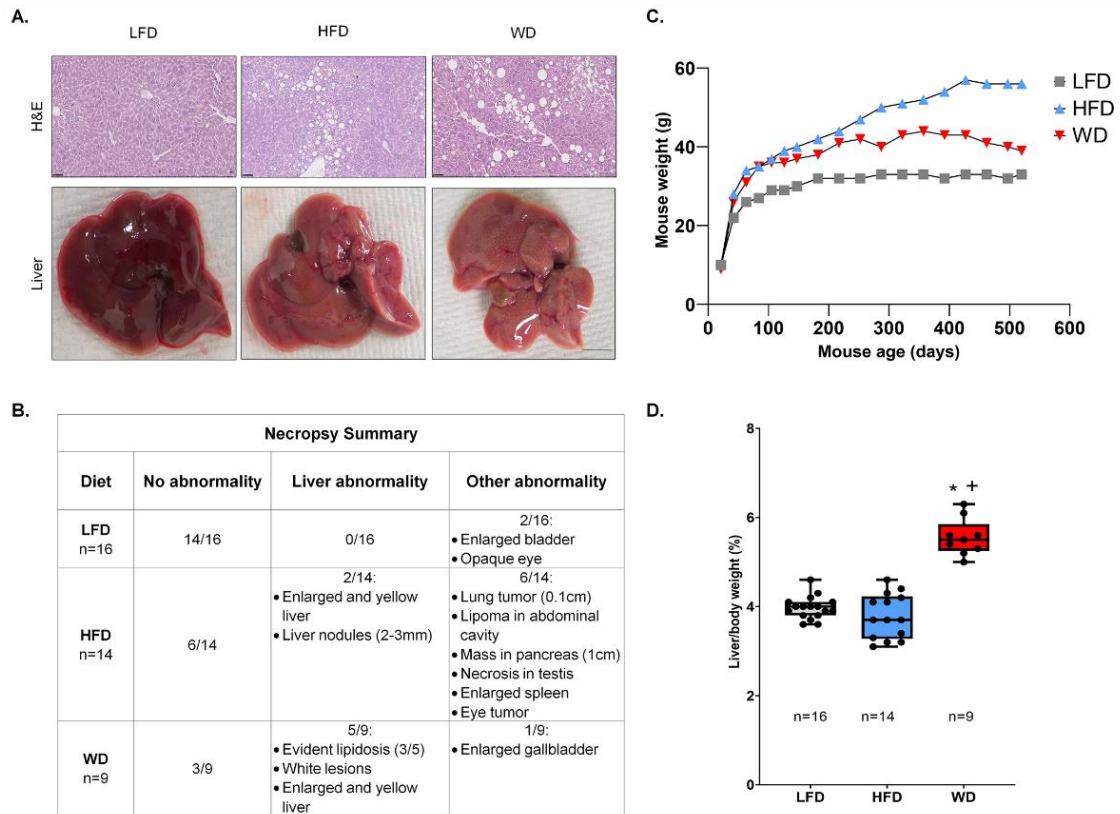
Figure 8. **A Western diet induces severe liver damage**. A, hematoxylin and eosin
(H&E)–stained liver sections (top panel) and representative pictures of liver (bottom
panel) for the respective diets. B, necropsy quantification on the percent of mice that had
liver abnormalities like lipidosis or masses or other abnormalities in tissues other than
the liver (Chi-Square test, normal versus liver abnormal and nonliver abnormal
p≤0.0001). C, mouse weight from weaning at 21 days to 520 days. D, liver weight
normalized to body weight. Scale bars: 100 μm. Bars represent the mean ± WD,
Western diet. SEM, *p value ≤0.

Table 3. University of Pittsburgh Medical Center (UPMC) histological system for components of NAFLD activity score (NAS) and fibrosis staging

Steatosis	Description	Score
	<5%	0
	5%-25%	1
	25%-50%	2
	>50%	3
Lobular Inflammation	Description	Score
number of inflammatory foci per field of view	no foci	0
	<2 foci/200x	1
	2-4 foci/200x	2
	>4 foci/200x	3
Fibrosis	Description	Score
	F0-no fibrosis	None 0
	F1A-mild perisinusoidal fibrosis in zone3 only	Mild 1
	F1B-moderate perisinusoidal fibrosis in zone 3 only	Moderate 2
	F1C-portal or periportal fibrosis only	
	F2-perisinusoidal fibrosis in combination with portal or periportal fibrosis	Moderate 3
	F3-bridging fibrosis	Massive 4
	F4-cirrhosis	

Oil Red O (ORO) quantification revealed that mice fed a WD had approximately twice as much liver steatosis as mice fed a HFD ($p < 0.05$, Figure 9A) and mice fed a HFD developed significantly more steatosis than those receiving the LFD ($p < 0.05$, Figure 9A). Qualitative histologic assessment based on the percentage of steatosis was modified from the UPMC/NAS scoring system where 0=<5%, 1=5%-25%, 2=25%-50%, and 3=>50% confirmed this finding. Over 50% of mice receiving a WD had a score of 3, while mice fed a HFD had a score of 1 and 2 (Figure 9A). Hepatic inflammation was quantified by immunohistochemistry (IHC) to detect a cluster of differentiation (CD) 3, a marker of T-cell activation (223), which was elevated in mice with NASH phenotypes (Figure 9B) (Since inflammation is a phenotype observed primarily in NASH and not in NAFLD, only WD was used for this analysis using the LFD as the control diet). More than 60% of mice fed a WD had more lobular inflammation than those fed a LFD (Figure 9B). Finally, liver fibrosis was quantified by picrosirius red (PSR) staining combined with polarized imaging as well as Masson-Trichrome staining. Mice on a WD developed more fibrosis than those on the HFD (Figure 9C and Figure 10A). A very similar pattern was observed by pathology using the fibrosis stage scoring system by UPMC(222). At least 50% of the mice fed a WD had a fibrosis score of 4 representing bridging fibrosis and some cirrhosis while at least 90% of mice in a HFD had a score of 3 representing portal and/or periportal fibrosis (Figure 9C). In addition, liver fibrosis observed in a WD was associated with an upregulation in the mRNA levels of *Col1a1*, known to be a prominent extracellular matrix protein produced in abundance in hepatic fibrosis (Figure 10B). In aggregate, the data demonstrates that in the mouse models, a HFD appears to induce NAFLD, while a diet high in fat and high sugar, typical of a WD, induces a NASH phenotype.
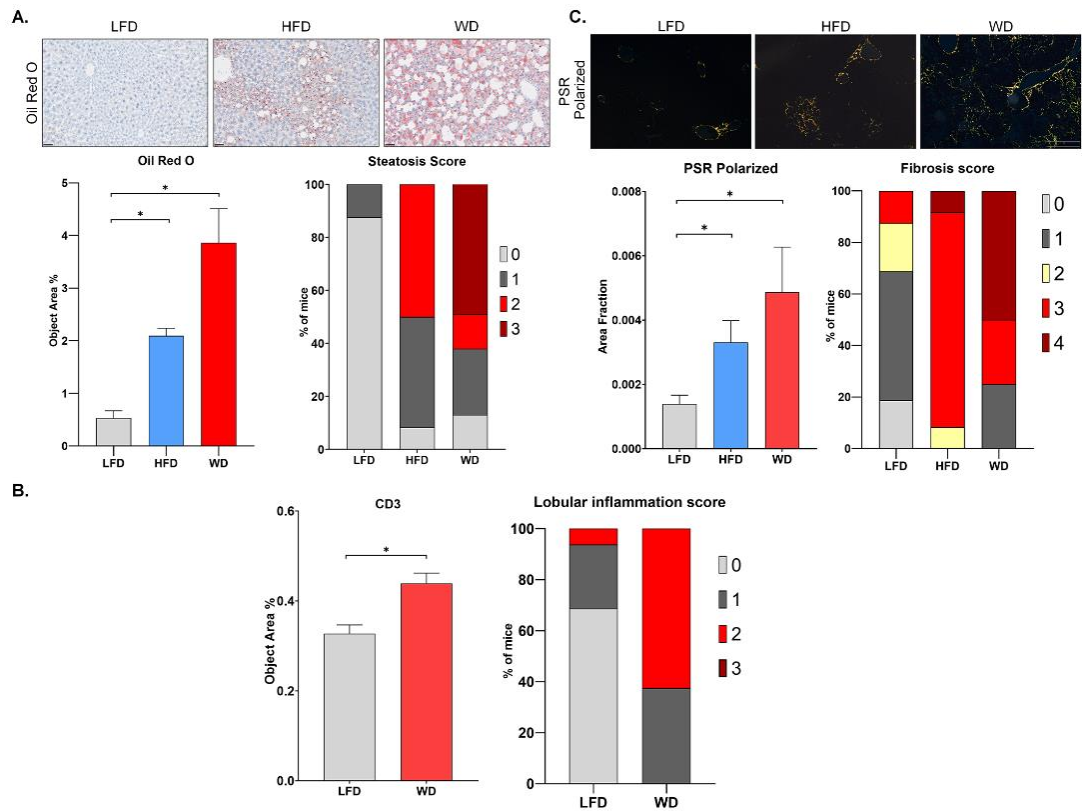
Figure 9. A Western diet induces NASH-like phenotypes. A, lipids were visualized by Oil Red O (ORO) staining in frozen sections. (n = 3 per diet). Percent of mice scored by the level of hepatic steatosis. (LFD n = 16, HFD n = 12, and WD n = 8). B, CD3 immunohistochemistry quantification (n = 3–4 per diet). Percent of mice scored by the level of hepatic lobular inflammation. (LFD n = 16, HFD n = 12, and WD n = 8). C, picrosirius red (PSR) staining in paraffin-embedded tissue using a microscope polarized view. (n = 7 per diet). Percent of mice scored by the level of hepatic fibrosis (LFD n = 16, HFD n = 12 and WD n = 8). Scoring systems are specified in supplemental Fig. S1A. Scale bars: 100 μm. Bars represent the mean ±SEM, *p value ≤0.05 versus control (LFD), +p value ≤0.05 HFD versus WD. Mann-Whitney test. HFD, high-fat diet; LFD, low-fat.
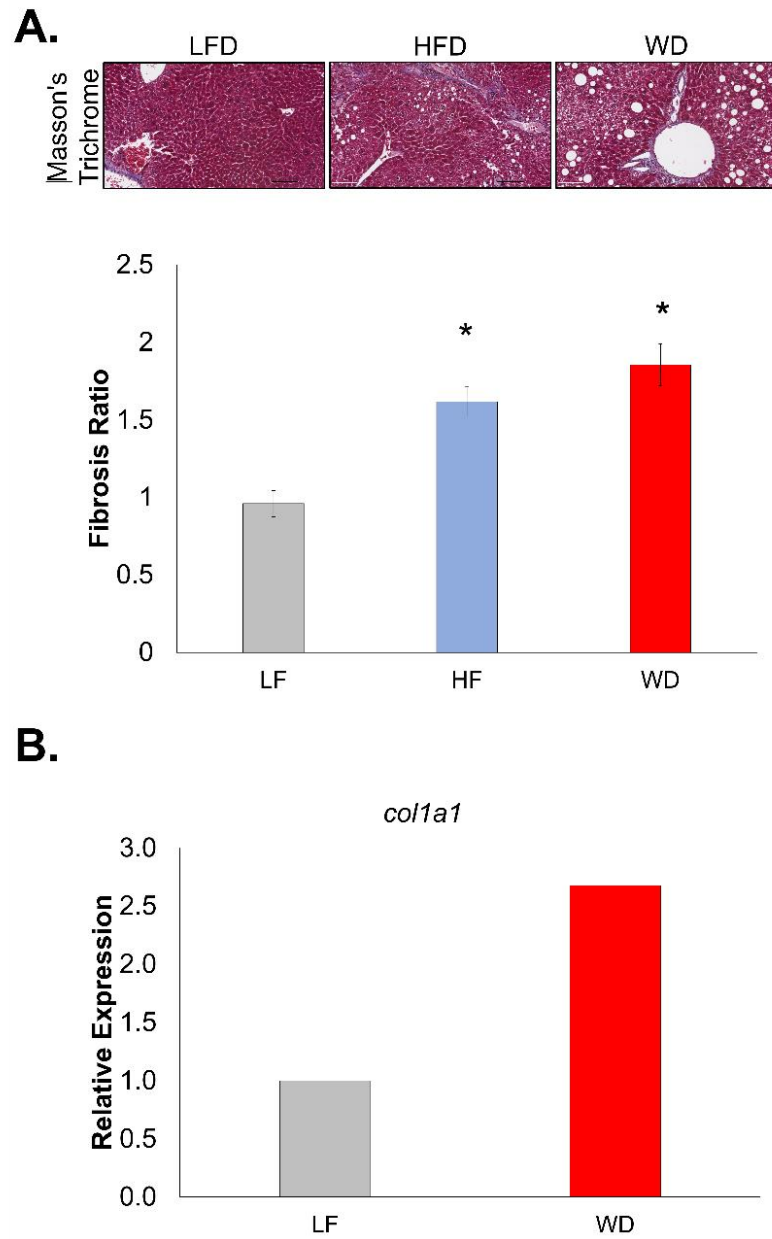
Figure 10. A WD induced a higher level of fibrosis. A, Masson's Trichrome staining quantification in paraffin-embedded tissue. Scale bars: 100µm. n=2-3 per diet. B, mRNA expression by qPCR for cola1 (n=2 mice per diet).

4.3.2 NAFLD and NASH upregulate specific N-glycan structures

To characterize the N-glycan modifications induced by NAFLD/NASH in our mouse model, we used Matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS) technology and N-Glycosidase F (PNGase F PRIME™) enzyme for specific cleavage of N-glycan structures from the glycoprotein at the Asparagine (Asn) residue. Paraffin-embedded liver tissues were processed following standard protocols as in Methods (9,32) (Figure 11A). MALDI-IMS revealed that livers from mice fed the HFD and WD had upregulation in N-linked glycan modifications when compared to livers from mice fed a LFD. A HFD induced glycan modification in high mannose (i.e., 1905.476 m/z) and complex (i.e., 1663.582 and 2174.769 m/z) glycans. Livers from mice receiving a WD had an increased intensity specifically in hybrid (i.e., 1622.557 m/z), complex branched glycan, and in certain fucosylated N-linked glycan (i.e., 2174.769, 1809.755, 2393.954 m/z) compared to livers from LFD mice (Figure 11A, 11B). In addition, representative images demonstrated the specificity of N-glycan structures in each tissue and revealed a higher expression of fucosylated glycans around portal triads of WD liver tissue (Figure 11C). Furthermore, due to the upregulation of fucosylated glycan structures, we investigated the effects of a WD in N-glycan fucosylation by gene expression of Fucosyltransferase 8 *(Fut8)*, the main enzyme capable of adding a core fucose structure and more commonly studied in HCC. *Fut8* increased expression had a correlation with advanced liver disease in mice fed a WD (Figure 11D). Overall, N-glycosylation profiles appeared to be altered as a function of diet. A list of all N-glycans reported in this analysis and representative structures can be found in Table 4.
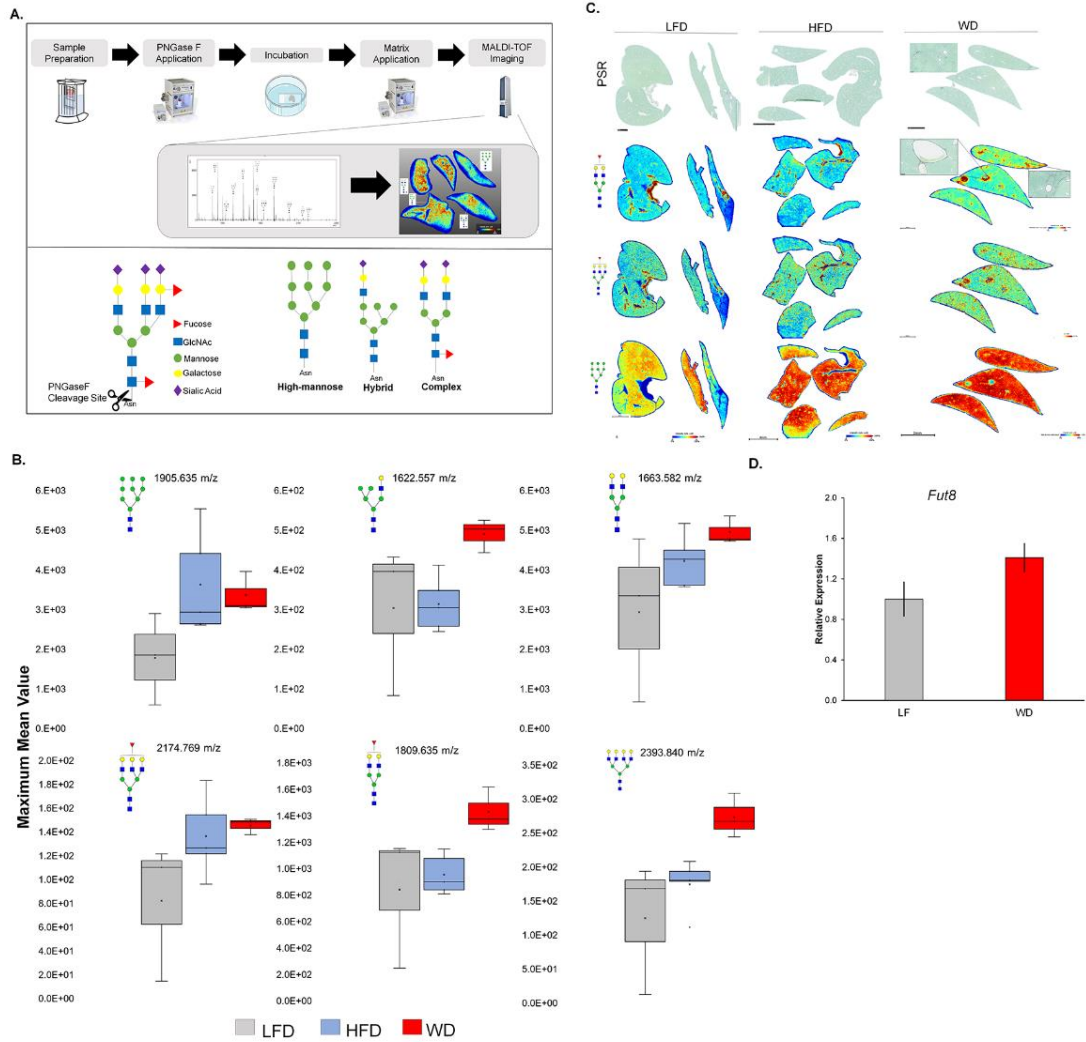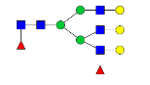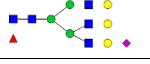
Figure 11. NAFLD and NASH upregulate specific N-glycan structures. A, workflow

for tissue preparation for N-linked glycan analysis across tissue for spatial information by

matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-IMS)

using time of flight (TOF) instruments (top). Labeling system and types of N-glycans

structures. Structural assignments of glycan structures were made based on the

composition of HexNAc, where more than four HexNAc were considered tri-antennary

while more than four HexNAc with limited number of galactoses were considered

bisecting (bottom). B, maximum mean value of m/z with changes between diets

generated by SCiLs analysis software. m/z values correspond to N-glycan structures. (n = 3–5 mice per diet). C, representative images of N-glycans structures upregulated (1809.635 m/z, 2174.771 m/z and 1905.635 m/z) with histopathological correlation of fucosylated glycans in portal triad areas shown by pic- rosirius red (PSR) staining. D, mRNA expression by qPCR for fucosyltransferase 8 (Fut8) normalized to Gapdh. (n = 3 per diet). Bars represent the mean ±SEM. HFD, high-fat diet; LFD, low-fat diet; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; WD, Western diet.

Table 4. Observed mass (m/z values) with the theoretical values, hexose compositions, and proposed glycan structure for the 18-month-old mouse cohort. Observed mass values are representative from low fat diet.

Observed Mass (m/z)	Theoretical Mass (m/z)	Mass Error	Glycan Structure	Proposed Glycan
933.319	933.317	2.678	Hex3HexNAc2+ 1Na	
1079.374	1079.374	-0.278	Hex3dHex1HexNAc2 + 1Na	
1095.370	1095.369	0.730	Hex4HexNAc2 + 1Na	
1136.402	1136.396	5.279	Hex3HexNAc3 + 1Na	
1241.415	1241.427	-10.069	Hex4dHex1HexNAc2+ 1Na	
1257.422	1257.422	-0.477	Hex5HexNAc2 + 1Na	
1282.452	1282.454	-1.637	Hex3dHex1HexNAc3+ 1Na	
1298.448	1298.449	-0.847	Hex4HexNAc3 + 1Na	
1339.474	1339.475	-0.671	Hex3HexNAc4 + 1Na	
1403.474	1403.480	-3.990	Hex5dHex1HexNAc2+ 1Na	
1419.473	1419.475	-1.197	Hex6HexNAc2+ 1Na	
1444.502	1444.507	-2.907	Hex4dHex1HexNAc3 + 1Na	
1460.501	1460.502	-0.273	Hex5HexNAc3+ 1Na	
1485.532	1485.533	-0.673	Hex3dHex1HexNAc4+ 1Na	
1501.526	1501.528	-1.465	Hex4HexNAc4+ 1Na	
1542.555	1542.555	0.259	Hex3HexNAc5 + 1Na	
1581.527	1581.528	-0.316	Hex7HexNac2 + 1Na	
1606.537	1606.559	-13.818	Hex5dHex1HexNAc3+ 1Na	
1622.557	1622.554	1.170	Hex6HexNac3 + 1Na	
1647.586	1647.586	-0.182	Hex4dHex1HexNAc4 + 1Na	
1663.582	1663.581	0.601	Hex5HexNAc4+ 1Na	
1688.611	1688.613	-1.125	Hex3dHex1HexNAc5 + 1Na	
1704.606	1704.607	-0.762	Hex4HexNAc5+ 1Na	
1743.582	1743.581	0.860	Hex8HexNAc2 + 1Na	
1773.578	1773.579	0.563	Hex4dHex1HexNAc3NeuGc1 + 2Na	

1809.635	1809.639	-2.210	Hex5dHex1HexNAc4 + 1Na	
1825.634	1825.634	0.000	Hex6HexNAc4+ 1Na	
1850.677	1850.665	6.322	Hex4dHex1HexNAc5 + 1Na	
1866.674	1866.660	7.232	Hex5HexNAc5+ 1Na	
1905.635	1905.633	0.997	Hex9HexNAc2 + 1Na	
1928.688	1928.645	22.295	Hex6HexNAc3NeuGc1 + 1Na	
1953.647	1953.676	14.843	Hex4dHex1HexNAc4NeuGc1 + 1Na	
1971.676	1971.692	-8.013	Hex6dHex1HexNAc4 + 1Na	
1992.650	1992.653	1.505	Hex5HexNAc4NeuGc1 + 2Na	
1995.656	1995.703	-23.249	Hex4HexNAc5NeuAc1+ 1Na	
2012.705	2012.718	-6.359	Hex5dHex1HexNAc5 + 1Na	
2028.706	2028.713	-3.351	Hex6HexNAc5 + 1Na	
2067.681	2067.686	-2.466	Hex10HexNAc2 + 1Na	
2069.695	2069.740	-21.451	Hex5HexNAc6 + 1Na	
2100.732	2100.734	-1.237	Hex5dHex1HexNAc4NeuAc1+ 1Na	
2101.738	2101.755	-7.945	Hex5dHex3HexNAc4+ 1Na	
2122.706	2122.724	-8.526	Hex5dHex1HexNAc4NeuAc1 + 2Na	
2157.755	2157.756	-0.509	Hex5HexNAc5NeuAc1+ 1Na	
2158.752	2158.776	-11.024	Hex5dHex2HexNAc5+ 1Na	
2174.769	2174.771	-0.965	Hex6dHex1HexNAc5+ 1Na	
2179.723	2179.745	-10.505	Hex5HexNAc5NeuAc1 + 2Na	
2231.769	2231.793	-10.708	Hex6HexNAc6+ 1Na	
2277.767	2277.782	6.585	Hex6dHex1HexNAc4NeuGc1 + 1Na	
2319.809	2319.809	0.301	Hex6HexNAc5NeuAc1+ 1Na	

2320.802	2320.829	-11.461	Hex6dHex2HexNAc5+ 1Na	
2341.768	2341.798	-12.767	Hex6HexNAc5NeuAc1 + 2Na	
2393.840	2393.845	-2.339	Hex7HexNAc6 + 1Na	
2465.858	2465.866	-3.568	Hex6dHex1HexNAc5NeuAc1+ 1Na	
2487.848	2487.856	-3.496	Hex6dHex1HexNAc5NeuAc1 + 2Na	
2539.912	2539.903	3.582	Hex7dHex1HexNAc6+ 1Na	
2905.013	2905.035	-7.814	Hex8dHex1HexNAc7+ 1Na	

4.3.3 N-glycosylation modifications in human NASH

To confirm N-glycan modifications observed in our *in vivo* model based on disease progression, human liver NASH biopsies were analyzed for N-linked glycans. NASH biopsies were scored using the NASH CRN fibrosis scoring (FS) system: FS 0, FS 1, FS 2, FS 3, and FS 4 (Table 5). Patients across fibrosis stages exhibited abnormal aminotransferase levels (consistent with hepatocellular injury typical of NASH), but as expected, these liver test abnormalities did not appear to be associated with the level of fibrosis (Table 6). In-situ glycan analysis was done as previously mentioned in our mouse model.

Table 5. NASH CRN system for fibrosis scoring with description and number of samples in each group. Some patients had more than one biopsy (sample) collected

Group	Fibrosis Score	Description	Number of samples
FS 0	0	No fibrosis	8
FS1	1 (1a)	Mild perisinusoidal fibrosis	8
	1 (1b)	Moderate perisinusoidal fibrosis	
	1 (1c)	Portal/periportal fibrosis	
FS2	2	Perisinusoidal and portal/periportal fibrosis	9
FS3	3	Bridging fibrosis	15
FS4	4	Cirrhosis	11

Table 6. Patient clinical information from NASH biopsies. Some clinical information was not obtained for all patients.

Fibrosis Score	0	1	2	3	4
Number of patients	8	8	10	12	11
Age (mean, SD)	60(6.4)	53(12.6)	58.5(8.0)	63(6.6)	58(19.5)
Gender (M/F)	0/7	5/3	2/8	6/5	4/6
BMI (kg/m²) (mean)	35	32	32	33	32
Biluribin total (mg/dL) (mean)	0.41	0.62	0.54	0.69	0.73
AST (U/L) (mean)	28	38	51	51	32
ALT (U/L) (mean)	37	63	64	48	48
Alk Phos (U/L) (mean)	90	83	81	85	85
INR (mean)	0.97	0.96	1.02	1.04	1.03

Initially, we examined those patients with no fibrosis (FS 0) but with low or high

steatosis (*defined in Methods*) and compared their glycan profile (Figure 12A).

Consistent with what was observed in animals fed a HFD or WD, increased levels of the

overall glycan profile, specifically, high mannose glycan and complex/fucosylated

glycans were observed in the biopsies of patients with high steatosis, but no evidence of

fibrosis (Figure 12A). Indeed, as representative images show, these N-glycans were

altered specifically in areas with the highest levels of steatosis, suggesting an effect

upon N-linked glycosylation.

Next, we examined N-glycosylation changes specific to fibrosis scores for all

patients. Fibrosis was also assessed using histomorphometric analysis of collagen area

(Figure 13A). Upregulation in fucosylated glycan structures correlated with patients with

all scores of fibrosis (Stages 1-4). However, the greatest increase in fucosylation was

observed in patients with more advanced fibrosis (FS3 and FS4) (Figure 12B).

Representative images show the specificity of the high intensity of these fucosylated

glycans to specific regions of the tissue. Fibrosis score correlations were specific to

fucosylated glycans as high mannose glycans, had no significant changes when

analyzed by fibrosis stages (Figure 12C). Representative images show that the regions

within the tissue where there is an increase in the intensity of high-mannose structures

are different from that observed in fucosylated images. To further explore the specificity

of glycan structures to pathological regions within the tissue, we used 2D heatmap

representative images of N-glycan structures and compared them with Pico Sirius Red

(PSR) staining images to demonstrate N-glycan changes occurring at specific

histopathological areas of the tissue, specifically fibrosis. Since some of the major glycan

modifications observed were fucosylation and high mannose structures, we focused on

89

histopathological changes in tissue occurring where these glycan structures were expressed at high intensity. A representative liver specimen with a fibrosis score of 4 was selected that included some steatosis and high levels of fibrosis to demonstrate the expression specificity of glycan structures within the tissue. The higher intensity of fucosylated glycans was primarily expressed in fibrotic areas while high mannose structures were expressed in fatty areas or non-fibrotic areas (Figure 12D). Overall, we elucidate the origin of N-glycan modifications by its specificity with histopathological changes like fibrosis and steatosis. A list of all N-glycan structures and their representative structures is in Table 7. Figure 13B describes the tissue selection method for MALDI-IMS imaging in NASH biopsies. Table 8 details the order organized by groups in which the MALDI-IMS data was acquired.
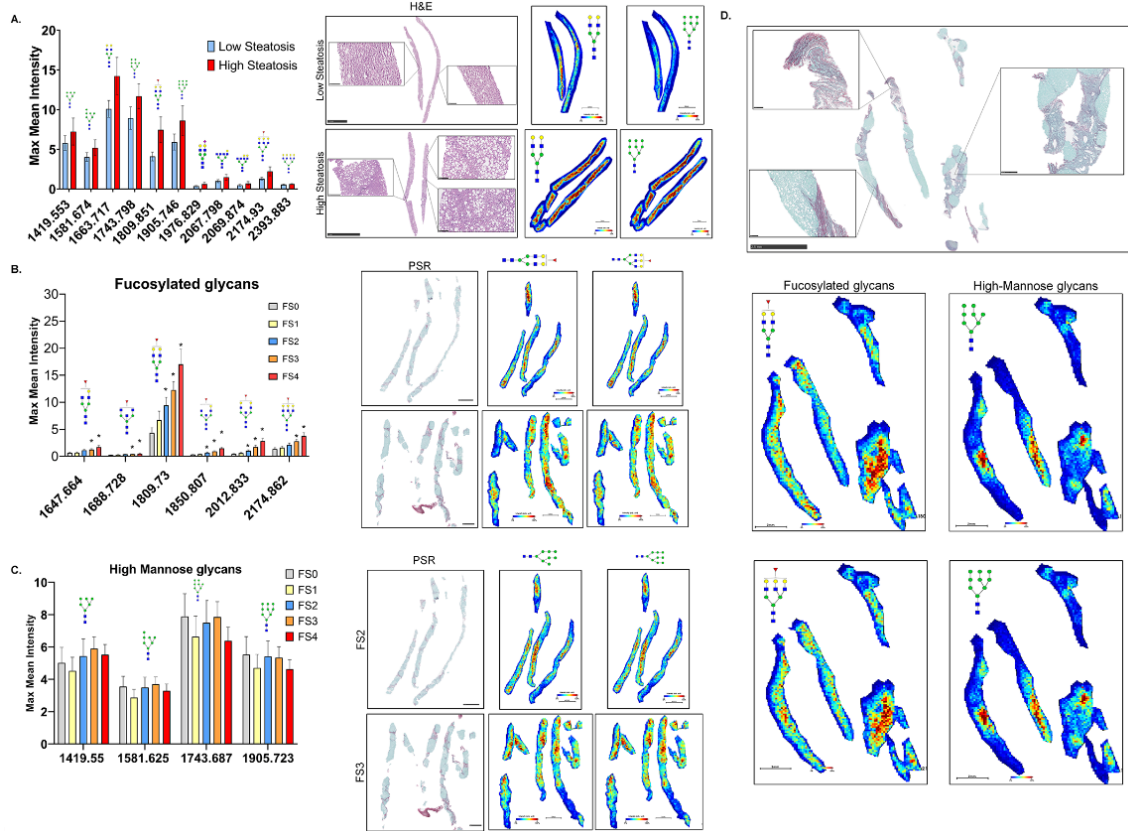
Figure 12. N-glycosylation modifications in vivo are validated in human NASH biopsies and correlate with fibrosis score and histo- pathological changes in tissue. (A) maximum mean value of m/z of patients scored an FS 0 but classified as low or high steatosis (left). Representative H&E and 2D heatmap images of 1663.575 m/z and 1905.630 m/z (right). Bars represent the mean ±SEM. (B) maximum mean value of m/z of fucosylated (red triangle) glycans with significant changes between fibrosis scores generated by SCiLs analysis software (left). 0 versus FS 1, FS 2, FS 3 or FS 4. Mann-Whitney test. (C) maximum mean value of m/z of high mannose glycans with changes between FS Representative PSR and 2D heatmap images of 1809.638 m/z and 2174.775 m/z (right). Bars represent the mean ± SEM, *p value ≤ 0.05 FS generated by

SCiLs analysis software (left). Representative PSR and 2D heatmap images of 1743.687 m/z and 1905.630 m/z (right). Bars represent the mean ±SEM. D, PSR staining from patient with a fibrosis score of 4 with representative images of 2D heatmap showing changes in glycan intensity based on histopathological events in the tissue. Zoom in pictures represent fibrotic (red staining) and/or fatty areas of tissue that correlate with fucosylated (left) or high mannose structures (right) increased intensity. FS, fibrosis score; H&E, hematoxylin and eosin; PSR, picrosirius red.
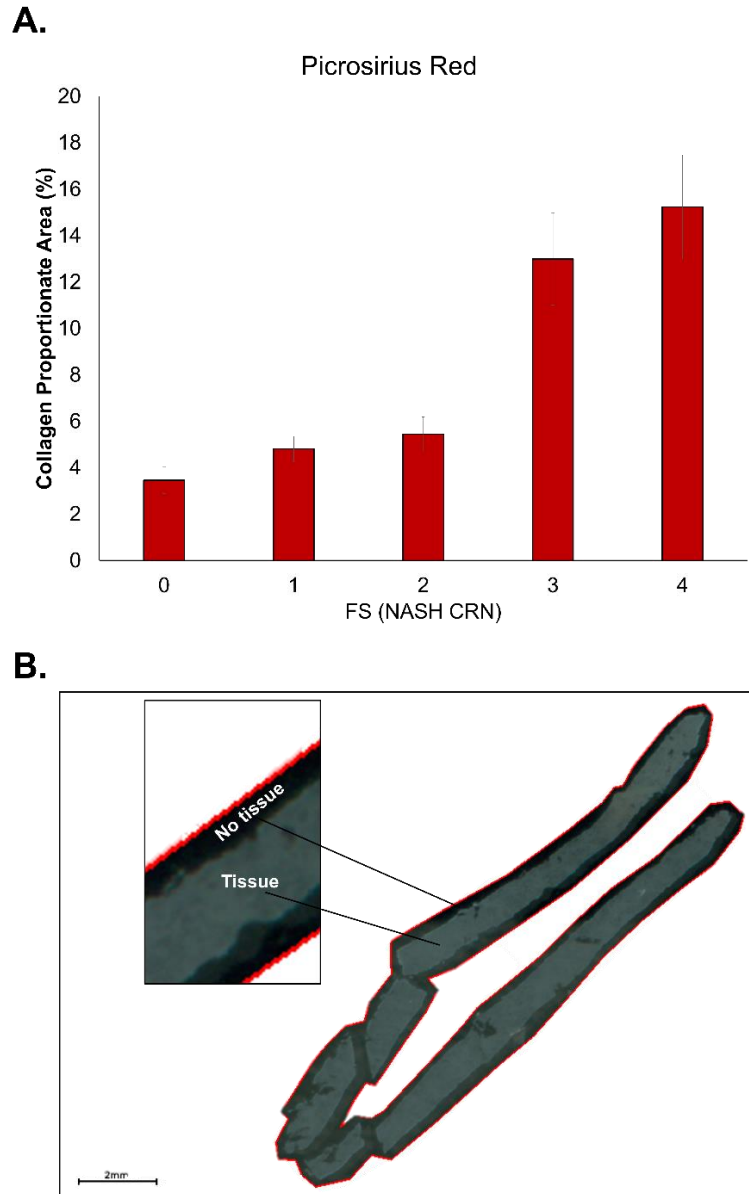
A.

Picrosirius Red

B.

Figure 13. (A) Collagen Proportionate area percent quantification of Picrosirius Red

(PSR) staining by fibrosis score. (B) NASH biopsy with no 2D heatmap intensity (from

Figure 2A) of any glycan showing how the tissue is selected for imaging. Zoom in image

shows as space between the tissue and where the tissue was selected for imaging

Table 7-Observed mass (m/z values) with the theoretical values, hexose compositions, and proposed glycan structure from analysis of NASH patient biopsies.

Observed Mass (m/z)	Theoretical Mass (m/z)	Mass Error	Glycan Structure	Proposed Glycan
1079.373	1079.374	1.759	Hex3dHex1HexNAc2 + 1Na	
1095.378	1095.369	7.487	Hex4HexNAc2 + 1Na	
1136.403	1136.396	5.808	Hex3HexNAc3 + 1Na	
1257.417	1257.422	4.293	Hex5HexNAc2 + 1Na	
1298.452	1298.449	2.773	Hex4HexNAc3 + 1Na	
1339.466	1339.475	6.643	Hex3HexNAc4 + 1Na	
1419.466	1419.475	6.622	Hex6HexNAc2 + 1Na	
1444.508	1444.507	0.900	Hex4dHex1HexNAc3 + 1Na	
1460.506	1460.502	2.738	Hex5HexNAc3 + 1Na	
1542.553	1542.555	0.906	Hex3HexNAc5 + 1Na	
1581.528	1581.528	0.063	Hex7HexNac2 + 1Na	
1622.565	1622.554	2.896	Hex6HexNac3 + 1Na	
1647.589	1647.586	1.760	Hex4dHex1HexNAc4 + 1Na	
1663.575	1663.581	3.847	Hex5HexNAc4+ 1Na	
1688.621	1688.613	4.915	Hex3dHex1HexNAc5 + 1Na	
1704.603	1704.608	2.932	Hex4HexNAc5+ 1Na	
1743.591	1743.581	5.735	Hex8HexNAc2 + 1Na	
1809.638	1809.639	0.331	Hex5dHex1HexNAc4 + 1Na	
1825.632	1825.634	1.149	Hex6HexNAc4+ 1Na	
1850.667	1850.665	0.811	Hex4dHex1HexNAc5 + 1Na	
1866.652	1866.660	4.392	Hex5HexNAc5+ 1Na	
1891.692	1891.692	0.000	Hex3dHex1HexNAc6+ 1Na	
1905.630	1905.633	1.626	Hex9HexNAc2 + 1Na	
1955.694	1955.697	1.380	Hex5dHex2HexNAc4+ 1Na	
1971.692	1971.692	0.000	Hex6dHex1HexNAc4 + 1Na	
1976.662	1976.666	2.276	Hex5HexNAc4NeuAc1 + 2Na	

1996.716	1996.723	3.856	Hex4dHex2HexNAc5+ 1Na	
2012.713	2012.718	2.583	Hex5dHex1HexNAc5 + 1Na	
2028.713	2028.713	0.295	Hex6HexNAc5 + 1Na	
2067.683	2067.686	1.305	Hex10HexNAc2 + 1Na	
2069.753	2069.740	6.281	Hex5HexNAc6 + 1Na	
2122.714	2122.724	4.522	Hex5dHex1HexNAc4NeuAc1 + 2Na	
2158.776	2158.776	0.277	Hex5dHex2HexNAc5+ 1Na	
2174.775	2174.771	1.655	Hex6dHex1HexNAc5+ 1Na	
2215.790	2215.798	3.384	Hex5dHex1HexNAc6+ 1Na	
2231.793	2231.793	0.000	Hex6HexNAc6+ 1Na	
2276.718	2276.718	0.000	Hex6dHex1HexNAc5 + 1SO4 + 2Na	
2320.824	2320.829	2.240	Hex6dHex2HexNAc5+ 1Na	
2341.783	2341.798	6.477	Hex6HexNAc5NeuAc1 + 2Na	
2377.838	2377.850	5.046	Hex6dHex1HexNAc6+ 1Na	
2393.850	2393.845	1.754	Hex7HexNAc6 + 1Na	
2487.845	2487.856	4.542	Hex6dHex1HexNAc5NeuAc1 + 2Na	
2539.901	2539.903	0.708	Hex7dHex1HexNAc6+ 1Na	
2905.003	2905.035	2.788	Hex8dHex1HexNAc7+ 1Na	

Table 8. Order in which MALDI-IMS data was acquired, samples were prepared and ran for MALDI-IMS experiments randomly. Samples are organized by groups in this table for a better description of the order (i.e., samples in Group 1 were prepared and ran first, followed by samples in group 2, etc.). ID: identification number given to each sample based on the order they were registered in our laboratory database.

Fibrosis Score	ID	Group	Fibrosis Score	ID	Group
0	20	1	3	12	2
0	21	1	3	14	2
0	22	1	3	15	3
0	23	1	3	26	3
0	24	3	3	27	3
0	40	3	3	28	2
0	41	3	3	29	2
0	42	3	3	30	2
1	18	2	3	31	1
1	19	1	3	44	3
1	36	3	3	52	2
1	37	3	3	53	2
1	38	3	3	54	2
1	39	1	3	55	2
1	47	1	3	58	1
1	48	1	4	1	1
2	8	2	4	2	1
2	16	3	4	3	1
2	17	2	4	4	2
2	35	2	4	5	2
2	34	2	4	12	2
2	45	1	4	25	2
2	46	2	4	60	2
2	50	2	4	61	2
2	51	2	4	62	2
			4	63	2

4.3.4 Core fucosylated N-glycan modifications are specific to fibrotic areas in human NASH

Alterations in fucosylation described here, are known to occur at different linkages of the glycan structure and are catalyzed by a family of Fucosyltransferases (FUTs 1-11). The specific fucose linkages are of interest as fucose residues attached in an alpha-1,6 linkage, or core fucosylation catalyzed by FUT8 is the most reported modification in severe liver damage (82,147). Since the MALDI-IMS method utilized in this study is unable to identify the specific fucose linkages of the particular glycans with only the use of the PNGase F PRIME™ enzyme. We implemented the use of the enzyme Endo F3 Prime™ to further identify the nature of the fucosylation observed in Figure 11. Endo F3 Prime™ has a preference to cleave core-fucosylated glycans (Figure 14A) (128). The cleavage of Endo F3 Prime™ between the two core N-acetylglucosamine residues results in a mass shift of 349.137 a.m.u for core fucosylated glycans compared to N-glycans released by only PNGase F PRIME™. Treatment with the enzyme Endo F3 Prime™ revealed eight main glycans found to be core-fucosylated with altered intensity based on the fibrosis scores (Figure 14B and Table 6). Representative images of the intensity of core fucosylated glycans in tissue show that these structures are expressed throughout the tissue at early fibrotic stages (FS0 and FS2) but as the disease progresses and fibrosis increases core fucosylated glycans are expressed only in fibrotic areas and at a very low intensity or not present at all in non-fibrotic areas (Figure 14C). However, core fucosylated hybrid (1581.528 m/z and 1460.500 m/z) and core fucosylated bisecting (1745.500 m/z) glycans were not present in fibrotic areas, instead these were highly present in non-fibrotic areas and/or steatotic areas (Figure 15). To further explore and confirm the expression of non-core fucosylated

glycans we added treatment with PNGase F PRIME™ after Endo F3 Prime™ treatment. We analyzed high mannose glycan intensity and show that there was a significant decrease in these structures with a high fibrotic score, suggesting these structures to be more abundant in steatotic tissues instead of fibrotic tissue (Figure 14E). Treatment with both enzymes revealed that 1809.500 m/z and 2539.901 m/z are mainly core fucosylated since these glycans had a lower intensity and a different expression pattern in the tissue compared to the core fucosylated form, suggesting having some terminal fucosylation. The 1850.665 m/z glycan had a similar intensity at similar areas as the core fucosylated form, suggesting it also has a terminal fucosylated form. As mentioned before, high mannose glycans were consistently expressed in non-fibrotic areas (1419.472 m/z and 1905.637 m/z). Finally, 1663.584 m/z biantennary glycan was expressed throughout the tissue, which is expected as this glycan is commonly found in different types of tissue at high levels (Figure 14D). Overall, we found a significant positive correlation between core fucosylated glycans: (1501.509 m/z, r=0.7656, p=0.0023 and 1663.586 m/z, r=0.6055, p=0.0248) and a negative correlation in high mannose glycans: (1905.632 m/z, r=-0.7852, p= 0.0049 and 1743.585 m/z, r=0.5349, p=0.0790) based on the fibrosis scores (Figure 14F). Our results suggest that core fucosylated glycans could have a role in disease progression, specifically in fibrosis.
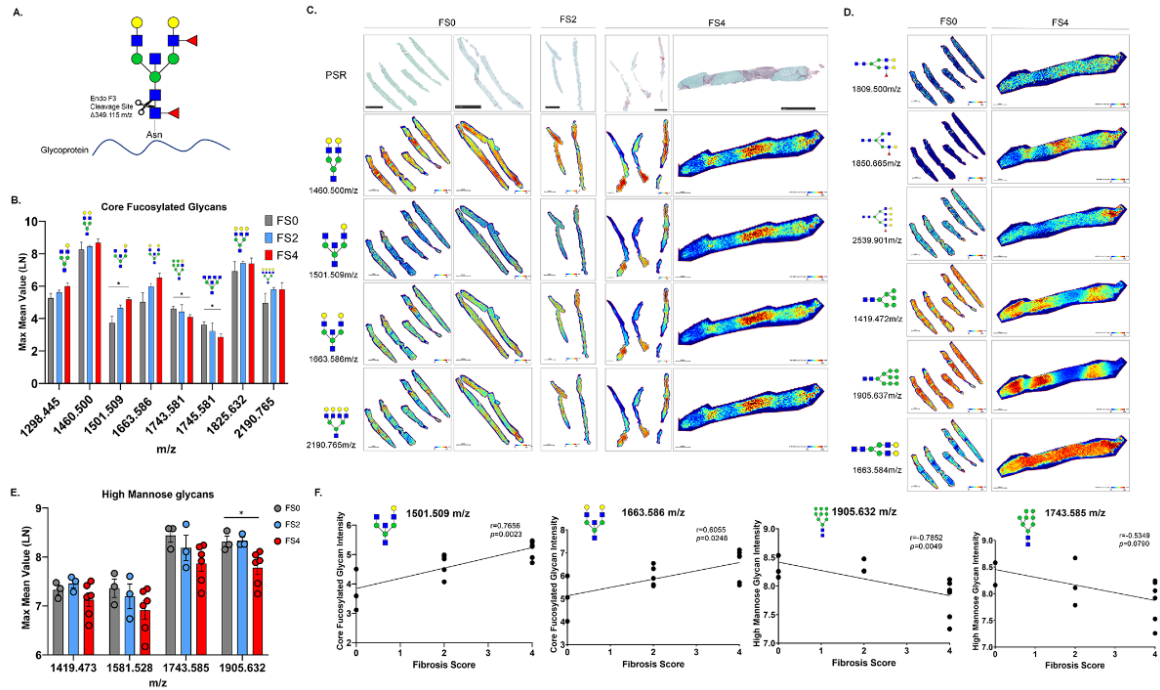
Figure 14. Core fucosylated glycans are specific to fibrotic areas in human NASH biopsies. (A) cartoon description of Endo F3 Prime cleavage site on core fucosylated (red triangle) glycans. (B) maximum mean value normalized to the natural logarithm (LN) of m/z values of core fucosylated glycans with significant changes between fibrosis scores 0 (n = 3), 2 (n = 5), and 4 (n = 6) generated by SCiLs analysis software. A core fucosylated glycan will have a mass shift with Endo F3 Prime treatment, (PNGase F PRIME mass shifts): 1298.445 m/z (1647.589 m/z), 1460.500 m/z (1809.638 m/z), 1501.509 m/z (1850.667 m/z), 1663.586 m/z (2012.713 m/z), 1743.581 m/z (2092.718 m/z), 1745.581 m/z FS 2 or FS 4. Mann-Whitney test. (C) representative Picrosirius red (PSR) staining and 2D heatmap images of the expression pattern of core (2094.771 m/z), 1825.632 m/z (2174.775 m/z), and 2190.765 m/z (2539.901 m/z). Bars represent the mean ± SEM. *p value ≤ 0.05 FS 0 versus fucosylated glycan structures in NASH

99

biopsies with a score of 0, 2, and 4 where red staining is representative of fibrotic areas. (D) representative 2D heatmap images of the expression pattern of terminal fucosylated, high mannose, and complex glycans with PNGase F PRIME treatment after Endo F3 Prime treatment. (E) maximum mean value normalized to LN of m/z values of high mannose glycans with significant changes between fibrosis scores 0, 2, and 4 generated by SCiLs analysis software with PNGase F PRIME treatment after Endo F3 Prime treatment. (F) Spearman correlation plots between glycan structures and fibrosis scores, where each data point represents a patient. Core fucosylated glycans: 1501.509 m/z (1850.667 m/z) and 1663.586 m/z (2012.713 m/z) and high mannose glycans 1905.630 m/z and 1743.591 m/z. FS, fibrosis score; NASH, nonalcoholic steatohepatitis.
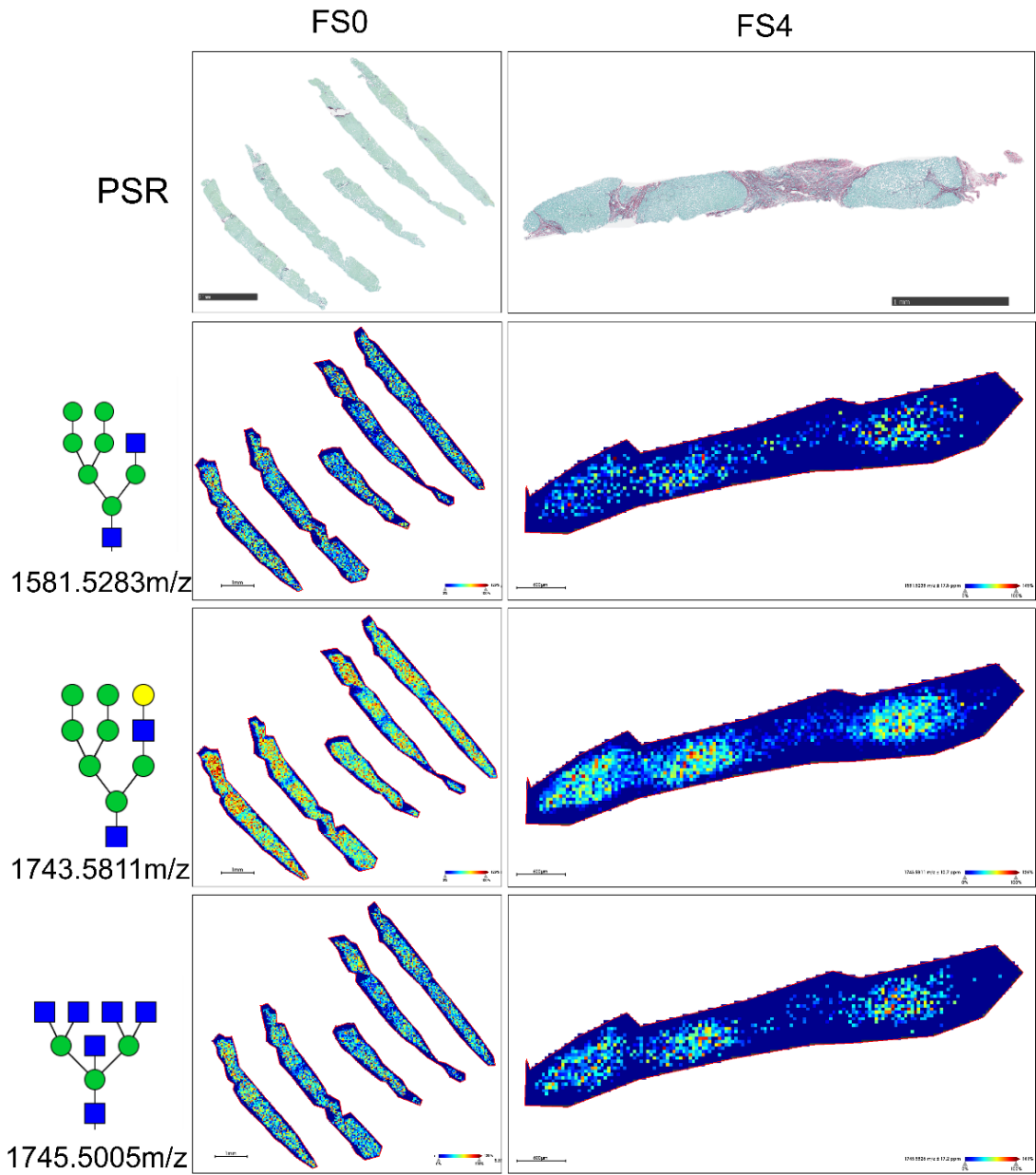
Figure 15. **2D heatmap images of core fucosylated glycans with EndoF3 treatment in NASH biopsies**. PNGaseF mass shift: 1930.6656m/z, 2092.7184m/z, and 2094.7718m/z.

4.4 Discussion

Alterations in glycosylation have been reported in metabolic diseases like fatty liver disease, diabetes, and NASH(89,96,204,224,225). The use of serum glycan markers is the most common method to identify specific N-linked glycan alterations that can be utilized for biomarkers strategies in fatty liver disease and NASH(90,97). However, the elucidation of liver-specific modifications and the origin in tissue of these glycan modifications is not well studied. In this study, we characterized the N-glycan profiles in NAFLD/NASH mouse models and human NASH biopsies and show that alterations of complex/fucosylated and high mannose structures are the main modifications observed in liver tissue. The correlation of histopathological liver damage with specific N-glycan structures that have been linked to liver cancer formation and metastasis suggests the possibility that these N-glycan structures could be useful to detect patients that might progress to cirrhosis or HCC.

While it is known that glycosyltransferases and glycan-processing enzymes are required for many biological functions in the liver like development and liver regeneration, alterations in metabolic diseases and cancers also suggest an oncogenic role for tumor formation and metastasis (4,42,48). For example, gene alterations of fucosyltransferases have also been reported to be upregulated in HCC (68,69,158,226). Similarly, altered glycosylation profiles in branching, sialyation, and fucosylation are known to be glycan signatures in liver, colon, and pancreatic cancer (58,129,227,228). Previous data from our group demonstrated that increased branching and fucosylation levels, correlated with a reduction in survival time in HCC patients (76). Here, we study early liver damage in this case NAFLD/NASH, and similarly, find a significant alteration of complex/fucosylated glycans and high mannose glycans that correlates to fibrotic and

fatty areas within the tissue. It is important to note that the level of fibrosis observed between our study models was significantly lower in our mouse model. It is well-known that fibrosis induction in diet-based mouse models that can mimic human fibrosis has been a limitation without the use of chemical-based agents (164).

Fucosyltransferases (FUTs) 1-11 are the main glycosyltransferases able to catalyze the addition of a fucose residue to different linkages within the glycoprotein(84,208). For example, the isoform L3 of serum glycoprotein alpha-Fetoprotein glycoform (AFP) has core fucosylation alterations in its glycosylation site and is currently used as an HCC biomarker (85,229,230). In our study, we observed significant changes in fucosylation from early stages of the disease, suggesting that core fucosylation could be a driver of liver damage in NASH. Further studies are needed to determine if these were catalyzed by FUT8 and to determine the role of other N-glycan transporters in NASH.

Our data suggest that advanced liver disease correlates with high expression of fucosylated and/or core fucosylated glycan structures that may be detectable in the serum and act as biomarkers of disease progression. This is important because even though NAFLD/NASH are considered early liver damage, diagnostic options are very limited (10,11). Patients can suffer from undetected/undiagnosed NASH and be diagnosed at very late stages of the disease with cirrhosis or HCC (14). Our results demonstrate that alterations in fucosylation can be explored for promising strategies for early detection.

One unique advantage of this study is that all glycan imaging analysis was performed using a MALDI-TOF system which can be clinically accessible and is currently being used in some centers. This instrument allowed us to observe and

103

quantify specific glycosylation changes that correlated nicely with spatial histopathological changes like fibrosis, considered to be one of the most important features of disease progression associated with increase mortality and liver-related complications (231) .

To conclude, we were able to verify that the mouse model used here matched the human disease, validating this model for other NASH studies investigating glycosylation. This study coupled conventional laboratory techniques with mass spectrometry glycan imaging techniques to identify a typical N-glycan liver tissue profile in mouse and human NAFLD and NASH. In vivo studies and human biopsies with low level of fibrosis demonstrated that N-glycan changes can be detected at early stages of disease before liver damage is visible in a pathological assessment. Overall, this study elucidates the N-linked glycosylation alterations in-situ and suggests that disease progression, specifically fibrosis could be the main driver of core fucosylated glycans in NAFLD/NASH, but further studies would be needed to validate the role of N-glycosylation on extracellular proteins in fibrotic tissue.

Chapter 5: N-linked glycosylation alterations identify NAFL before histological annotations.

5.1 Introduction

Liver disease can be caused by a variety of factors including obesity, type 2 diabetes, metabolic disorders, viruses, and genetic variants (9,149). One of the first signs of disease in the liver is Non-alcoholic fatty liver (NAFL), characterized by a steatotic liver, which is known to be one of the most common types of liver damage affecting a high percentage of the population (8). While NAFL is considered a benign liver injury with minimal symptoms, if left untreated, it can progress to non-alcoholic steatohepatitis (NASH) when there is severe inflammation and fibrosis, in addition to steatosis (149). The full scope of NAFL and NASH, both induced independently from alcohol and/or drugs is known as Non-alcoholic fatty liver disease (NAFLD). With liver disease progression, liver injury is considered to be reversible by improving lifestyle habits related to diet and physical activity (232). However, NAFLD is considered a silent disease where patients are asymptomatic during its progression, causing diagnosis to be made often at the late and irreversible stages of the disease. After NAFLD, liver disease can progress to cirrhosis, Hepatocellular Carcinoma (HCC), and Cholangiocarcinoma (CCA) (11,16). Advanced liver disease diagnoses are on the rise due to the overwhelming high rates of obesity and other metabolic diseases (2). In addition, NASH-induced HCC is estimated to become the main cause of liver transplants in the US and it has become a public health concern (10,203). Currently, there are serum-based noninvasive biomarkers to identify NAFLD, HCC, and CCA, however, these lack specificity and sensitivity (27,44,203,233). The gold standard for the diagnosis of these liver diseases varies between imaging strategies, histological and pathological analysis, liver biopsy using a scoring system, and/or a combination (13,26,44,52). However, these have presented some disadvantages, making their use for a definitive diagnosis in the

clinic challenging. Overall, there is an urgent need to find reliable biomarker options that can identify specific stages of liver disease or at least signs of progressive liver disease.

Due to the heterogeneity of all liver diseases previously mentioned (NAFLD, HCC, and CCA) and of the liver organ itself, multiple genetic changes, epigenetic changes, and microenvironmental factors have been categorized as risk factors for liver diseases and their progression (4,234). Glycobiology is a field that has been considered for providing molecular information, specifically in the alterations of N-glycan structures and how these can correlate with diseases (66,72,153,208). N-linked glycosylation is one of the most common post-translational modifications on proteins, where specific sugar residues are added to the glycoprotein, this N-glycan structure attachment occurs at the Asn reside of the glycoproteins (70). While N-linked glycosylation has many biological roles, like protein folding, cell signaling, and immune response, it is also used in the clinic as a biomarker for liver cancers: AFP-L3 for HCC and CA19-19 for CCA (57,153,229). N-linked glycan alterations have been widely studied in different liver diseases, including cirrhosis, HCC, fatty liver, and NASH (73,77,154,204). Different modifications have been identified such as fucosylation, sialylation, and, extensive branching. However, these studies have focused on characterizing the N-glycan-related changes at a specific stage of liver disease progression (78,90,97). It is important to understand these alterations with time in the context of liver disease because it is known that the glycome can be altered by environmental factors affecting the role of N-glycan-related enzymes and N-glycan structures according to disease over time (71,73). In addition, knowing if these glycan modifications come before or after pathological alterations will help identify the roles they play in disease progression.

In this study, we followed the progression of NAFL, NASH, and advanced liver disease in a mouse model to characterize the N-glycan changes occurring over time. Here, we provide a full profile of N-linked glycan structures throughout the progression of liver diseases induced by environmental factors, elucidate the origin of these modifications in situ and identify modifications with promising use for liver disease screening.

5.2 Methods

5.2.1 Diet and mouse models

All animals received humane care according to the criteria outlined in the Guide for the Care and Use of Laboratory Animals. Animal studies were approved by the university's Institutional Animal Care and Use Committee (IACUC). C57BL/6J male mice were purchased from Jackson laboratory at weaning age. Mice were given about 10 days after arrival at our facility for acclimation before any studies started. All studies described here, started when mice were at 4 weeks of age and were separated randomly into respective groups and time points. The diet study included two groups a normal chow (Research Diets), and a western diet (40 kcal% Fat, 20 kcal% Fructose and 2% Cholesterol) (Research Diets D09100310). A total of 48 mice were used: day 0: 3 mice/group, month 2: 3 mice/group, month 4: 6 mice/group, month 6: 6 mice/group, and month 9: 6 mice/group. For the diet/carcinogen study, a total of 31 mice were used: day 0: 3 mice/NC DEN and WD DEN (the mice in the NC and WD were the same used in the diet study) month 4: 3 mice/group, and month 11.5: 3 mice for NC, NC/DEN, WD; and 4 mice for WD/DEN. Mice were fed ad lithium. Histological and scoring of NALFD characteristics was done by an independent veterinary pathologist using H&E staining of the liver tissue section, following an established scoring system (235). Veterinary

pathologists scored mice independently and provided at least 5 scores (from 5 different fields) per sample, scores were averaged for mice in the same group.

5.2.2 Tissue processing for MALDI-IMS

Liver tissue was harvested at respective timepoint processed for formalin-fixation, flash frozen, and serum was collected. Tissue blocks were cut into 5um sections and processed for N-glycan MALDI-IMS following previously published protocols (124). In summary, paraffin-embedded tissues were first heated at 60˚C for an hour, followed by dewaxing steps including xylene, ethanol, and water washes. Next, tissues were taken thru antigen retrieval for 20 minutes at 95˚C in citraconic buffer pH 3.0. Tissues were cooled down by buffer exchange to water slowly to avoid any tissue loss. To cleave N-glycans, PNGaseF prime was sprayed using an HTX sprayer and tissues were left incubating at 37 ˚C for two hours. Finally, tissues were sprayed with CHCA matrix. Samples were imaged using A timsTOF Flex trapped MALDI-QTOF mass spectrometer (Bruker Daltonics) (m/z 500–4000) operating in a positive mode 6.2.4 Data processing and analysis.

5.2.3 Data analysis

Data generated from MADI-IMS experiments were processed using analysis software SCiLS Lab 2020 imaging software (Bruker). SCiLs-generated N-glycan spectras were normalized to the total ion count (TIC). N-glycan structure annotations for each m/z (mass to charge ratio) value were made based on an in-house database of known N-glycan generated using GlycoWorkBench and GlycoMod for annotation or to MS/MS data previously done by our group (217,236). Maximum mean values for each m/z value corresponding to a specific N-glycan structure were normalized to the total

amount of m/z values per sample. Overall, the relative intensity of each m/z value was used for N-glycan quantifications.

5.2.4 IHC Multiplex

Liver tissues were processed after necropsy for formalin fixation. Formalin-fixed embedded tissues were cut at 5um sections and processed for multiplex immunofluorescence using the Roche method for various markers within the same tissues. Markers used here included: anti-aSMA, anti-CD3, anti-F480, and anti-CD8a; using DAPI as the nuclear counterstaining. Stained slides were imaged using the Vectra Polaris Automated Quantitative Pathology Imaging System (PerkinElmer, Inc.).

5.3 Results

5.3.1 Diet and chemically induced liver disease progressive mouse model

To study the full progression of liver disease, liver injury was induced by diet and/or carcinogen treatment (173,176,178,235,237). Mice were harvested at different time points hypothesized to represent different stages of liver disease: from early liver injury (Month 2) to advanced liver disease Month 9 and cancer-like characteristics (Month 11.5). Liver disease progression was studied in two different studies: 1) a diet-inducible NAFL/NASH study, where groups in this study included a normal chow (NC) and a western diet (WD), where timepoints were established based on the stages of disease including Day 0, Month 2, Month 4, Month 6, and Month 9 (Figure 16A). 2) a chemically inducible advanced liver disease study in the background of NALFD, which was a combination of diet and a carcinogen dose. The carcinogen used was Diethylnitrosamine (DEN), commonly used in liver-related studies to induce liver tumors (188). Thus, this study included the combination of diet and DEN: NC, NC/DEN, WD,

and WD/DEN. Timepoints for this study included Day 0, Month 4, and Month 11.5 (Figure 16B). Mouse weights were collected every week for both studies (during nine-month and eleven-and-a-half-month time points) and liver tissue, and serum was harvested at the respective time point. Figure 16C shows the weight changes in the diet study, and as expected mice in the WD gain significantly more weight throughout the study compared to mice fed an NC. Liver pictures harvested at necropsy demonstrate that liver injury in mice fed a WD can be observed at a time point of two months, similarly, liver pictures at Months 4, 6, and 9 demonstrate continuous liver damage including, liver enlargement, discoloration, and rounded edges (Figure 16D). In the diet/carcinogen study, mice in the WD and WD/DEN had a very consistent weight gain until approximately 32 weeks whereas mice in the WD/DEN had a subtle weight decrease, and a more significant weight drop at 40 weeks compared to the WD (Figure 16E). While it is known that mice treated with DEN can lose weight, we hypothesize that this weight loss might be due to the severe liver damage these mice were enduring during this time point, which has been observed by others using this carcinogen (238). Necropsy liver pictures for the diet/carcinogen study had a similar pattern of liver damage to what was observed in the diet study, in addition to a significant number of nodules and stiff liver in the WD/DEN at time point 11.5. Liver from mice fed an NC in combination with the carcinogen also had a minimal liver injury relative to those fed the WD/DEN, this suggests that a WD is a strong inducer of liver disease (Figure 16F).
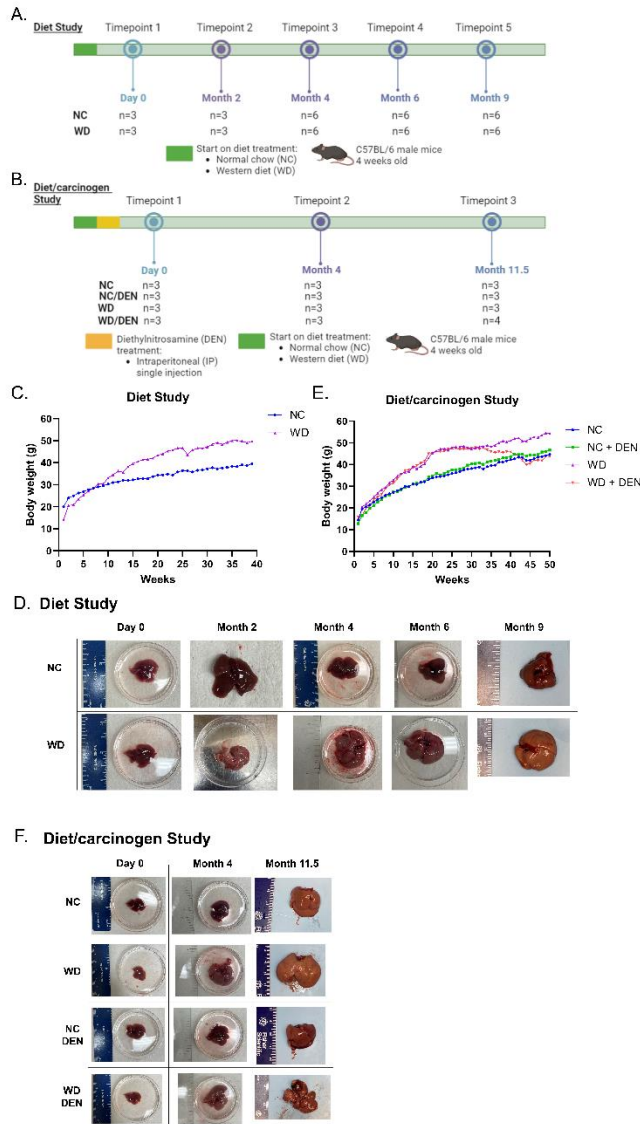
Figure 16. Mouse livers in a western diet developed severe progressive liver damage. (A) Diet-inducible NAFL/NASH study overview (B) Chemically inducible advanced liver disease study in the background of NALFD overview, (C) Mouse weights in diet study throughout 9 months, (D) Necropsy pictures taken at harvesting for each respective timepoint (diet study), (E) Mouse weights in diet/carcinogen study throughout 11.5 months, (F) Necropsy pictures taken at harvesting for each respective timepoint (diet/carcinogen study).

Next, to confirm that liver disease was successfully induced, and to confirm the level of diseases, veterinary pathologists performed histological scoring utilizing the same predictor factors used clinically for NAFLD diagnosis (macrosteatosis, microsteasotis, inflammation, fibrosis). In addition, hyperplasia was also scored, since this is an abnormal change in normal cells prone to becoming cancer cells. Table 9 lists the specific description used for histological analysis. For the diet study, as Figure 17 shows, at month 2, some mice started showing signs of steatosis, at month 4 there was a significant increase in steatosis with some not significant inflammation and hyperplasia, at month 6 high levels of steatosis continued in addition to inflammation in more mice, and by month 9 all of the predictor factors were increased in mice fed a WD, suggesting a full spectrum of NAFLD (Figure 17A).

Based on histological scoring and histological changes observed, we proposed labeling the time points by the disease status observed. That is, in the diet study, even though there were minimal histological changes at month 2, in the prognostic factors we measured, livers were enlarged, and mice were obese, which is consistent with what is observed in humans with NAFLD. For months 4 and 6, there were significant levels of fat accumulation, consistent with what is seen in NAFL in humans. Similarly, at month 9, we were able to observe all the prognostic markers elevated, which is consistent with NASH in humans. We concluded that this diet study was able to induce the intended disease, and thus represents an appropriate model for the study of NAFL/NASH(Figure 17B).

For the diet/carcinogen study, at month 4, mice in the WD and WD/DEN also had a higher level of steatosis with no other histological reports made for the other factors. However, by month 11.5, all of the prognostic markers were present for mice in the WD and WD/DEN, including fibrosis. Mice in the NC/DEN had elevated levels of steatosis,

113

and some inflammation, which could be accounted to the DEN treatment and aging. While mice in the WD, and WD/DEN had the higher level in all these factors, specifically WD/DEN, which had the highest level of macrosteatosis, inflammation, and fibrosis, which are the factors linked to more advanced diseases (Figure 17C).

In the diet/carcinogen study, as in the other study, month 4 demonstrated what is commonly seen in NAFL while month 11.5 had all the prognostic factors of NASH, in addition to other histological changes that are observed in advanced liver disease and those with a possible progression to HCC. We concluded that this diet/carcinogen study was able to induce the intended diseases and propose this as a NASH/advanced liver disease study (Figure 17D).

Table 9. Scoring system used by veterinary pathologists for liver tissues in both studies

| Timecourse study of the progression of liver diseases induced by diet and carcinogen treatment (NAFLD scoring and tumor burden	
Objective used	10x
Fields	5 fields per lobe, chosen randomly
Scoring steatosis and hypertrophy	Macrovesicular steatosis, Microvesicular steatosis, hypertrophy scored at (0=<5%, 1=5-33%, 2=33-66%, 3=>66%
Scoring inflammation	0,1=0.5-1 foci, 2=1-2 foci, 3=>2 foci
Scoring fibrosis	Binary scoring (yes/no)
Notes	Hypertrophy could be originating from macrosteatosis
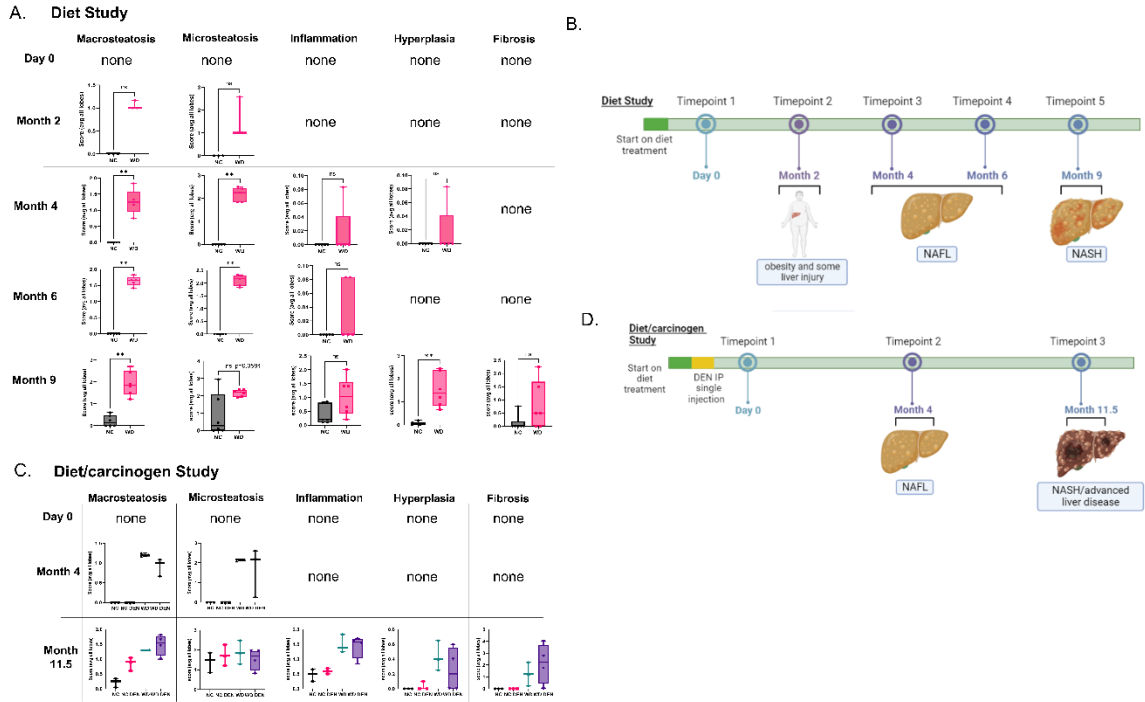
Figure 17. Pathological scoring revealed NAFL, NASH, and advanced liver disease characteristics. (A) veterinary pathology scores quantifications from liver tissue based on macrosteatosis, microsteatosis, inflammation, hyperplasia, and fibrosis for each time point, (B) Proposed model of liver disease induction based on histological changes observed in a diet study, (C) veterinary pathology scores quantifications from liver tissue based on macrosteatosis, microsteatosis, inflammation, hyperplasia, and fibrosis for each time point, (D) Proposed model of liver disease induction based on histological changes observed in a diet/carcinogen study

5.3.2 Characterization of N-linked modifications in NAFLD progression

To characterize the N-linked glycan profile in liver disease progression, we used Matrix-Assisted Laser Desorption Ionization (MALDI) Imaging Mass Spectrometry (IMS). We were first interested in the full N-glycome characterization of NAFL/NASH progression based on the type of N-glycans that were identified in all samples. Analysis was based on the following types of N-glycans: high mannose, fucosylation, extensive branching (non-fucosylated), bisecting (non-fucosylated), bi-antennary (non-fucosylated), tri-antennary (non-fucosylated), and hybrid (non-fucosylated). Figure 18A outlines the N-glycan structure distribution for all groups; out of all the N-glycans identified in all samples, 44% N-glycans were fucosylated, 9.8% were high mannose, and 11.4% were bi-antenna (non-fucosylated). Other N-glycans had a less than 10% distribution. We furthered our analysis based on the relative intensity of each structure according to the group and stages previously proposed, the N-glycan types that were changing between groups included high mannose, fucosylated, and bi-antennary (non-fucosylated) N-glycans. At the obesity/some liver damage stage, there was an increase in high mannose, but no changes in fucosylated glycan or biantennary glycan with the WD. However, at more advanced stages (NAFL and NASH), diet induced a decrease in high mannose structures with an increase in fucosylated and bi-antennary N-glycans (Figure 18B). It is important to note, that relative to early liver damage in the WD, fucosylated N-glycans increased while high mannose N-glycans slightly decreased. This suggests the importance of fucosylation in the development and during liver injury (Figure 18B).

Based on the different distribution of fucosylated and high mannose N-glycans we focused on these to further analyze if there were any changes based on disease

stage. Analysis of all high mannose N-glycans grouped revealed a decrease in these structures in the WD and this was consistent at all disease stages starting at proposed NALF (months 4 and 6) up to proposed NASH (month 9) (Figure 18C). Analysis of all fucosylated N-glycans groups revealed an increase in these glycans also starting at NAFL up to NASH (Figure 18D). Figure 18E, tracks single high mannose and fucosylated N-glycan structures at each stage of the disease and confirms the trend previously observed. In addition, it demonstrates how early these specific modification start in liver disease (month 4). Overall, N-glycan modifications can be detected very early in liver disease progression and their alterations are consistent with the disease progression.
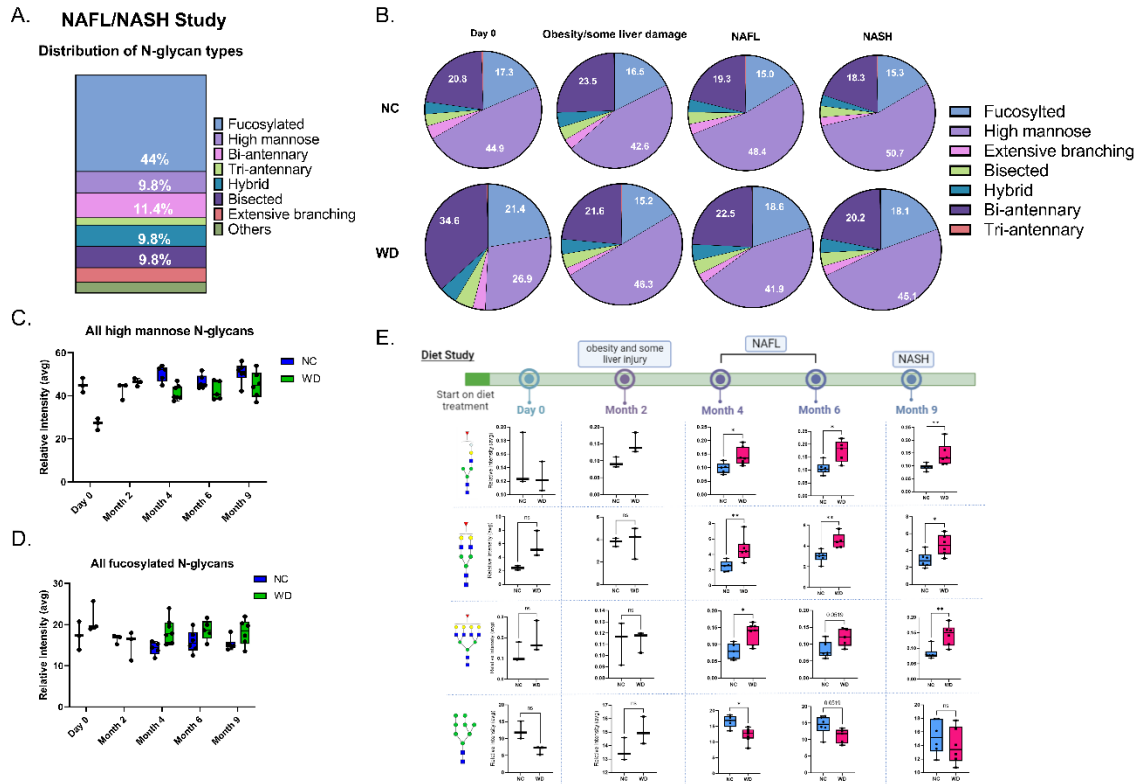
Figure 18. N-glycan modifications can be detected from early liver disease up to NASH. (A) Distribution of N-glycan types in the diet study, (B) Distribution of N-glycan modification within timepoint and diet, (C) Relative intensity of all high mannose N-glycans between diet, (D) Relative intensity of all fucosylated N-glycans between diet, (E) Relative intensity quantification of individual high mannose and fucosylated N-glycans at each time point.

5.3.3 Characterization of N-linked modification in advanced liver disease progression in the background of NAFLD

A very similar analysis was done with the NALFD/advanced liver disease study and as with the other study, the N-glycan types identified in these samples were fucosylated high mannose, and bi-antennary (Figure 19A). When analyzing their relative intensity according to groups and stages, the high mannose downregulation with diet and upregulation of fucosylated N-glycans with the WD and WD/DEN was consistent, including at the lastest timepoint (Figure 19B). A very similar trend of high mannose and fucosylated N-glycans was observed. When analyzing the sum of all high mannose (Figure 19C) and fucosylated (Figure 19D) N glycans, a decrease, and increase, respectively were observed. Figure 19E includes an analysis of single N-glycans with disease progression, this analysis confirms the trend of decreasing high mannose and increasing fucosylated N-glycans even at an advanced liver disease. Similar to the analysis in the NAFL/NASH study, bisected fucosylated and sialylated N-glycan structure was also found to be altered at an advanced liver disease stage in the WD and WD/DEN. This analysis suggests the expression of fucosylated and high mannose N-glycan are consistently modified in the progression of liver disease.
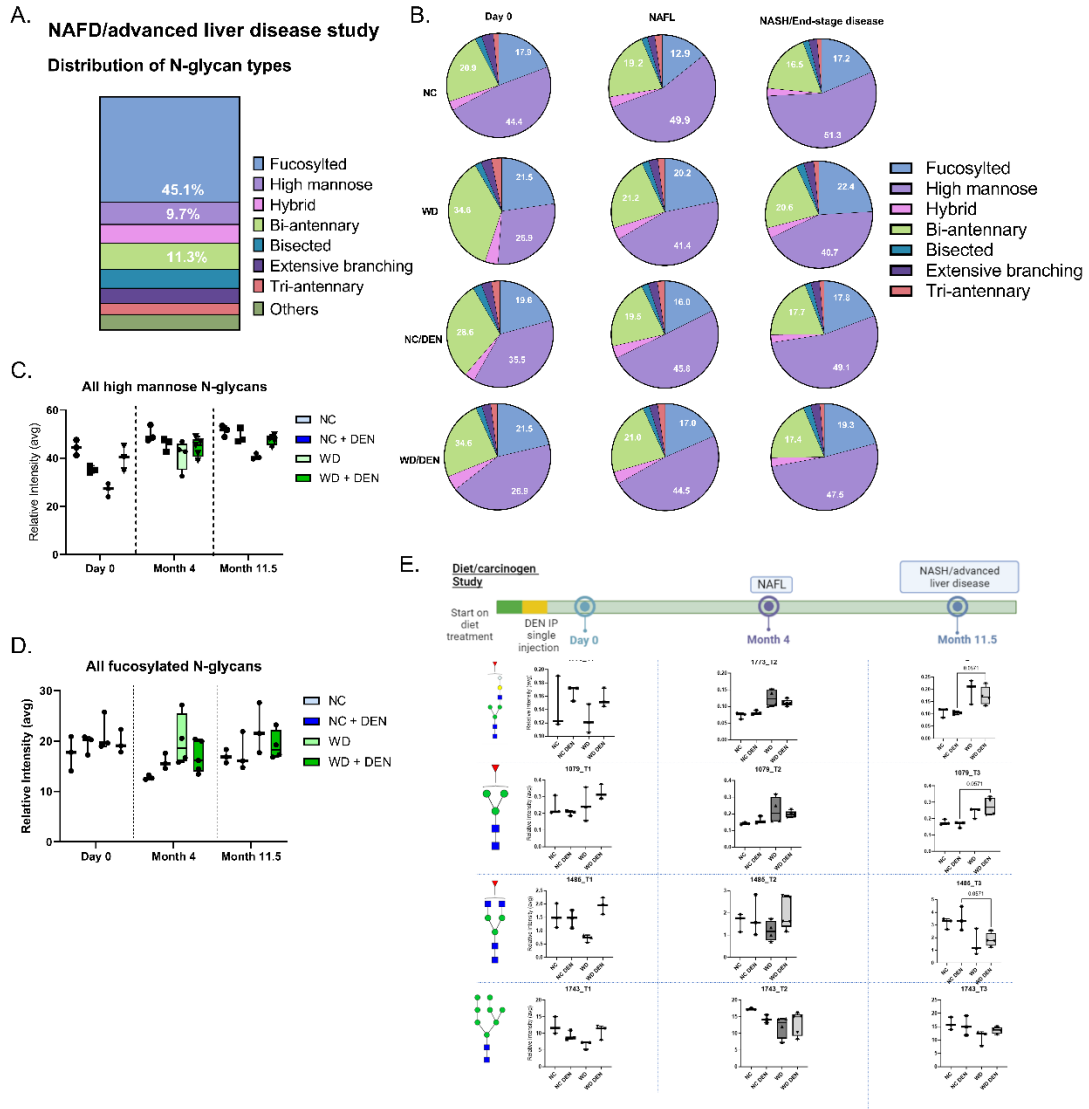
Figure 19. N-glycan modifications can be detected from early liver disease to NAFLD and up to advanced liver disease. (A) Distribution of N-glycan types in the diet/carcinogen study, (B) Distribution of N-glycan modification within timepoint and diet/treatment, (C) Relative intensity of all high mannose N-glycans between treatment, (D) Relative intensity of all fucosylated N-glycans between treatment, (E) Relative intensity quantification of individual high mannose and fucosylated N-glycans at each time point.

5.3.4 Histological and disease characteristics correlations

To take advantage of the spatial N-glycan profiling advantage MALDI IMS offers, we were interested in the pathology associated with the N-glycan modifications identified. We were specifically interested in the WD/DEN group since these mice had significant liver enlargement, liver discoloration, and nodules in multiple lobes at necropsy and had all of the characteristics of early liver disease and advanced liver disease based on the pathological scoring analysis. Veterinary pathology annotated H&E scans with all of the histological modifications observed, and while no HCC was diagnosed from these livers, hepatocyte hyperplasia was identified, suggesting that if these mice were left longer most likely they would have developed HCC. Annotations made on these samples also confirm that we are seeing very late stages of liver disease progression (Figure 20A). We overlayed these annotations with MALDI-IMS data and identified N-glycan structures that were specific to the annotations made by pathology. High mannose, bi-antennary, bisecting, and double fucosylated N-glycan structures were specific to areas annotated as nodules while single fucosylated N-glycans were more specific to areas outside of the nodules (Figure 20A). Next, we overlayed the MALDI IMS images, H&E, and immunohistochemistry (IHC) multiplex images to further analyze the immune cell profile and fibrosis changes in the WD/DEN group. This analysis revealed that N-glycan modifications are not specific to immune cell populations, CD4, CD8, or macrophages or fibrosis marker, α-Smooth muscle actin (αSMA) (Figure 20B-D) which lay outside the areas of the major N-glycans described. This suggests that while inflammation and fibrosis are one of the main characteristics of disease progression, N-glycan modifications are independent of the histological characteristics analyzed here and should be explored further for their origin with other types of cell populations.
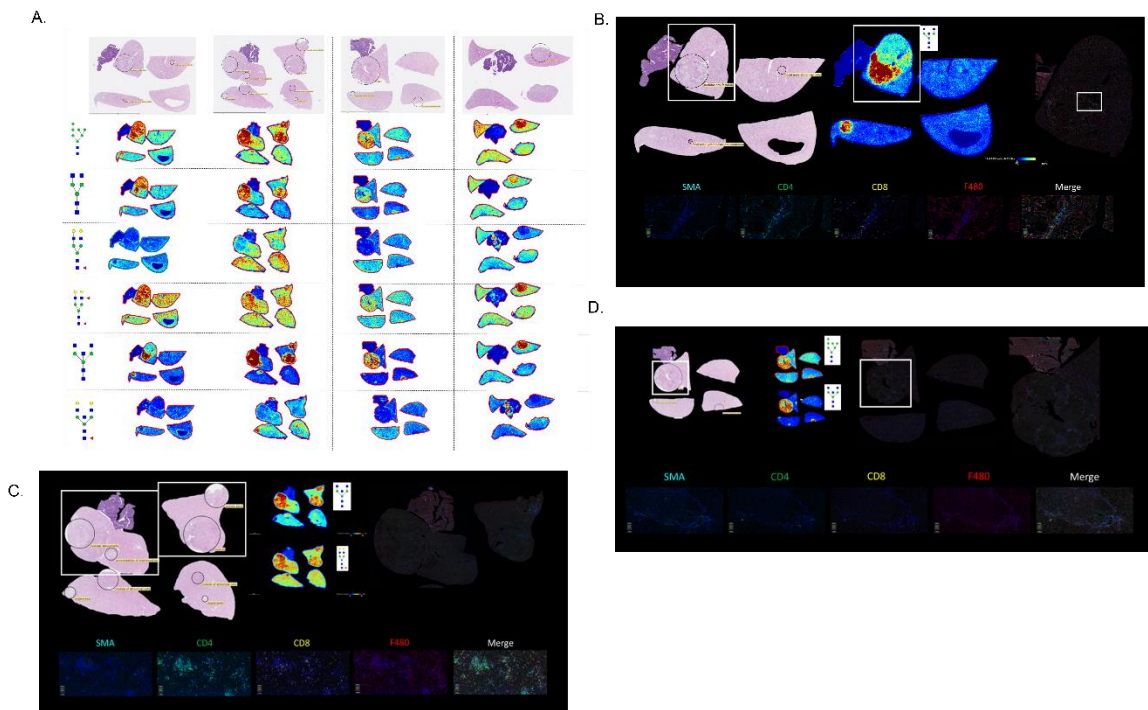
122

Figure 20. N-glycan modifications identified correlated with histopathological areas but are independent of immune cell populations (A) H&E staining of mice in the WD/DEN treatments (top panel), representative images of N-glycans structures for each tissue (bottom panels), (B-D) H&E scans, MALDI-IMS representative image, and IHC multiplex images of same tissues as in A. bottom panels are snapshots at 20x magnification for each cell marker.

5.4 Discussion

N-linked glycosylation has been previously exploited for the identification of serum biomarkers for advanced liver diseases like cirrhosis or hepatocellular carcinoma, specifically using glycoproteins with specific N-linked glycan modifications (91,229). Similarly, some studies have also identified N-glycan structure modifications in the serum of NASH patients (89). A research gap with these prior studies is that the initiation and the changes with time of these N-glycan modifications are still unknown. More specifically, previous studies did not address if the N-glycan modifications observed in cirrhosis or HCC can also be identified with minor liver injury. Elucidation of this timing in N-glycan structures would help identify possible N-glycan modifications involved in disease progression and promising biomarkers for the identification of early liver disease. In this study, we investigated the timing of N-glycan modifications with liver disease progression of early liver injury, NAFL, NASH, and advanced liver disease using MALDI-IMS N-glycan analysis coupled with pathology analysis. Overall, we concluded that N-glycan modifications start at very early stages of liver disease, even before histology analysis can identify liver injury and these same N-glycan modifications can be detected at advanced stages of the disease. Our results suggest that N-linked glycosylation, specifically fucosylation might have a role in disease progression.

Diet and carcinogen Diethylnitrosamine (DEN) treatments used in this study induced the expected level of liver disease. However, mice treated with the combination of diet and DEN carcinogen did not lead to the development of HCC. We hypothesized this might be due to a late start of DEN treatment, which in our studies was at 4 weeks old, instead of the common model of DEN injections at 2 weeks old. However, based on histological analysis by pathology, we still consider that the level of diseases that were

124

induced in both studies (a diet-inducible NAFL/NASH study and a chemically inducible advanced liver disease study in the background of NALFD) were consistent with what is observed clinically (obesity, steatosis inflammation, fibrosis, and weight loss at advance disease) and was relevant for N-glycan characterization.

The opposing trend of N-glycan modifications between high mannose and fucosylated N-glycan structures that were observed in both studies could be explained by alterations in the N-glycan processing pathway, where fucosyltransferases could be overtaking and inducing some of the disease phenotypes observed while high mannose N-glycans are decreasing with disease due glycoproteins involved in diet/carcinogen processing being modified further into branched and fucosylated structures. To further explore these findings, the expression of glycosyltransferases and N-glycan processing genes should be investigated further.

An unexpected finding in this study was that at day 0 in both studies, we found slight N-glycan alterations. While these were not statistically significant, there are still some changes within groups. It is important to note that the main difference within the group is that mice were left in the respective treatment for at least 4 hours and harvested after-within the day 0-time point. we hypothesized that this might be a very early response to diet in terms of the metabolic processing of glucose intake present in a high-caloric diet. It is also important to note that the trend of upregulation/downregulation between fucosylated and high mannose structure was more significant in mice in the WD relative to mice in the WD/DEN. Suggesting that the use of a carcinogen might induce a different mechanism of disease and that a high-caloric diet by itself is enough to observe an N-glycan profile that is consistent with advanced liver disease.

Importantly, we elucidated that N-glycan modifications can be observed at very early stages of the disease. Based on pathology scoring analysis, the earliest significant liver tissue injury (steatosis and inflammation) that can be observed by histology is at month 6. However, our N-glycan data demonstrated that N-glycan modifications start with a slight trend at month 2 and become significant by month 4. IHC multiplex imaging of immune cell populations and fibrosis markers confirmed that the origin of these N-glycan modifications is independent of an inflammatory and fibrotic response histologically.

To conclude, this study identifies the N-glycan modifications during stages of liver disease progression, we report that N-glycan modifications can be observed at the early stages of the disease, even before histological analysis can identify tissue injury. This study suggests N-glycan modifications should be considered useful as early disease markers and could also be playing a role in liver disease progression to inflammation, fibrosis, and advanced liver disease characteristics.

Chapter 6: Analysis of N-linked glycan alterations in tissue and serum reveals promising biomarkers for intrahepatic Cholangiocarcinoma

Analysis of N-linked Glycan Alterations in updates Tissue and Serum Reveals Promising Biomarkers for Intrahepatic Cholangiocarcinoma

Shaaron Ochoa-Rios, Calvin R.K. Blaschke, Mengjun Wang, Kendell D. Peterson, Andrew DelaCourt, Stéphane Elie Grauzam, David Lewin, Peggi Angel, Lewis R. Roberts, Richard Drake, and Anand S. Mehta

Corresponding Authors: Anand Mehta: mehtaa@musc.edu

6.1 Introduction

Cholangiocarcinoma (CCA) is an epithelial cancer arising in the biliary mucosa lining the ducts that carry bile from the liver to the small intestine (49). CCA is the second most common type of liver cancer after hepatocellular carcinoma (HCC) and is a highly lethal cancer (50). CCA is classified based on the anatomic location into intrahepatic CCA (iCCA), perihilar CCA (pCCA), or distal (dCCA) subtypes (51). Cholangiocarcinomas frequently consist of small nests of epithelial cancer cells surrounded by dense stromal regions of cancer-associated fibroblasts, immune cell populations, and extracellular matrix. In addition, glandular formations are also a histologic characteristic of cholangiocarcinoma (239). CCA risk factors can be subtype-specific: Diabetes, obesity, smoking, chronic viral hepatitis, cirrhosis, chronic pancreatitis, and chemical exposure have been linked to iCCA compared to other subtypes (28,50,51). Other risk factors have been associated with geographic regions, where primary sclerosing cholangitis (PSC) is the most common risk factor in Western countries and liver fluke infection in Southeast Asia (28). Despite advances made so far in understanding the risk factors and biological mechanisms of the disease, a definitive diagnosis of CCA at an early stage continues to be challenging. For CCA diagnosis, a combination of diagnostic methods is necessary and a biopsy should be performed when feasible and taken into consideration for the final diagnosis (51). Conventional diagnostic techniques do not account for the heterogeneity of the tumor location, size, and pathological and cellular characteristics (49,50). Consequently, there is an urgent need for improved biomarkers and treatments for CCA. Glycosylation, one of the most common post-translational modifications, regulates biological functions including cell-cell communication, protein folding, and receptor signaling (49,70). Dysregulation in glycosylation has been reported in many cancers,

including CCA (49,80,91,240). N-linked glycosylation is the addition of glycan structures to the glycoprotein at the asparagine residue. This process occurs through a well-established biosynthetic processing pathway in the endoplasmic reticulum (ER) and Golgi apparatus. The composition of these N-glycan structures is influenced by the abundance of glycosyltransferases and glycosidases, which add and remove monosaccharides, respectively, as well as the availability of nucleotide-monosaccharide donors (70,84). N-glycan alterations like fucosylation (addition of a fucose residue), sialyation (addition of a sialic acid residue) and complex branching (addition of a GlcNAc residue) have well-established associations with many different cancers, including HCC (227,241–243). Alterations in fucosylation can be classified as (1) core fucosylation or (2) outer-arm fucosylation. Fucosyltransferases (FUTs 1-11) catalyze the addition of the respective fucose residue at a specific linkage of the N-glycan structure. FUT8 is the only enzyme capable of adding a fucose residue at the α1-6 linkage of the N-glycan structure (known as core fucosylation). While the remaining FUTs can add fucose residues at different linkages on the antenna of the N-glycan structure (known as outer-arm fucosylation). Due to the many biological roles, N-glycosylation is responsible for, N-glycan structures can also be altered in healthy tissue. We have previously reported N-glycan alterations like high mannose, and bi-antennary N-glycans with limited fucosylation and branching in healthy liver tissue (76). N-glycosylated antigens have been widely used as biomarkers for different types of cancers: alpha-fetoprotein (AFP) for HCC, prostate-specific antigen (PSA) for prostate cancer, and Carbohydrate Antigen 19-9 (CA19-9) for CCA. However, the low specificity of CA19-9 reduces its clinical utility, and a higher-performing biomarker for CCA is urgently needed. Previous research has identified alterations in the abundance and/or N-glycosylation of certain serum or plasma glycoproteins that are correlated to liver

damage caused by CCA and CCA tumor progression (49,80,240). We, and others, have previously correlated changes in glycosylation in the tissue and serum of individuals with the development of HCC (68,76,78,154,244–247). However, it is unclear if a similar change would be observed in the other major type of liver cancer, CCA.

To address this gap, we utilized matrix-assisted laser desorption ionization (MALDI) imaging mass spectrometry (IMS) for the characterization of the N-glycan-related molecular changes occurring in tissue and serum. A total of three cohorts were analyzed, consisting of two tissue cohorts: a discovery cohort (n=104 cases) and a validation cohort (n=75), and one independent serum cohort consisting of patients with iCCA, HCC, or benign chronic liver disease (n=67). Finally, we use the identified N-glycomic profiles for the development of a potential biomarker that could distinguish iCCA from other types of liver damage (75,248).

6.2 Materials and Methods

6.2.1 Tissues and Tissue Microarrays

Initial analysis was performed using normal liver tissue (HuFPT074), hepatocellular carcinoma (HCC) tissue (HuCAT081), and intrahepatic cholangiocarcinoma (iCCA) tissue (HuCAT086) (Biomax, Inc., Rockville, MD). H&E stains of each tissue were annotated by a pathologist with the tumor, adjacent to the tumor, and fibrotic regions. Subsequently, analysis was performed using a tissue microarray (TMA) (#LV2081, Biomax, Inc) that contained 208 cores with 103 cases (duplicated cores per case): consisting of fifty HCC, twenty iCCA, one clear cell carcinoma cyst, five metastatic HCC (spleen, chest wall, cerebrum, costal bone, and lymph node), two hepatic cyst, eight tissues with cirrhosis and dysplastic nodules, ten hepatitis infected tissues, two adjacent normal tissues, and six independent normal tissues. The cores were 1.0 mm in diameter

and validated by pathology. The validation tissue cohort TMA (DLV753, US Biolab, Rockville, MD) contained 75 cases with 75 cores: forty-five cases of HCC, twenty-three cases of iCCA, two cases of mixed carcinoma, and five cases of normal liver tissue. Cores were 1.5 mm in diameter and validated by pathology. All tissue samples were formalin-fixed paraffin-embedded (FFPE) cut into a 5 µm thick section and unstained before analysis.

6.2.2 Serum samples

Samples were from 30 patients with iCCA, 17 samples with primary sclerosing cholangitis (PSC), and 20 patients with other liver diseases but not iCCA or PSC saw at Mayo Clinic, Rochester, MN between January 2000, and May 2010. Peripheral blood was collected from each participant at the time of the office visit before treatment. Sera were stored at -80°C. The following data elements were abstracted from the medical record: demographics (age, gender, ethnicity, race, weight, height), medical history, etiology of liver disease, laboratory data including CA19-9 and AFP, and imaging results (US, CT, or MRI). Histopathology results and radiologic findings from the medical records of all patients were reviewed to ascertain the diagnosis of iCCA and identify tumor location. The diagnosis of iCCA in all patients was confirmed by histopathology. The anatomic location of cholangiocarcinomas were categorized as "intrahepatic" if the mass lesion arose within the hepatic parenchyma and did not extend to or involve the secondary branches of the biliary trees as demonstrated either by computerized tomography imaging, magnetic resonance imaging or endoscopic retrograde cholangiopancreatography findings. The etiology of liver disease was based on the laboratory, imaging, and histopathology results and the judgment of the treating physician. For patients with viral hepatitis, anti-HCV

antibody, serum HCV RNA, HBV surface antigen, HBV e-antigen, and HBV DNA levels were recorded.

6.2.3 Enzymes and Reagents

Trifluoroacetic acid, Harris-modified hematoxylin, and α-cyano- 4-hydroxycinnamic acid (CHCA) were obtained from Sigma- Aldrich. HPLC grade methanol, ethanol, aceto-nitrile, xylene, hydrogen peroxide, and water were obtained from Fisher Scientific. Recombinant peptide N-glycosidase F (PNGase F) PRIME™ and endoglycosidase F3 Prime™ (Endo F3) were obtained from N-Zyme Scientifics (Doylestown, PA).

6.2.4 Tissue preparation for MALDI-IMS

Unstained FFPE tissue slides were processed using standardized N-glycan imaging workflows of matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-IMS) as described (76,127) Briefly, tissues were heated at 60°C for 1 hour and cooled to room temperature before deparaffinization. The slides were washed with xylene to remove the paraffin and then rehydrated using a series of water and ethanol washes. Antigen retrieval was performed using citraconic anhydride (Thermo Scientific) as the buffer and placed in a decloaker for 30 minutes. The buffer was then cooled to room temperature and buffer exchange was performed to replace the slides in 100% water. To release the N-glycans, an M5 TM-Sprayer™ Tissue MALDI Sample Preparation System (HTX Technologies, LLC) was used to spray 0.5 mL of 0.1 µg/µl aqueous solution PNGase F PRIME™ (N-zyme Scientifics) as previously described (76,217). To elucidate the core fucosylated N-glycan profile, we treated the validation TMA with the enzyme Endo F3 Prime™, which preferentially cleaves core-fucosylated N-glycans between the two core N-acetylglucosamine residues and results in a mass shift of 349.137 m/z for core fucosylated N-glycans compared to N-glycans released only by PNGaseF (29). Following

the spray of the respective enzyme, the slides were placed in a humidified chamber and incubated at 37°C for 2 hours. Slides were then desiccated and dried before matrix application. The matrix was assembled using α-cyano-4-hydroxycinnamic acid (0.042 g CHCA in 6 mL 50% acetonitrile/49.9% water/0.1% TFA) and sprayed using the same M5 TM-Sprayer™.

6.2.5 Serum preparation for MALDI-IMS

Serum samples were processed for N-glycan analysis as previously described (126). Briefly, 1 ul of serum was diluted in 2 ul of sodium bicarbonate (100Mm, pH 8.0). After mixing, 1 ul was spotted on a hydrogel-coated slide. Serum samples were spotted in triplicate and a blank well was always included in the same slide as the serum samples. Spots were left to immobilize onto the slide at room temperature for 1 hour and washed a total of three times with Carnoy's solution and one time with double distilled water. PNGase F PRIME™ and matrix were sprayed for tissue samples described above.

6.2.6 N-Glycan Imaging using MALDI-IMS

All tissues and serum samples were imaged using a timsTOF Flex MALDI-QTOF mass spectrometer (Bruker Daltonics) (m/z 500–4000) operating in a positive mode. Focus Pre TOF parameters were set as followed: transfer time 120.0 μs and pre-pulse storage 25.0 μs. For whole tissues (normal, HCC, and iCCA) and TMAs images were collected at 100 μm raster. For serum samples, images were collected at 150 μm raster. All images were collected at 200 laser shots per pixel.

6.2.7 Data Processing

Data analysis was done in SCiLs lab 2021 imaging software (Bruker) for analysis of the mass range m/z 500 to 4000. SCiLs-generated N-glycan spectra were normalized

to the total ion count (TIC). N-glycan structure annotations for each m/z (mass to charge ratio) value were made based on an in-house database of known N-glycan generated using GlycoWorkBench and GlycoMod for annotation or to MS/MS data previously done by our group (63,219). For tissue, the maximum mean value of each m/z was extracted and used to calculate the relative intensities for each of the N-glycan peaks identified. For serum analysis, the following steps were applied to the extracted maximum mean value data: a blank well was included in the experimental plan and was subtracted from each N-glycan. Only N-glycans present in at least 80% of the samples were used for analysis. The standard sample was used to create normalization factors for each N-glycan on each slide, where the intensity of the individual slide's standard was divided by the average intensity of the standards across all slides for each N-glycan. Each slide's N-glycans were then multiplied by the corresponding normalization factor and the glycan intensities were converted to the relative intensity. A list of all N-glycans identified in tissue and serum can be found in Supplementary Table 1.

6.2.8 Statistical analysis

N-glycan data was explored with a boxplot and scatter plot (outliers were included). Descriptive statistics were presented as mean values +/- standard deviation unless otherwise stated. Statistical inference between two groups was applied with a t-test or Mann-Whitney test based on the distribution of the data, normality of data was checked by the Shapiro-Wilk test. To explore the statistical inference of associations between outcomes with glycans, we applied logistic regression for a binary outcome, ordinal logistic regression, tree analyses, and random forest algorithm for multiple ordinal outcomes (data not shown). Logistic regression was also used to derive the rate of change and corresponding p-value of each glycan change from those without iCCA to those with iCCA

in both TMA and serum. Receiver operating characteristic curves (ROC curves) were built to assess the discriminating ability of individual glycans or panels of glycans. The Area under the ROC curve (AUC) was also chosen as a criterion in each step of feature selection to remove noninformative glycans or redundant glycans during optimization.

To alleviate the bias of feature selection, we applied various methods such as step-wise logistic regression, Lasso algorithm, correlation filter, random forest algorithm for feature selection to optimize the combination of glycans, and cross-validations (Leave-one-out cross-validation (LOOCV) and 3-fold CV) were further applied during feature selection to avoid bias (249). Based on the robustness, interpretability, and suitability for the glycan panel structure of interest, we chose a logistic regression model to build a predictive algorithm with selected features. The relative importance of the chosen glycans in the algorithm was derived. The performance of the predictive algorithm was explored by apparent validation (classification) and LOOCV. AUC was calculated, its standard error was derived using a bootstrap method with 2000 iterations, and the 95%CI of AUC and corresponding p-value were derived from the standard error. Statistical comparison between two AUCs was performed using Delong's test.

Cluster analysis was also applied to explore the similarity of the data structure of the glycans of interest based on statistical distances. Tanglegram (Cophylo plot) of serum and TMA were plotted to explore the congruence of the two dendrograms (250,251). Principal component Analysis (PCA) was used for further exploring glycan data structure based on its variation-covariance (information of each glycan), biplots of selected glycans of serum and TMA are provided for visual inspection in the relationship of principal components.

Statistical analysis was performed by Graph Pad Prism 9.0 software package (Graph Pad, Inc, San Diego CA.) and R (version 4.1, https://www.r_project.org)

6.2.9 Study approval

Human serum samples were obtained from the Mayo Clinic Hepatobiliary Neoplasia Registry and Biorepository under an IRB-approved protocol. This study was approved by Mayo Clinic Institutional Review Board and met all guidelines set forth by the 1975 Declaration of Helsinki for good clinical practice. All participants provided informed consent for this study.

6.3 Results

6.3.1 N-glycan alterations correlate to histopathological changes

To elucidate the in-situ N-glycan changes that occur in iCCA tissue, we utilized MALDI-IMS methodology. Relative intensities across each tissue are presented via a heat map of individual m/z (mass to charge ratio) values, where blue is low abundance and red is high abundance. Here, m/z values are representative of specific N-glycan structures. In Figure 21, specific N-glycans that are altered with the development of iCCA, and HCC are shown based on the histopathological changes annotated by a pathologist. N-glycans at m/z 1809.646 and 2012.717 (bi-antennary and bisected fucosylated N-glycans, respectively) were predominantly expressed in iCCA tissue, specifically in the tumor region while in the HCC tissue, these were present in only the fibrotic areas within the tumor region (Figure 21A-B, and 19E). However, an N-glycan at m/z 2393.840 (complex highly branched N-glycan) was highly expressed in HCC tissue compared to iCCA and was also expressed in the tumor region of iCCA tissue but at a lower intensity (Figure 21C

and 19E). Finally, an N-glycan at m/z 1905.630 (high mannose N-glycan) did not show a definitive localization to the histology of iCCA or HCC tissues (Figures 20D and 20E). Figure 22A-B includes H&E staining scans for the iCCA and HCC tissues used. Overall, we demonstrate by MALDI-IMS that the origin of specific N-glycan modifications can change based on the type of liver cancer or the histopathological changes in each tissue.
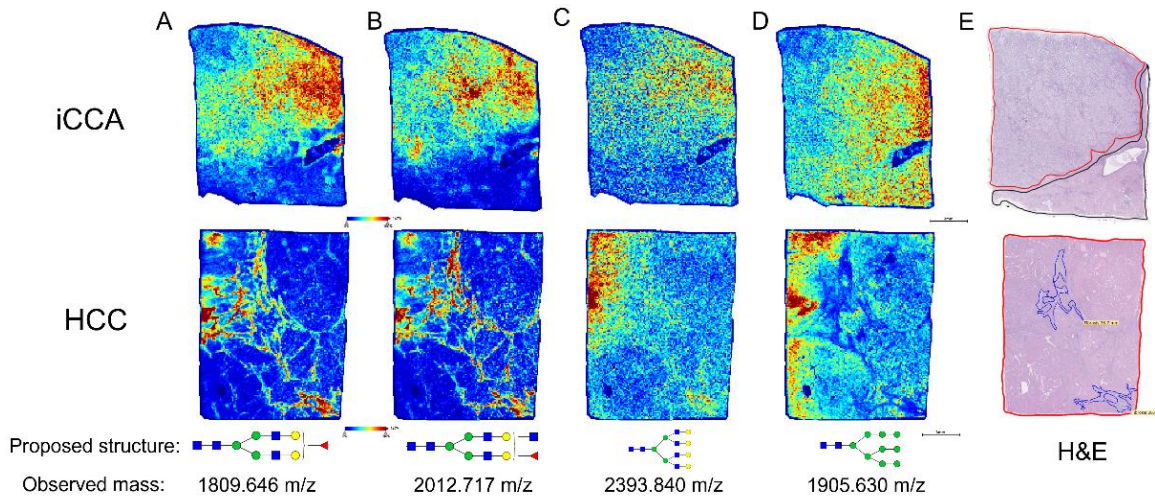
Figure 21. Bisected and biantennary fucosylated structures are highly expressed in iCCA tumor. Representative images of the relative intensity of biantennary fucosylated N-glycan (1809.646 m/z). (A) bisected fucosylated N-glycan (2012.717 m/z). (B) tetraantennary branched N-glycan (2393.840 m/z). (C) high mannose N-glycan (1905.630 m/z). (D) proposed N-glycan structures at the bottom correspond to the respective m/z value (observed mass). (E) H&E staining. Tumor regions are outlined in red, normal areas are outlined in black, and fibrotic regions are outlined in blue. Intrahepatic Cholangiocarcinoma (iCCA), and Hepatocellular Carcinoma (HCC). For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose.
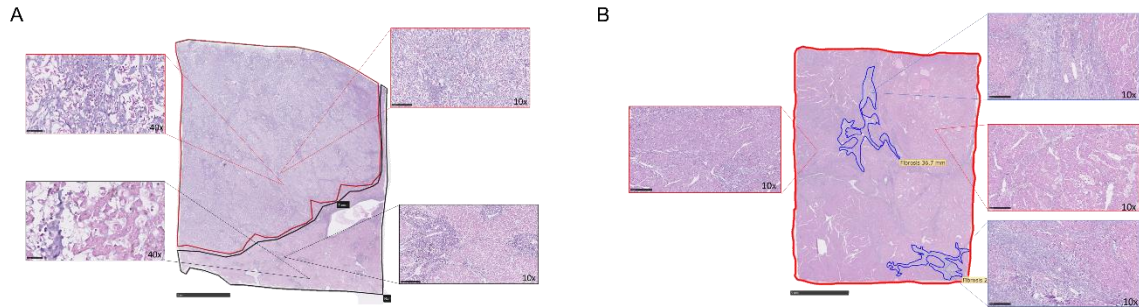
Figure 22. Hematoxylin & Eosin (H&E) staining of (A) Intrahepatic
Cholangiocarcinoma (iCCA) tissue, 10x (right) and 40x (left) magnification images from
respective areas of the tissue, and (B) Hepatocellular Carcinoma (HCC) tissue, 10x
magnification images from respective areas of the tissue. Tumor regions are outlined in
red, normal areas are outlined in black and fibrotic regions are outlined in blue.

6.3.2 Bisected fucosylated N-glycan alterations are specific to iCCA patients in tissue

To determine whether the N-glycan alterations previously identified in iCCA tissue could be observed in a larger set of tissue samples, we examined a discovery tissue microarray (TMA) consisting of a total of one-hundred-four tissue cores: fifty HCC, twenty iCCA, one clear cell carcinoma cyst, five metastatic HCC (spleen, chest wall, cerebrum, costal bone, and lymph node), two hepatic cysts, eight tissues with cirrhosis and dysplastic nodules, ten hepatitis infected, two adjacent normal tissues, and six independent normal tissues. Figure 23A-E shows representative images from this discovery TMA (TMA 1) demonstrating the N-glycan changes from MALDI-IMS N-glycan imaging data. As Figures 23 A-C show, specific N-glycans at m/z values of 1850.667 (presumed bisected N-glycan with one fucose residue), 1996.720 (presumed bisected N-glycan with two fucose residues), and 2158.791 (presumed bisected N-glycan with two fucose residues and with an additional galactose residue) were found primarily in iCCA tissue (Figure 23A-C). In contrast, N-glycan at m/z 2393.824 was specific to HCC tissue when compared to iCCA or normal samples (Figure 23D). Additionally, N-glycan at 1905.644 followed a similar trend as in Figure 21, where it was expressed in all groups with a decrease in the iCCA samples compared to other groups (Figure 23E). Normal tissue was associated with high mannose N-glycans, and bi-antennary type N-glycan with and without terminal sialic acid, with limited levels of fucosylation and branching, as we have observed before (76).

Next, we used a second TMA (TMA 2) for validation which consisted of a total of seventy-five tissue cores: forty-five HCC, twenty-three iCCA, two mixed HCC-iCCA, and five normal liver tissues (Figures 22F-J). We identified the same N-glycans in our validation TMA as in our discovery of TMA in iCCA tissues. As before, bisected fucosylated N-

glycans at m/z values of 1850.667, 1996.720, and 2158.791 are significantly increased in iCCA as compared to the other groups (Figure 23F-H), while N-glycans at 2393.824 and 1905.644 are either not elevated or under-expressed in iCCA (Figure 23I-J). Each N-glycan alteration from the two TMAs was further analyzed by comparing their relative intensity in iCCA tissue to non iCCA tissue which included normal, and other liver diseases (described in TMA 1). As figure 23K-M show, bisected fucosylated N-glycans were significantly altered in the iCCA tissue (Figure 23K-M) while complex highly branched and high mannose N-glycans were significantly decreased in iCCA tissue (Figure 23N-O). H&E staining of each TMA and Clinical information with the diagnosis for each core for small duct and large duct classifications is provided in Figure 24A-B, respectively. While only N-glycan changes between iCCA, HCC, and normal tissues are detailed here, Table 10 details N-glycan changes in other types of liver diseases present in TMA 1.
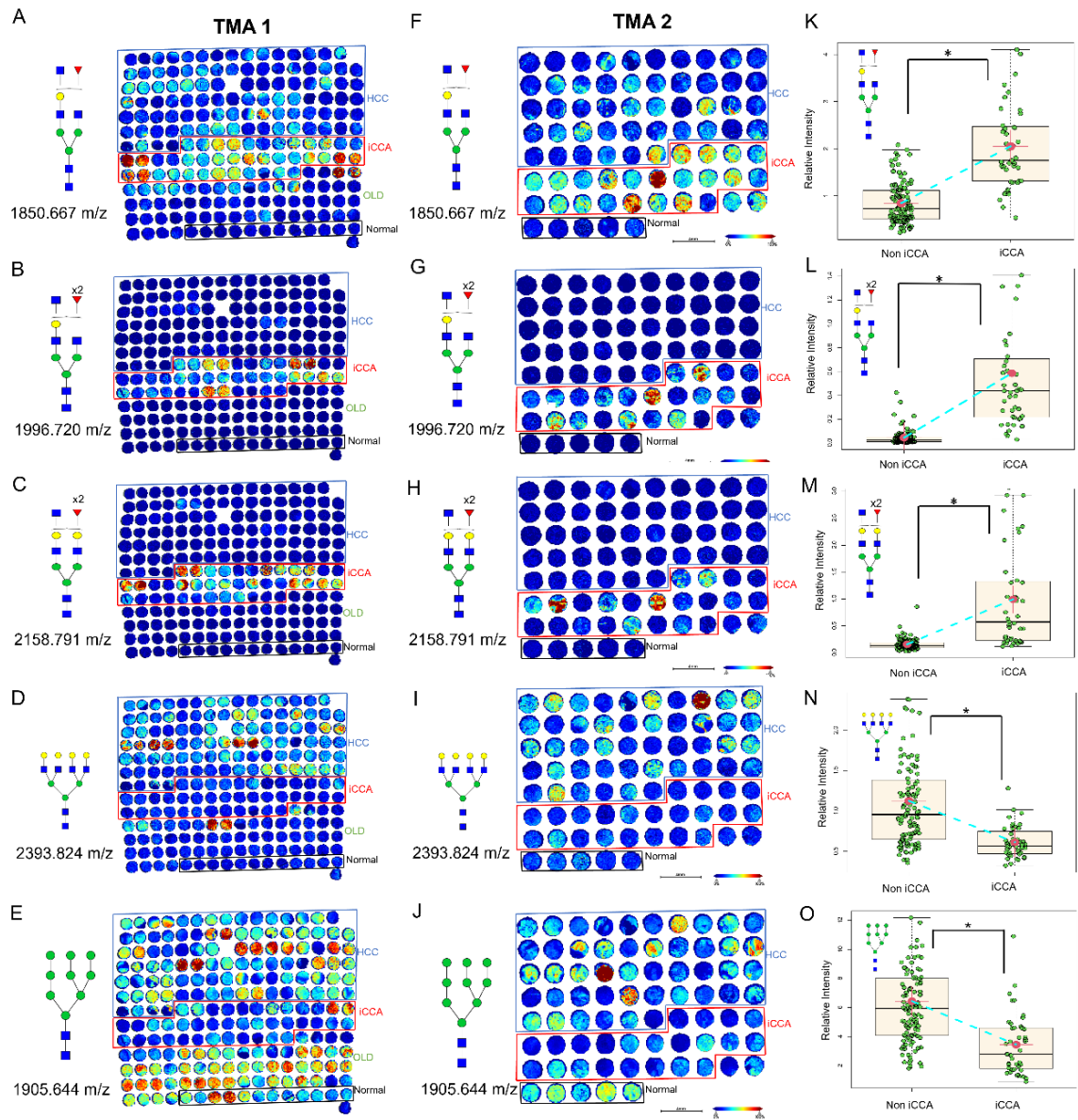
FIGURE 23. Bisected fucosylated N-glycan alterations are specific to patients with iCCA in tissue—analysis in a discovery and validation tissue cohort.

Representative images of the relative intensity of bisected/triantennary single fucosylated N-glycan (1850.667 m/z). (A) double fucosylated N-glycan (1996.720 m/z). (B) double fucosylated with a galactose N-glycan (2158.791 m/z). (C) tetraantennary branched N-glycan (2393.824 m/z) N-glycan (D) and high mannose N-glycan (1905.644

m/z) in TMA 1 (E). This TMA includes two cores per patient: 50 HCC, 20 iCCA, and 6 normal hepatic tissue. OLD: other liver diseases. (F–J) same N-glycans as in A–E but analyzed in a second independent validation tissue cohort (TMA 2). This TMA includes one core per patient: 45 HCC, 23 iCCA, and 5 normal hepatic tissue. (K–O) relative intensity quantification of both TMAs comparing Icca (n = 43) v non iCCA (n = 108). Each point in box plots represents a patient. The mass defect used for each m/z value is based on TMA 1 run. The asterisk indicates statistical difference (Mann–Whitney, P < 0.001), and error bars represent the SD. Darker red colors represent a higher intensity for the specific glycan while more blue tones represent less intensity. For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose.
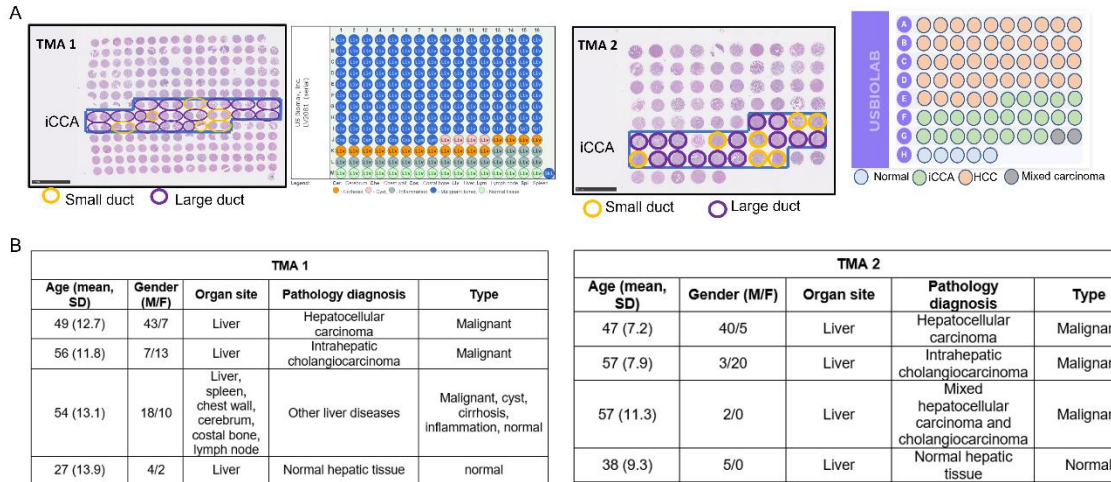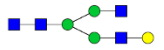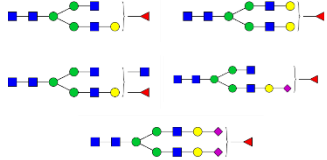
Figure 24. TMA H&E stainings and patient demographics. (A) TMA H&E staining with an outline that specifies the diagnosis for each core for TMA 1 (left) and TMA 2 (right). Mixed carcinoma: HCC and iCCA. Small (yellow) and large (purple) ducts classifications for each TMA. (B) Patient demographics for A. TMA 1 and B. TMA 2.

Table 10. Table details other pathology diagnoses included in TMA 1 with the proposed structure for the N-glycans highly expressed in each. These modifications are characterized based on 1-2 patients.

Pathology Diagnosis	Proposed structure
Metastatic HCC (spleen)	
Clear cell carcinoma	
Metastatic HCC (cerebrum)	
Metastatic HCC (lymph node)	
Metastatic HCC (chest wall)	
Metastatic HCC (costal bone)	
Hepatic Cyst	

6.3.3 Bisected core fucosylated N-glycan alterations are specific to iCCA patients in tissue

TMA 1 and TMA 2 representative images from two different bisected N-glycans each with two fucose residues (Figure 25B, 25C, and 25G, 25H) demonstrated a higher specificity to iCCA tissues relative to Figure 23A and 23F which is also a bisected N-glycan with only one fucose residue. Based on this observation, we were interested to elucidate the origin of this fucosylation, that is core fucosylation or outer-arm fucosylation (Figure 25A). We hypothesized that the specificity of bisected double fucosylated N-glycans in iCCA tissues was most likely due to core fucosylation since FUT8 and core fucosylation have been the most common modifications reported in cancer (243). Figure 25B-D, demonstrates a higher intensity and specificity of core fucosylated N-glycans to iCCA tissues in bisected N-glycans, suggesting that the fucosylation previously observed is core fucosylation. While a fucosylated tetra-antennary branched N-glycan (at 2190 m/z) had a high intensity of core fucosylated N-glycans, this was not specific to any of the groups (Figure 25E), confirming that the specificity of these N-glycan structures to iCCA samples is due to the combination of bisected and core fucosylated N-glycans.
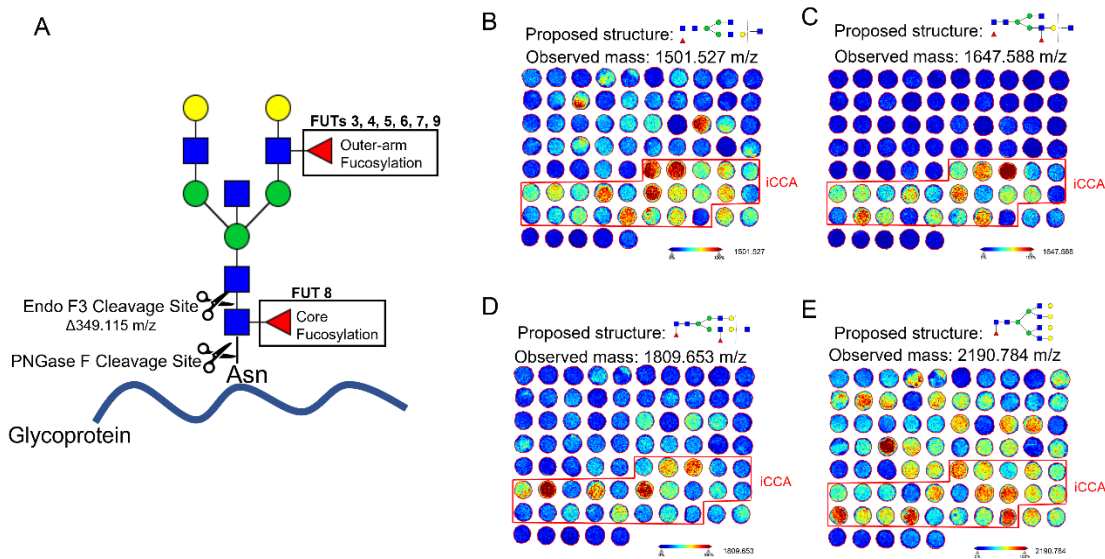
Figure 25. Bisected core fucosylated N-glycan alterations are specific to patients with iCCA in tissue. (A) Cartoon description of Endo F3 (including mass shift) and PNGase F cleavage sites, and the respective FUTs that catalyze the addition of the fucose residue. Representative images of core fucosylated N-glycans after Endo F3 treatment on TMA2, bisected core fucosylated N-glycan (B) bisected core, and outer-arm fucosylation N-glycan, (C) bisected core, and outer-arm fucosylation with two galactose residues N-glycan, (D) and tetraantennary core fucosylated N-glycan (E). For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose.

6.3.4 Characterization of N-glycosylation alterations in serum of iCCA patients

To determine the translational potential of the N-glycan modifications seen in tissue to serum assays, we utilized a MALDI-IMS serum N-glycan profiling (126)**.** Figure 26A outlines the workflow used to process serum samples. Serum imaging data were analyzed based on the relative intensity of each N-glycan and used for sample comparison between non-iCCA (samples from patients at risk of developing CCA or other types of liver disease including HCC, cirrhosis, hepatitis, PSC, and fatty liver diseases) and iCCA patients. Consistent with the trend observed in the discovery and validation TMAs, we found that the N-glycan bisected double fucosylated was significantly altered in this independent serum cohort (Figure 26B). Similarly, other fucosylated (in this case tri- and tetra-antennary) N-glycans (2174 m/z, 2320 m/z, and 2361 m/z) were also significantly altered in the iCCA cohort compared to non-iCCA (Figure 26C-E). In addition, we identified N-glycan at 1339.463 m/z, a non-fucosylated bisected N-glycan (known as the core structure of an N-glycan structure) to be significantly decreased in the iCCA tissues (Figure 26F). Overall, serum analysis revealed a very similar trend to what was observed in tissue, where bisected fucosylated and other fucosylated N-glycans continue to be highly altered in iCCA compared to any other group. Table 11 includes clinical patient information for this serum cohort.
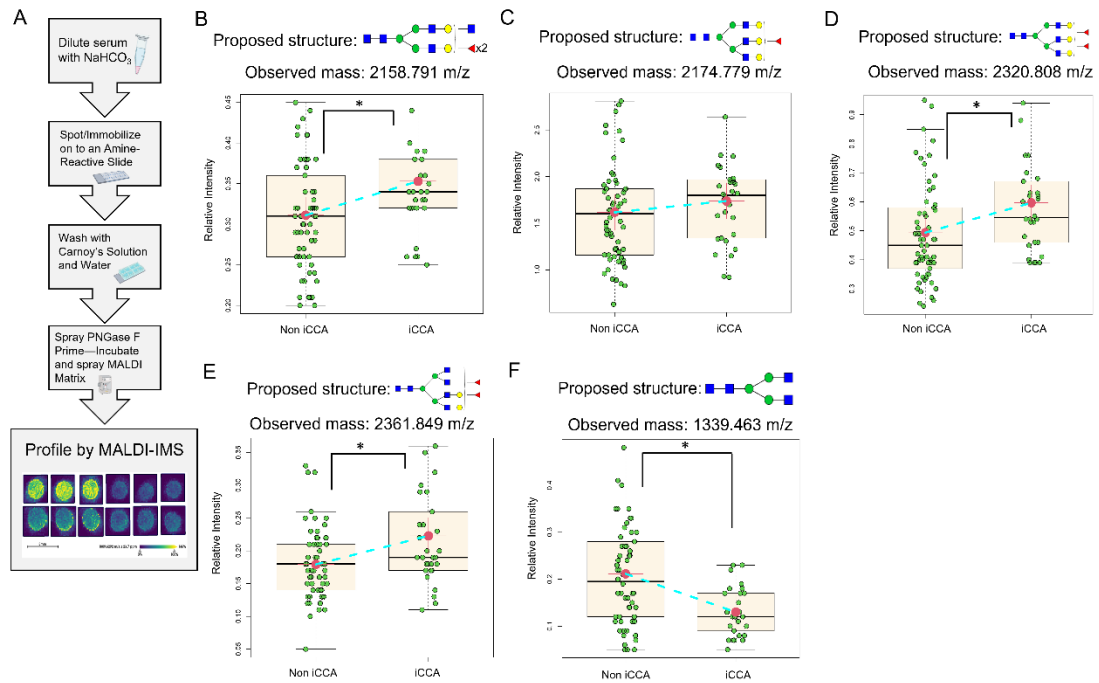
Figure 26. Fucosylated N-glycans are significantly altered in iCCA serum. (A) Serum N-glycan imaging workflow. Relative intensity quantification of bisected double fucosylated N-glycan (2158.791 m/z). (B) triantennary fucosylated N-glycan (2174.779 m/z). (C) triantennary double fucosylated N-glycan (2320.808 m/z). (D) and tetraantennary double fucosylated N-glycan (2361.849 m/z). (E) biantennary N-glycan (1339.463 m/z). (F) Each point in box plots represents a patient. The asterisk indicates statistical difference (Mann–Whitney, P < 0.001), and error bars represent the SD. NoniCCA n = 62 and iCCA n = 30. The mass defect used for each m/z value is based on TMA 1 run. For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose.

150

Table 11. Patient demographics for serum cohort. ALT (Alanine transaminase), AST (aspartate aminotransferase), AFP (Alpha-fetoprotein), and ALP (Alkaline phosphatase). Other liver diseases include nonalcoholic steatohepatitis, hepatitis C with cirrhosis, hepatic adenoma, benign fibrotic gallbladder disease, and diabetes. PSC (Primary Sclerosing cholangitis), HCC (Hepatocellular carcinoma), and OLD (Other liver diseases). Gray shading for missing clinical information

Serum Cohort		Total patients	Age mean (SD)	Gender (M/F)	CA19-9 (U/mL) (mean)	ALT (U/L) (mean)	AST (U/L) (mean)	Albumin (g/dL) (mean)	AFP (ng/mL) (mean)	ALP (IU/L) (mean)
Pathology										
Cholangiocarcinoma		30	56 (16.5)	17/13	123228.65	236	284	3.8	3.4	273
Non CCA	Primary Sclerosing cholangitis (PSC)	17	50 (10.9)		870.1176	410	845	3.4	10.5	343
	Hepatocellular carcinoma (HCC)	20	65 (8.4)	12/8		58	86	3.5	2783	150
	Other liver diseases	25	62 (9.5)	4/6						

6.3.5 Association of N-glycan modifications identified in tissue to the serum of iCCA patients

Next, we performed a comprehensive analysis of all the N-glycan modifications observed in tissue and serum to determine the association between N-glycans found in tissue and serum. As before, we used the relative intensity of each N-glycan for the respective dataset (discovery TMA and validation TMA were analyzed and are referred to here as TMAs). Analysis of TMAs and serum revealed a total of 12 N-glycans that followed the same trend (upregulation or downregulation) with a significant *p*-value in serum, TMA, or both for some N-glycans (Table 12). Next, these N-glycans were analyzed by their data structure with clustering analysis (Figure 27A) and found that 6 specific N-glycans had a very similar data structure (Figure 27B), where most of these 6 N-glycans were fucosylated with different N-glycan types: bi-antennary, bisecting, tri-antennary and tetra-antennary N-glycans (Figure 27B). We continued our analysis using the initial 12 N-glycans identified to determine their data structure based on variability (Entropy) by principal component analysis (PCA). Figure 27C for serum and Figure 27D for TMAs, show that N-glycans profile cluster into three specific types based on the level of fucosylation (First, bi-antennary, not fucosylated; second, single fucosylated; and third, double fucosylated). Biplot of first principal component (dimension 1) (accounts for 39% of the total variation in TMAs and serum datasets) and second principal component (dimension 2) (accounts for 17% of the total variation in TMAs and serum datasets). Overall, double fucosylated N-glycans demonstrated higher informative contributions in the data set. Figure 28 specifies the relative contribution per glycan to the first and second principal components indicated in plots with their relative contribution per N-glycan. A

complete N-glycan profile quantification between the serum and tissue is presented in Figure 29.

Table 12. Identified N-glycans with the same trend between TMA and serum.
[1]Observed mass-to-charge ratio (m/z) value. [2]The proposed glycan structure based upon the m/z value [3]Composition of the identified m/z value. [4]Rate of trend in TMA and serum datasets. [5]The p value in TMA and serum datasets.

m/z[1]	Proposed structure[2]	Formula[3]	Rate of trend[4]	p value[5]
1257.417 m/z		Hex5HexNAc2 + 1Na	TMA: -1.2654 Serum: -1.6486	TMA: 0.000021 Serum: 0.310389
1339.464 m/z		Hex3HexNAc4 + 1Na	TMA: -10.179 Serum: -11.1094	TMA: 0.000007 Serum: 0.001252
1501.527 m/z		Hex4HexNAc4 + 1Na	TMA: -4.3148 Serum: -3.0030	TMA: 0.000715 Serum: 0.033319
2158.792 m/z		Hex5dHex2HexNAc5 + 1Na	TMA: 8.4103 Serum: 8.3182	TMA: 0.000011 Serum: 0.012718
2174.779 m/z		Hex6dHex1HexNAc5 + 1Na	TMA: 0.5213 Serum: 0.4505	TMA: 0.000137 Serum: 0.295533
2320.808 m/z		Hex6dHex2HexNAc5 + 1Na	TMA: 0.8466 Serum: 2.8109	TMA: 0.148806 Serum: 0.022350
2361.849 m/z		Hex5dHex2HexNAc6 + 1Na	TMA: 42.0811 Serum: 9.7265	TMA: 0.000008 Serum: 0.009994
2465.871 m/z		Hex6dHex1HexNAc5NeuAc1 + 1Na	TMA: 3.5604 Serum: 3.1887	TMA: 0.001474 Serum: 0.022080
2487.861 m/z		Hex6dHex1HexNAc5NeuAc1 + 2Na	TMA: 0.5366 Serum: 1.5437	TMA: 0.271611 Serum: 0.010348
2539.896 m/z		Hex7dHex1HexNAc6 + 1Na	TMA: 0.5737 Serum: 1.4326	TMA: 0.01425 Serum: 0.507898
2685.953 m/z		Hex7dHex2HexNAc6 + 1Na	TMA: 2.1893 Serum: 1.6233	TMA: 0.004647 Serum: 0.473449
2852.954 m/z		Hex7dHex1HexNAc6NeuAc1 + 2Na	TMA: 0.9399 Serum: 3.8909	TMA: 0.446394 Serum: 0.160911
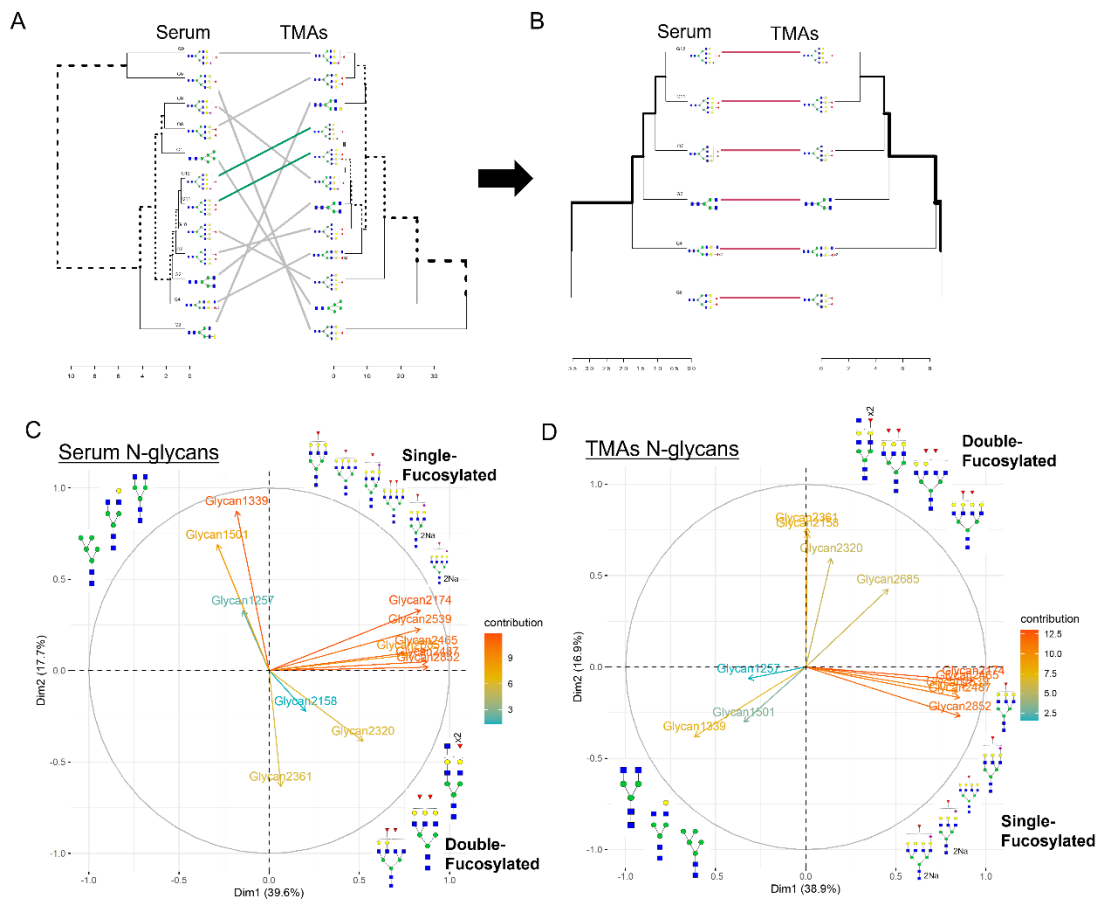
Figure 27. Profile data clustering reveals N-glycan grouping based on the number of fucose residues in iCCA serum and tissue. Dendrograms representing clustering for 12 N-glycans (A) and 6 N-glycans (B) in serum and TMAs. Glycan (G), G1: 1217, G2: 1339, G3:1501, G4:2158, G5:2174, G6:2320, G7:2631, G8:2465, G9:2487, G10:2539, G11:2685, G12:2852. PCA and their dimension (Dim) of serum (C) and TMAs (D) showing N-glycan clustering based on the number of fucose residues. 2Na (doubly sodiated N-glycan). For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose.

155
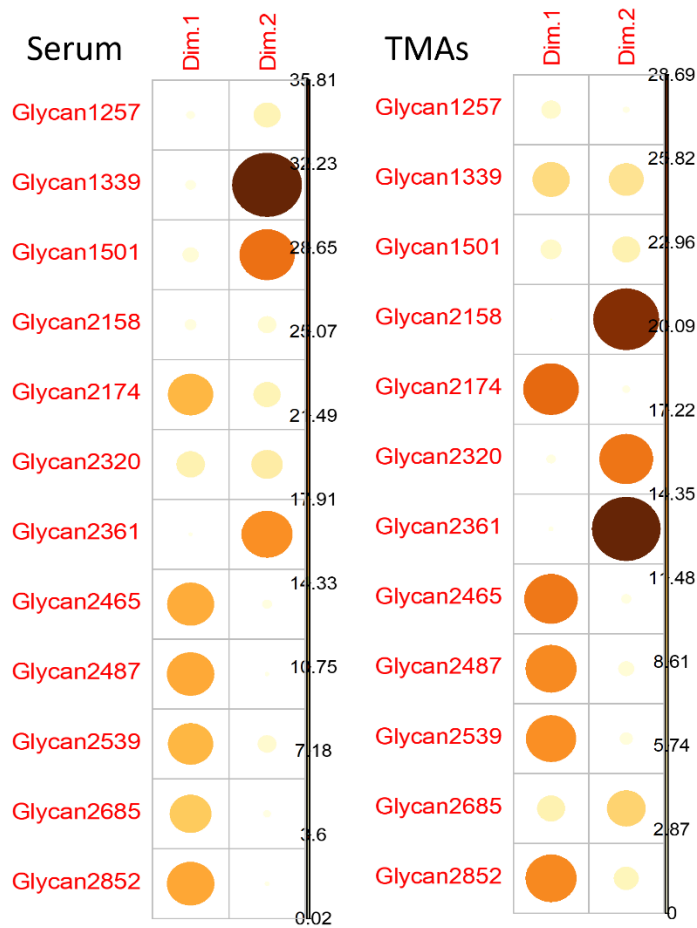
Figure 28. **Relative contribution of serum and TMA N-glycans.** (A) Relative contribution of each serum-glycan (left) and TMA-glycan (right) in the first and second principal components. The size and color of the circle represent a higher contribution of the glycan to the respective Dim. (Dimension).
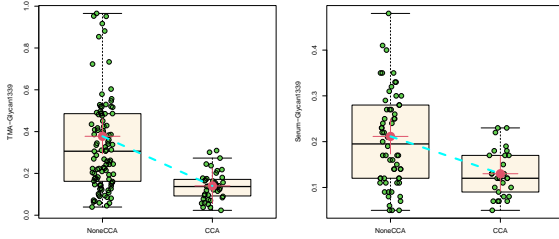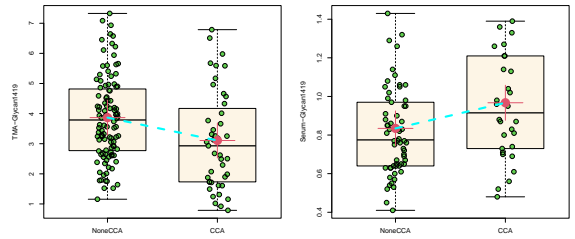
Glycan1136
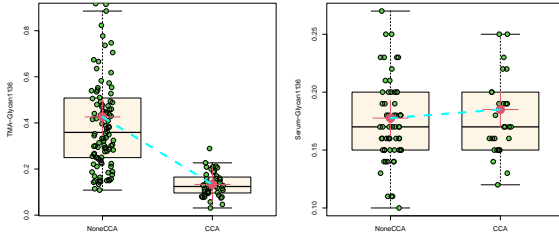
Glycan1257
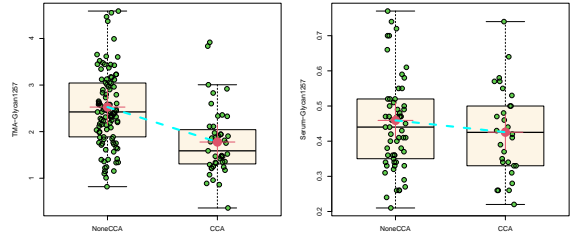
Glycan1282
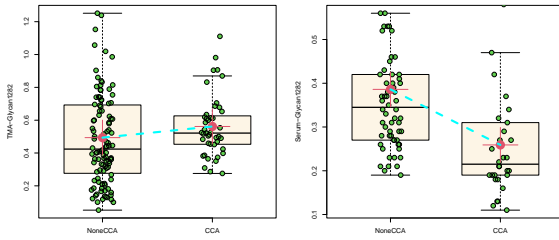
Glycan1298
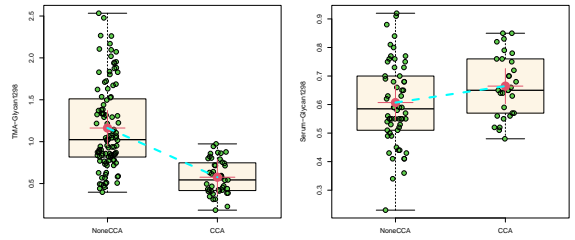
Glycan1339

Glycan1419

Glycan1444
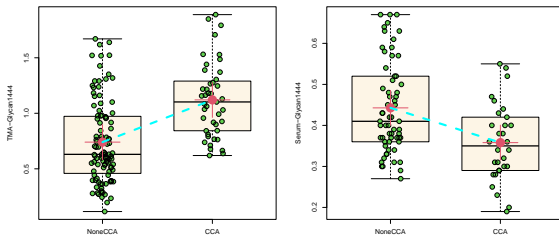
Glycan1485
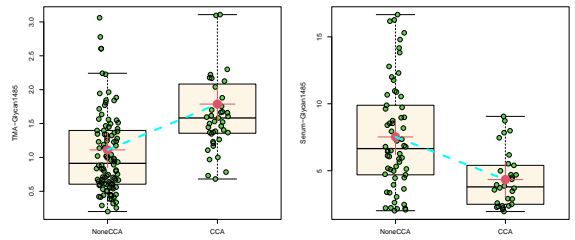
Glycan1501
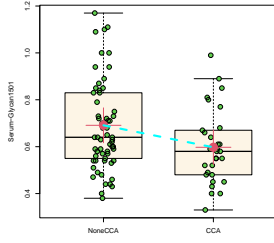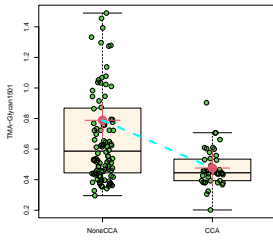
Glycan1581
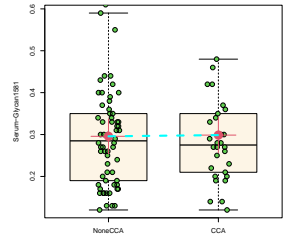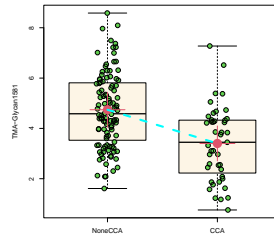
Glycan1606
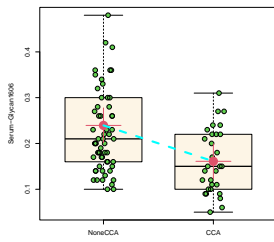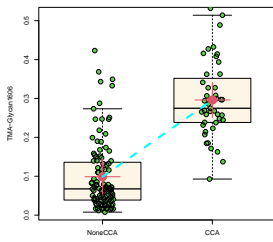
Glycan1611
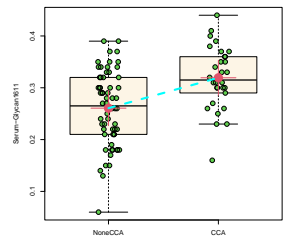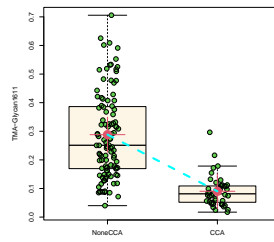
Glycan1622
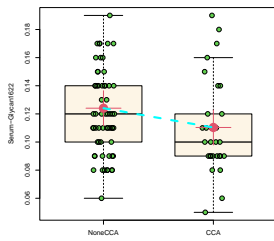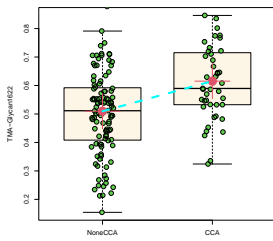
Glycan1647
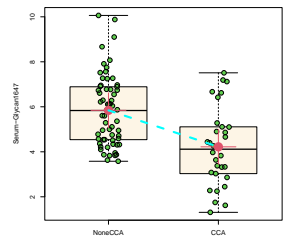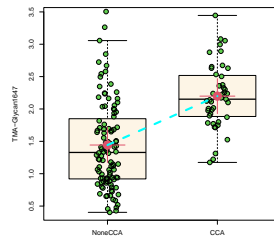
Glycan1663
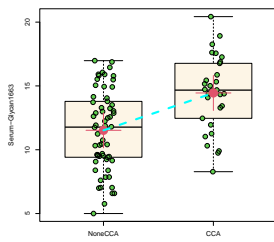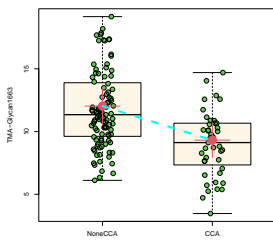
Glycan1688
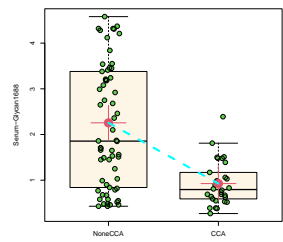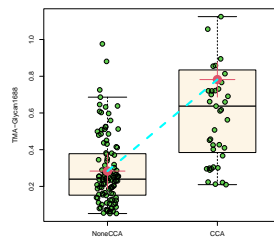
Glycan1704
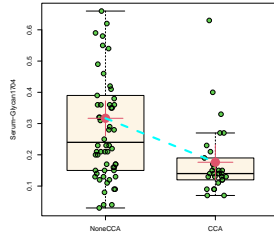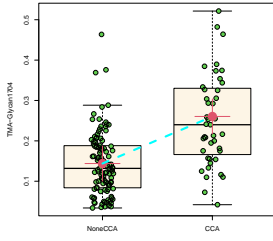
Glycan1743
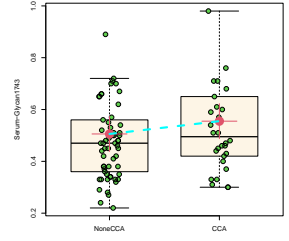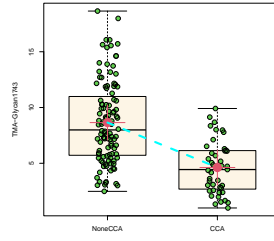
Glycan1809
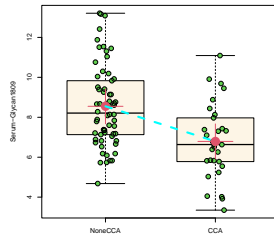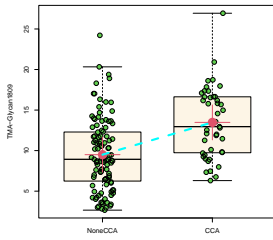
Glycan1825
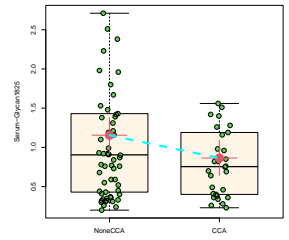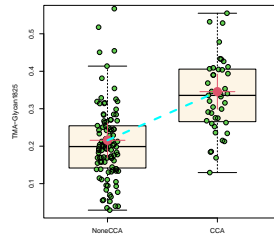
Glycan1850
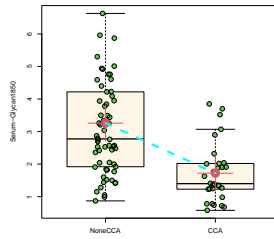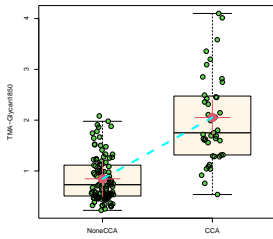
Glycan1866
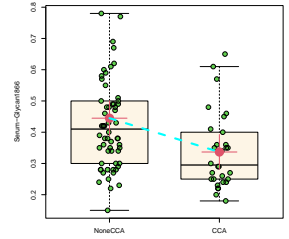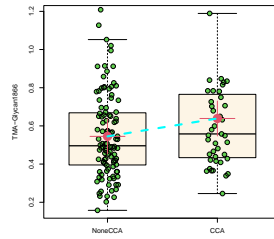
Glycan1905
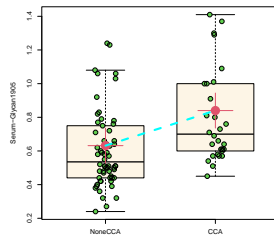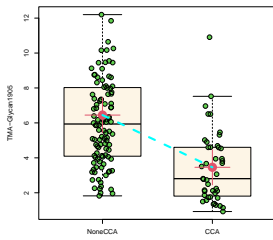
Glycan1954
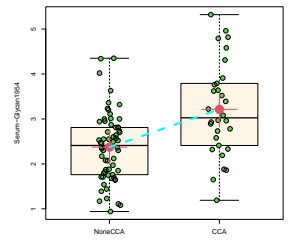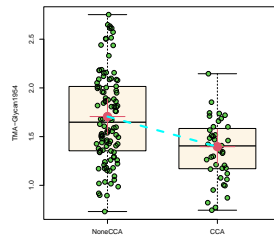
Glycan1976
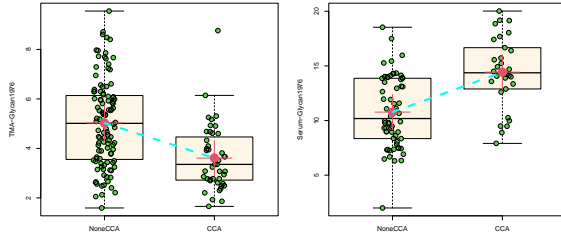
Glycan2012
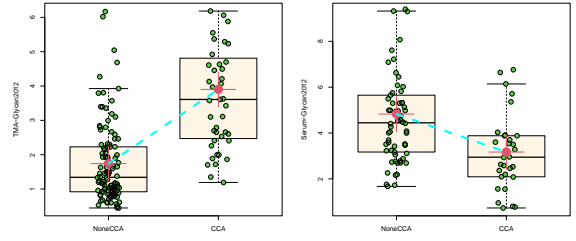
Glycan2028
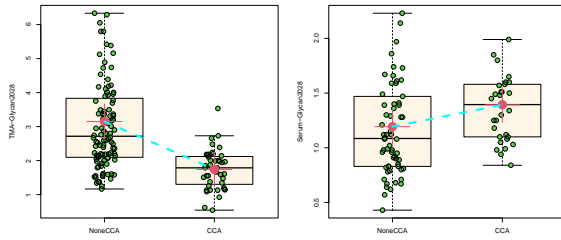
Glycan2100
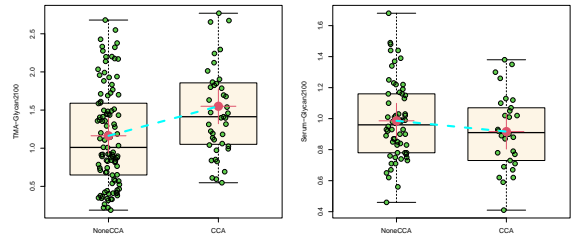
Glycan2122
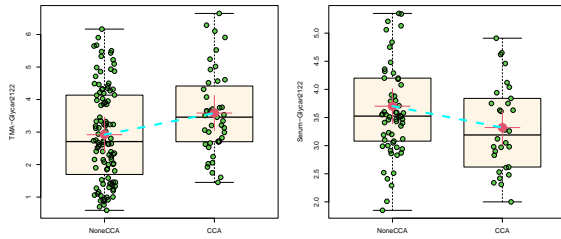
Glycan2158
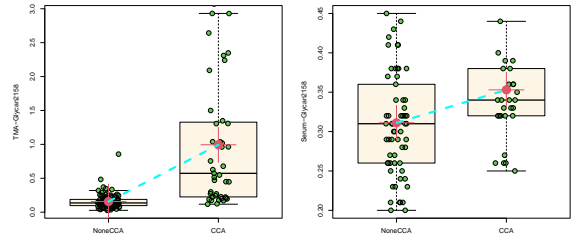
Glycan2163
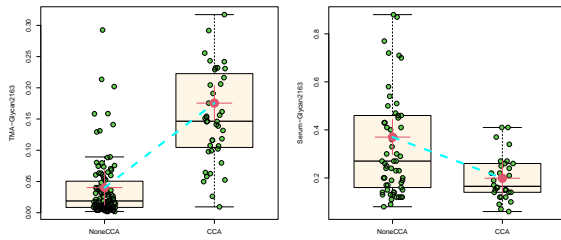
Glycan2174
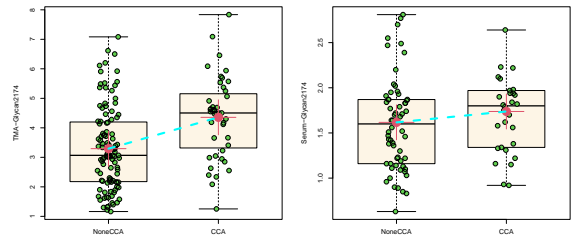
Glycan2303
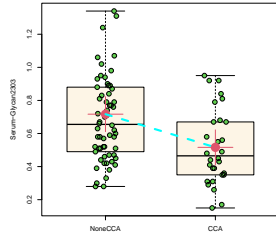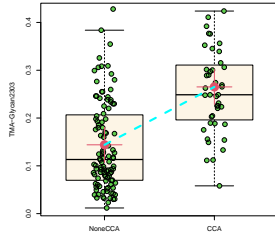
Glycan2319
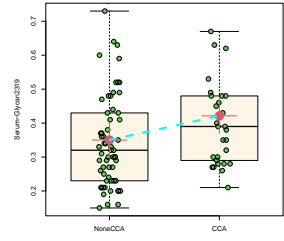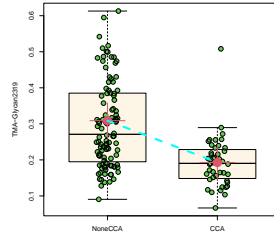
Glycan2320
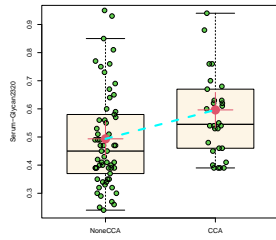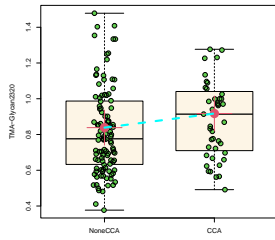
Glycan2325
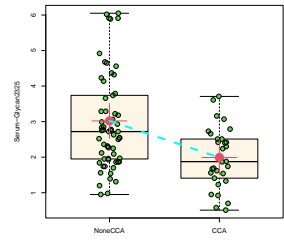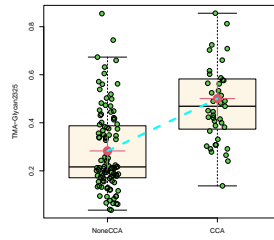
Glycan2341
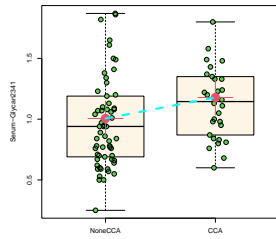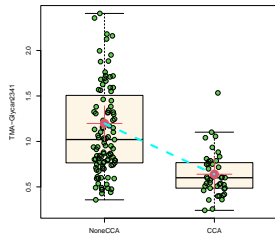
Glycan2361

Glycan2377

Glycan2393

161

Glycan2465   Glycan2487

Glycan2539   Glycan2654

Glycan2685   Glycan2706

Glycan2852
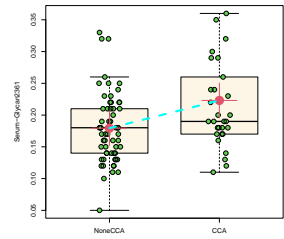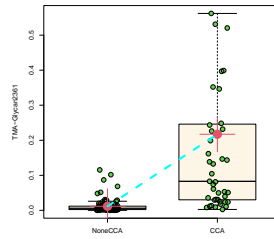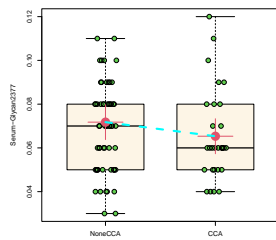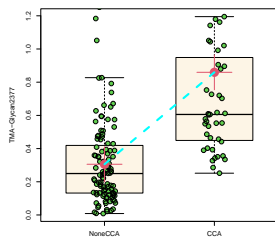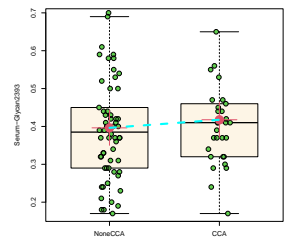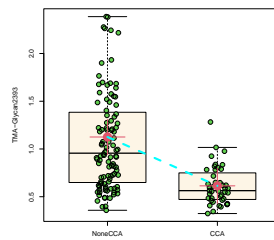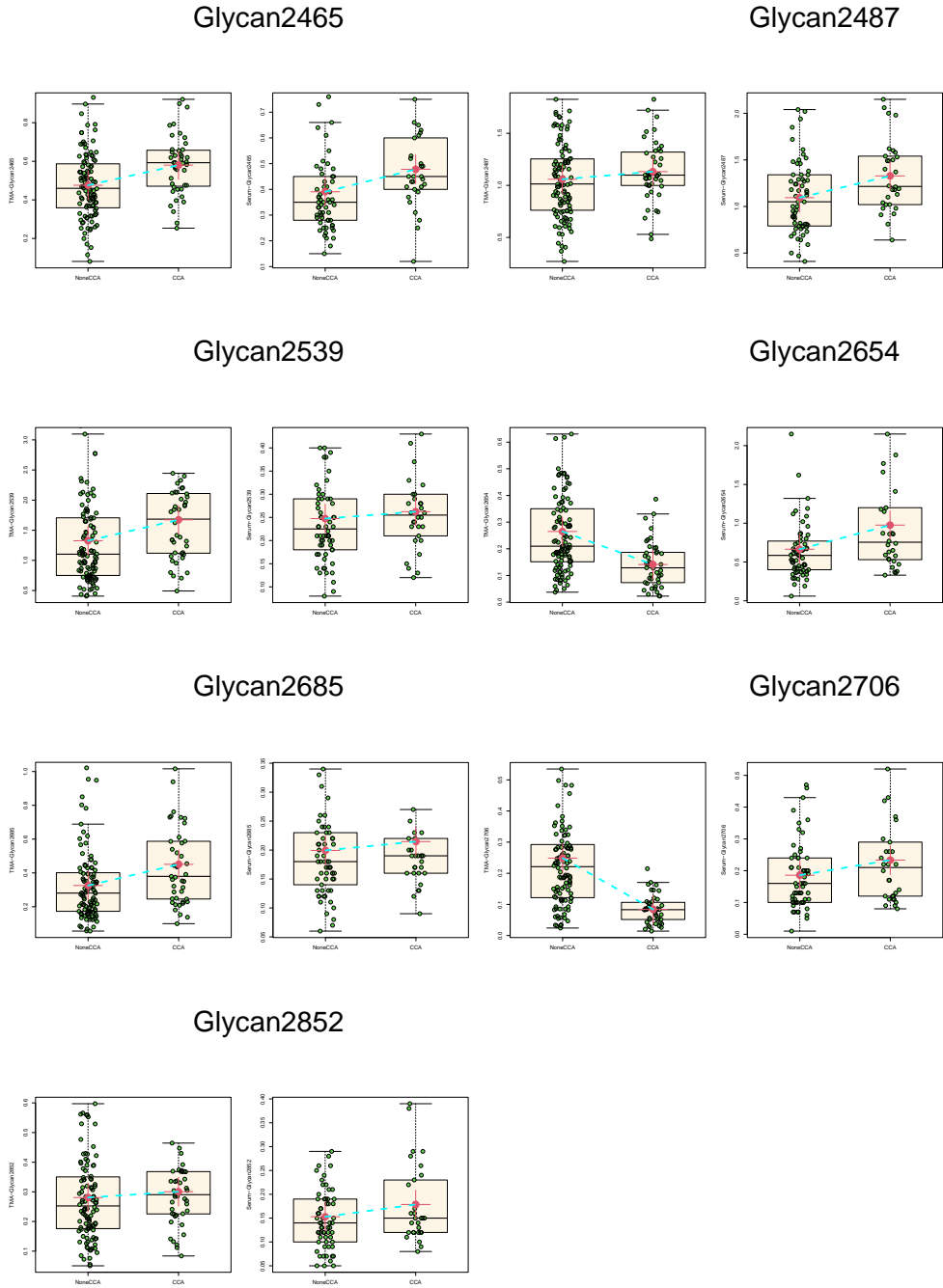
Figure 29. Relative intensity quantification of all N-glycans identified in serum (left row) and tissue (right row) analysis. Red font labeling for N-glycans follows the same trend between serum and tissue.

To identify specific N-glycans that could distinguish between patients with iCCA and non iCCA, we optimized the 12 N-glycans, this revealed three main N-glycans per dataset (Figure 30A). In the TMAs analysis, the N-glycans were at 1339 m/z, 1257 m/z (high mannose N-glycan), and 2158 m/z (Figure 30A). For the serum analysis, the N-glycans were a 1339 m/z, 2158 m/z, and 2361 m/z (highly branched tetraantennary, double fucosylated; Figure 30A). In addition, we investigated if the different types of iCCA (small duct and large duct) could express a different N-glycan profile, and found that there were no significant differences in any of these glycans, suggesting that these N-glycans are independent of the type of iCCA (Figure 30B). From these three main N-glycans, only the two common N-glycans (1339 m/z and 2158 m/z) between tissue and serum were further analyzed (Figure 30C). The ROC curve from common N-glycans combined in serum had an AUC of 0.7656 (Figure 31A); while these same common N-glycans in the TMAs had an AUC of 0.9317 (Figure 31B). Interestingly, N-glycan 1339 m/z and 2158 m/z had an opposing trend in serum (Figure 31C) and tissue (Figure 31D), where 1339 m/z was significantly decreased in iCCA while 2158 m/z was significantly increased. This analysis suggests that the use of bisecting fucosylated N-glycans and the core N-glycan structure have a powerful discrimination ability that could be a promising differentiator strategy between patients with iCCA and non iCCA.

163

A

		TMAs	Serum
Optimized N-glycans		1339 m/z 1257 m/z 2158 m/z	1339 m/z 2158 m/z 2361 m/z
Classification	**AUC**	0.9457	0.7796
	95% CI	0.9123-0.9792	0.6871-0.8720
LOOCV	**AUC**	0.9317	0.7344
	95% CI	0.8905-0.9730	0.6341-0.8347

B

C
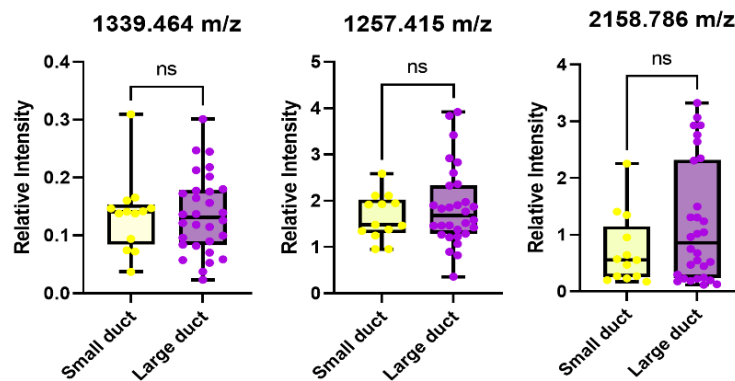
		TMAs	Serum
Common N-glycans		1339 m/z 2158 m/z	1339 m/z 2158 m/z
Classification	**AUC**	0.9317	0.7656
	95% CI	0.8903-0.9732	0.6673-0.8639
LOOCV	**AUC**	0.9203	0.7398
	95% CI	0.8734-0.9672	0.6373-0.8423

Figure 30. **Optimized and common N-glycans in serum and tissue were identified based on selected algorithm** (A) Three N-glycans in TMA and serum were identified after optimization. (B) Relative intensity quantification for both TMAs of the respective N-glycan based on small and large duct classification. (C) Two common N-glycans were identified between datasets. LOOCV (Leave-One-Out Cross-Validation). AUC (Area Under the Curve).
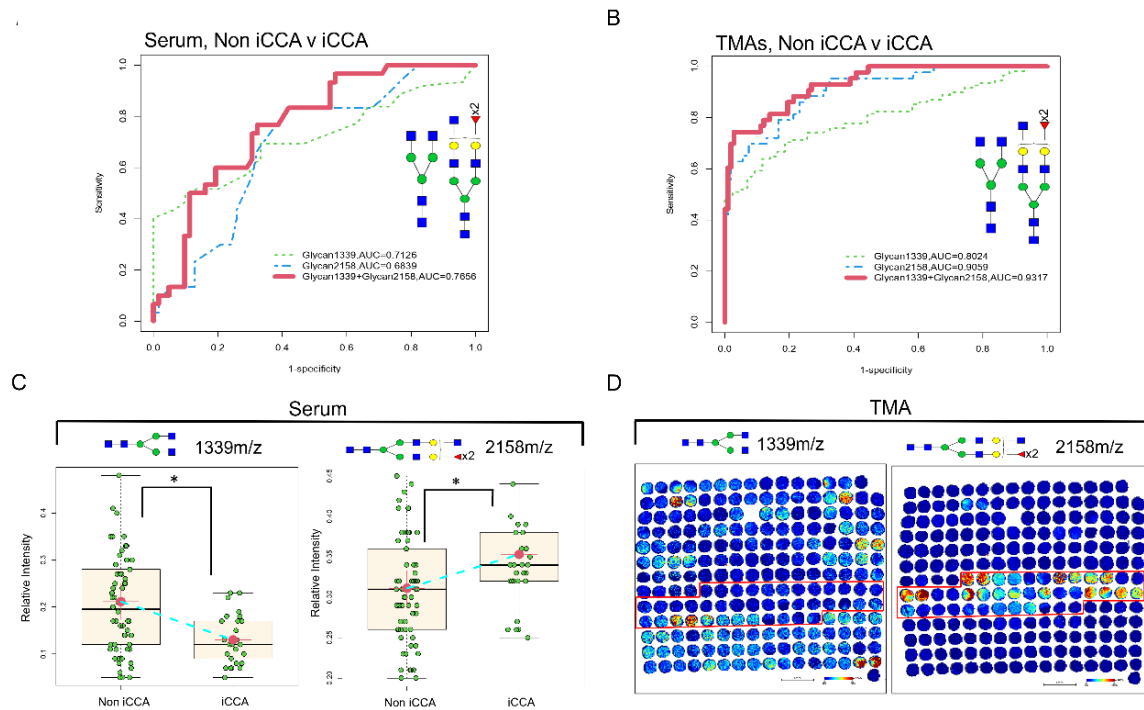
Figure 31. N-glycans from serum and TMA as promising biomarkers to differentiate iCCA from other liver diseases. (A) ROC curve for serum. NoniCCA n = 62 and iCCA n = 30. (B) ROC curve for TMA, non iCCA n = 108 and iCCA n = 43. For ROC curve labeling: N-glycan 1339 m/z (green-dashed line), N-glycan 2158 m/z (blue-dashed line), and a combination of N-glycans 1339 m/z and 2158 m/z (red-solid line). (C) Relative intensity quantification boxplots of N-glycan 1339 m/z and 2158 m/z in serum, non iCCA n = 62, and iCCA n = 30. (D) Representative images of TMA 1 showing the N-glycan intensity for 1339 m/z and 2158 m/z. The asterisk indicates statistical difference (Mann–Whitney, P < 0.001) and error bars represent the SD. For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose.

Next, we explored further the importance of 1339 m/z as one of the differentiators between iCCA and non iCCA in serum because this N-glycan modification has not been associated with liver cancer before to our knowledge. We excluded this N-glycan from the 12 N-glycans from our analysis and performed a feature selection from a random forest algorithm. We identified, 5 N-glycans: 2465 m/z and 2487 m/z (triantennary fucosylated with a sialic acid residue and same structure sulfated, respectively), 2158 m/z, 2320 m/z, and 2174 m/z. Figure 32A shows the importance of each N-glycan selected, with all N-glycans being bisected/triantennary fucosylated. Combining these N-glycans to determine their ability to distinguish between iCCA and non iCCA resulted in an AUC of 1 for classification and an AUC of 0.7151 for LOOCV (Figure 32B). This analysis demonstrates the importance of N-glycan 1339 m/z in our model because its exclusion resulted in the need for 5 N-glycans to compensate for the use as a powerful discriminator (Figure 32B). However, this still confirms our previous analysis in which fucosylated bisected and triantennary structures have important roles in iCCA.
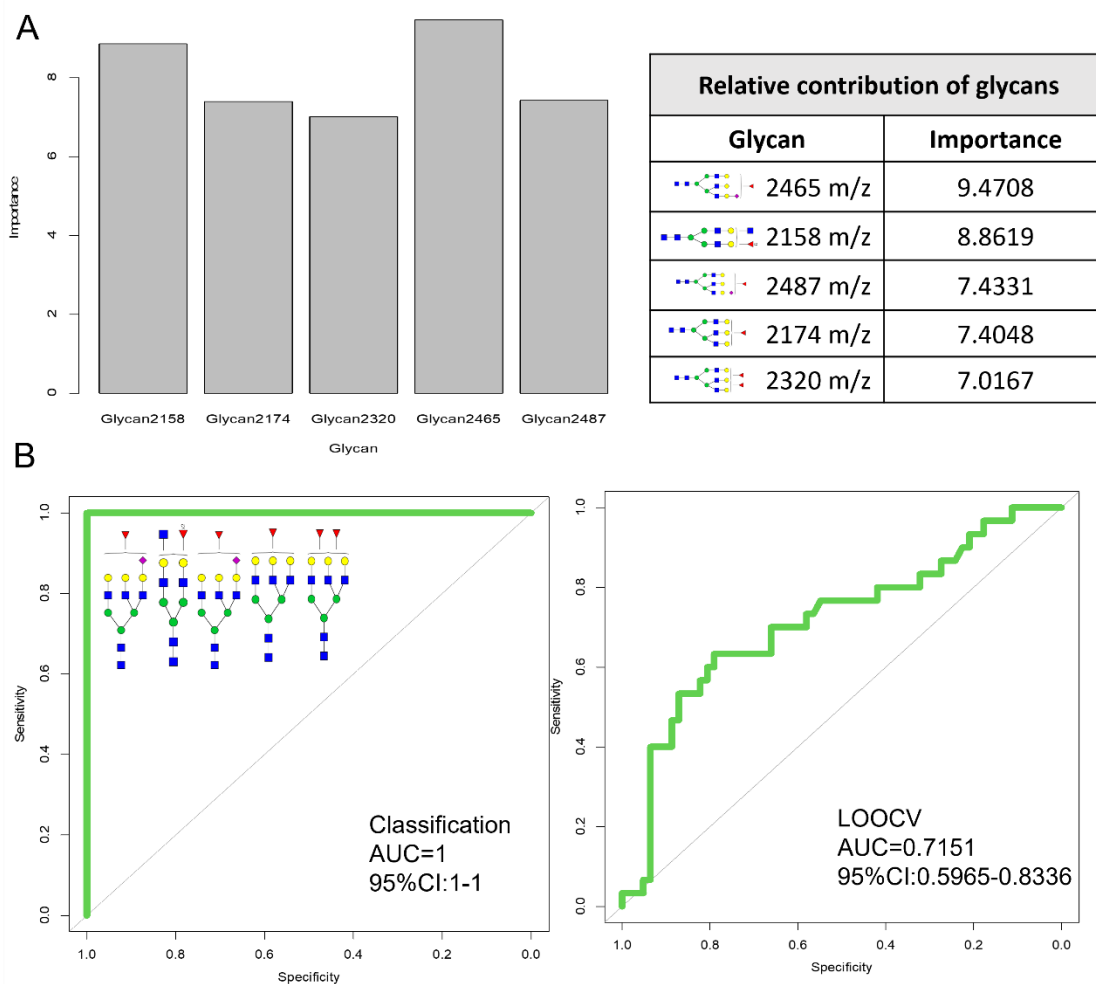
Figure 32. Importance and relative contribution of N-glycans based on selected algorithm. (A) Quantification of the relative contribution of N-glycans (left), table of N-glycans, and proposed structure with importance values (right) when removing N-glycan at 1339 m/z from the analysis. (B) ROC (Receiving Operator Characteristic) curve classification for serum (left) ROC curve LOOCV (right) of N-glycans in A.

Finally, to implement a clinical translational aspect for biomarker discovery using the N-glycan molecular changes presented here, we added CA19-9 information to our analysis. Serum CA19-9 is a biomarker used to identify those who might require diagnostic imaging for CCA. For this analysis, we focused on patients for whom we had CA19-9 information available (30 iCCA patients and 17 PSC patients). Strategies for the diagnosis of iCCA in PSC are urgently needed since early detection of iCCA can improve patients' survival and the conventional strategies share many features between conditions making diagnosis unsuccessful in most patients (51). Figure 33A lists the small capability of individual glycans (AUC: 0.500-0.727), CA19-9 (AUC: 0.512), age (AUC: 0.654), or liver enzymes (AUC:0.558-0.697) to differentiate between patients with iCCA or PSC. However, when we combined some of the most significant N-glycans (1339, 1257, and 2158), this resulted in a significant ability of this combination to differentiate between patients with PSC to those with iCCA (AUC: 0.8431, p=<0.00001) (Figure 32B and 33C). We demonstrate this is independent of CA19-9 since including CA19-9 information in the N-glycan combination, did not significantly improve biomarker performance (AUC: 0.8472) (Figure 33B, 33D) and Table 13. Finally, an additional multivariate model on N-glycans 1339, 1257, and 2158 and the clinical information available for PSC and iCCA patients revealed that the ability of these N-glycan to differentiate between diagnosis is independent of the clinical information applied (Figure 34 A-C). Overall, we demonstrate that the combination of specific N-glycan modifications can be a promising biomarker for identifying iCCA patients from those with PSC.
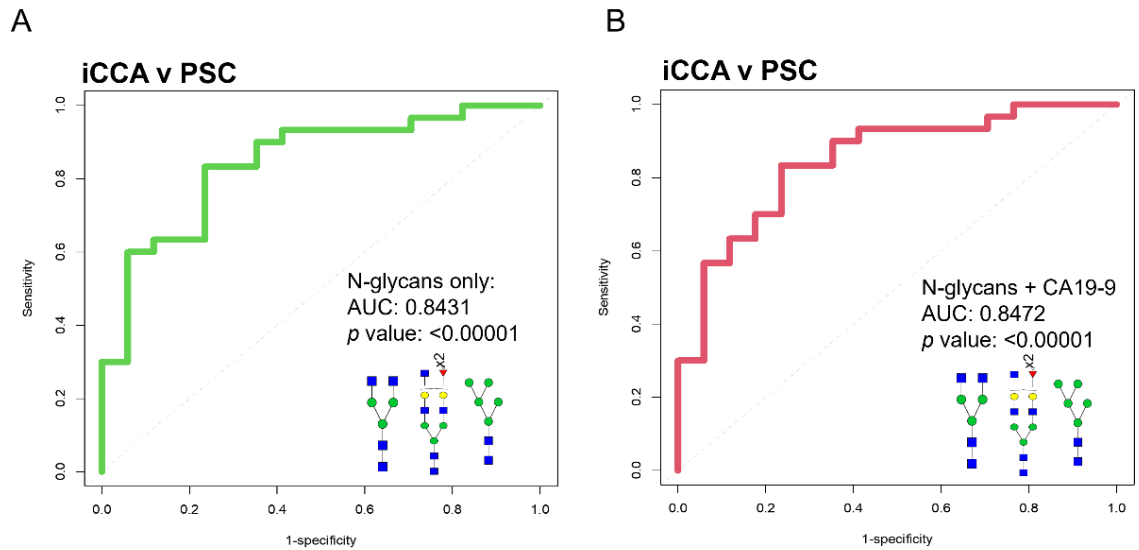
Figure 33. N-glycan combinations as promising biomarkers to differentiate iCCA from PSC. (A) ROC curve classification of a combination of 1339 m/z, 2158 m/z, and 1257 m/z. (B) ROC curve of combination of 1339 m/z, 2158 m/z, 1257 m/z, and CA19-9. For ROC curves: iCCA (n = 30) and PSC (n = 17). For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose

Table 13. List combinations of N-glycan as possible biomarkers to differentiate iCCA from PSC. For N-glycans, red triangle, fucose; blue square, N-acetylglucosamine; green circles, mannose; yellow circles, galactose

[1]Biomarker	[2]n	[3]iCCA:PSC	[4]AUC	[5]95%CI of AUC	[6]p value
Glycan1257+Glycan1339+Glycan2158	**47**	**30:17**	**0.8431**	**0.7257-0.9606**	**<0.00001**
Glycan1257+Glycan1339+Glycan2158+CA19-9	47	30:17	0.8472	0.7313-0.9629	<0.00001
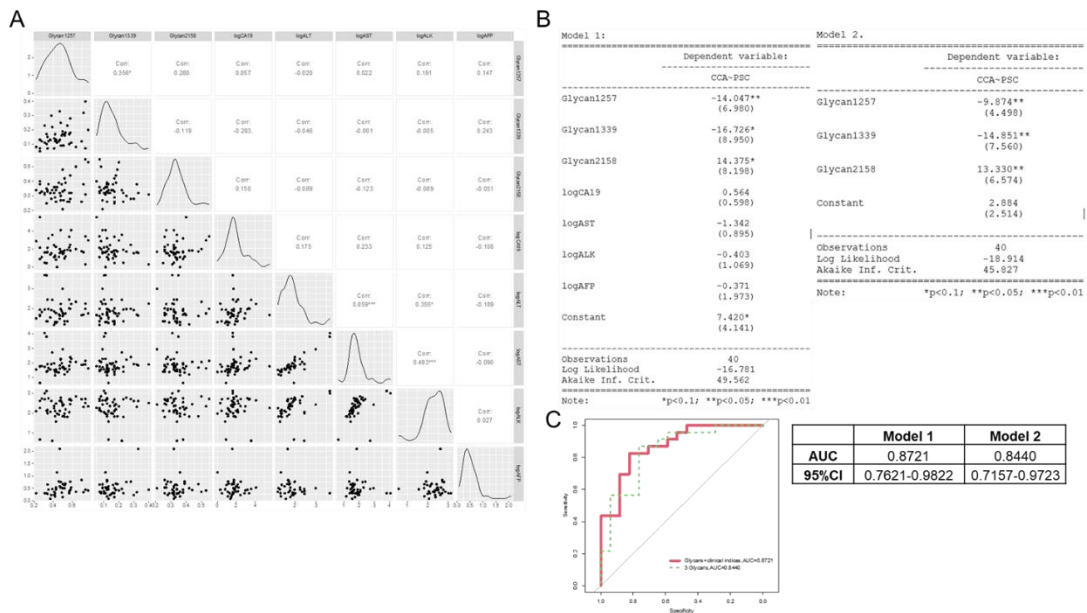
Figure 34. Multivariate model-Multiple logistic regression in serum samples. (A)
Scatterplots and correlations between N-glycans of interest (1339, 1257, 2158) and
clinical information available (ALT, AST, ALK, and AFP). (B) Multivariate model-Multiple
logistic regression in CCA and PSC serum samples (n=40). Model 1: three N-glycans of
interestto identify and clinical information available (left panel). Model 2: Only the three
N-glycans of interest (right panel). C. ROC curve of the combination of glycans and
clinical information (red-solid line) and only glycans (green-dashed line) (left),
classification performance table of the two models in B. p=0.5731, Delong's test between
model 1 and 2. CA19-9, ALT, AST, ALK, and AFP values were log-transformed for
plotting and modeling convenience. logALT was removed from the multiple logistic
regression analysis due to a high value of Variance inflation factor (VIF).

6.4 Discussion

Alterations in N-glycosylation have been long observed with HCC (78,152,154,252,253), and we and others have shown that increased levels of some fucosylated glycoproteins could be observed in the serum of patients with CCA (79,80,91,240). However, a study that elucidates the origin of these N-glycan modifications while also exploring the same modifications in iCCA serum as a possible biomarker has not been done before.

In this study, we identified bisecting, branching (tri-antennary), and fucosylation as specific N-glycan structure modifications in iCCA tissue and serum. These N-glycan modifications are known to be catalyzed by the following glycosyltransferases: First, β-1,4-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase (MGAT3) for bisecting, which has been previously reported to differentiate between iCCA and HCC in serum (254). In addition, MGAT3 has been considered a malignancy suppressor where its overexpression can inhibit metastatic profiles of cancer cells (61). Second, α-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A (MGAT5) for branching, which has been linked to malignancy and correlates with disease progression (114). In addition, its activation in many cancers has been reported to be through the upregulation of the RAS-RAF-MAPK signaling pathway (61). Finally, another major alteration observed in the CCA patients was increased fucosylation. We and others have shown that core fucosylation, catalyzed by alpha-1,6-fucosyltransferase (FUT8), is one of the main N-glycan modifications in early liver disease like non-alcoholic fatty liver disease (NASH) and liver cancers (77,82,91,159). FUT8 expression has also been shown to increase as cells undergo an epithelial-mesenchymal transition (EMT) by remodeling core fucosylation on the TGF-B receptor (255). Previous studies in CCA serum have reported increased

172

double fucosylation with tri-antennary N-glycans (79,80,91). Similarly, we demonstrate here that alterations in fucosylation are due to outer-arm and core fucosylation, but core fucosylation had a higher specificity to iCCA in tissue (Figure 25B-D).

This study elucidates N-glycan modifications that can be specific to each type of liver cancer in tissue. iCCA patients had a very specific expression of 2158 m/z (presumed bisected double fucosylated N-glycan), which was not present in HCC or normal samples (Figure 23C, 23H, and 23M). In contrast, only HCC patients expressed 2393 m/z (tetra-antennary branched N-glycan), which was not present in iCCA and/or normal samples (Figure 23D, 23I, and 23N). This suggests N-glycosyltransferases MGAT 3 and MGAT 5 have different roles in cancer progression according to the type of liver cancer. The changes in specific N-glycan structures according to the type of cancer observed here can be explained by the opposing roles of MGAT3 and MGAT5. MGAT3-generated bisecting structures are less readily bound and modified by MGAT5, resulting in a decrease in tri- and tetra antennary structures. Accordingly, increased intensity of bisecting structures correlated with decreased intensity of tetra antennary structures in iCCA patients, but HCC patients had no changes in bisecting structures and increased highly branched structures. Future studies should investigate the exact roles of MGAT3 and MGAT5 in the context of each type of liver cancer.

An interesting N-glycan modification that has not previously been linked to iCCA was 1339 m/z (bi-antennary non-fucosylated N-glycan) which in this study had a decreased intensity in iCCA tissue and serum. Statistical analysis revealed the importance of this modification when coupled with bisected fucosylated structures as a biomarker candidate for differentiating between iCCA v non iCCA and iCCA v PSC. We hypothesize

that the reduction in the expression of this N-glycan is due to this structure being modified to higher/complex N-glycan structures since this is the basic core of an N-glycan structure. A translational point of this study was correlating tissue and serum analysis for the characterization of N-glycan-related molecular changes since most of the serum N-glycoproteins are synthesized by the hepatobiliary system and can reflect liver health. Utilizing N-glycan alterations that were correlated in tissue and serum analysis improved biomarker detection and should be pursued for the investigation of other diseases, especially those impacting the liver. It is important to note that not all N-glycans followed the same trend from tissue to serum. We hypothesize these inconsistencies may be due to the large contribution of immunoglobulin G to the serum N-glycan profile.

Here, we utilized MALDI-IMS N-glycan imaging to identify the modifications that occur directly in iCCA patient tissue and serum samples. In our whole tissue analysis, we elucidate the histopathological origin of N-glycan structures in the iCCA and HCC tissue. Next, we analyzed two tissue TMA cohorts (a discovery and validation cohort) and an independent serum cohort. Finally, we used tissue (from discovery and validation cohorts) and serum N-glycan alterations and identified 2158 m/z (presumed bisecting double fucosylated) and 1339 m/z (bi-antennary non-fucosylated) as the most effective combination to distinguish between iCCA and non iCCA in tissue and serum. While 1339 m/z, 2158 m/z, and 1257 m/z (high mannose N-glycan) as biomarker candidates to distinguish between iCCA patients and PSC patients in serum.

Together, the data presented here suggest that the changes characterized here in tissue and serum could originate from cancer itself and prior analytical glycan tools were not able to detect these changes within the tissue. A matched serum and tissue cohort would be needed to confirm the origin of these N-glycan alterations. To conclude, we

propose the use of N-glycan alterations as promising biomarkers for their ability to differentiate between intrahepatic iCCA, HCC, PSC, and benign chronic liver disease. Further studies should be done using a larger set of serum samples to validate the value of the biomarker candidates proposed here.

Chapter 7: Generation of a hepatocyte-specific *Fut8* knockout mouse model

7.1 Fucosyltransferase 8 (*Fut8*) and core fucosylation

The study of N-glycan structure alterations can be challenging due to the complexity and diverse modifications in different cellular processes, some of which can be disease-specific. Elucidating the role of glycosyltransferases (enzymes that add a specific residue to build an N-glycan structure) is necessary to gain a better understanding of their roles in diseases. The major modification in eukaryotic systems is core fucosylation, catalyzed by glycosyltransferase Fucosyltransferase 8 (FUT8). FUT8 enables the activity of α1-6 fucosyltransferase which catalyzes the transfer of an L-fucose residue from GDP-Fuc to the inner GlcNAc residue of the core of the N-glycan, this is known as core fucosylation (Figure 35) (256). Mouse and human FUT8 is a type II membrane protein on the Golgi with 575 amino acids with no consensus for N-glycosylation. The human FUT8 is localized on chromosome 14 while the mouse FUT8 is localized on chromosome 12. The 3D structure of human FUT8 consists of three domains (1) N-terminal coiled-coil, (2) catalytic (3) C-terminal SH3 domain (256).

Many different reports have elucidated the numerous biological functions FUT8 has during development, normal cellular processes, cancer-related cellular processes, regulation and function of glycoproteins, immune response, and its value as a cancer biomarker (alpha-fetoprotein (AFP-L3)) for Hepatocellular Carcinoma (HCC) (61,93,139,159,160,229,257,258). Our group has demonstrated by spatial analysis that core fucosylation is one of the major modifications that occur directly in HCC, CCA, and NASH tissue (76,77). However, while alterations in core fucosylation and FUT8 have been characterized in cells, serum, and tissue (mouse and human), the role, and mechanism of FUT8/core fucosylation in the liver are still unknown. One of the major

limitations to elucidating the impact of fucosylation in liver disease is the lack of liver-

specific relevant animal models.
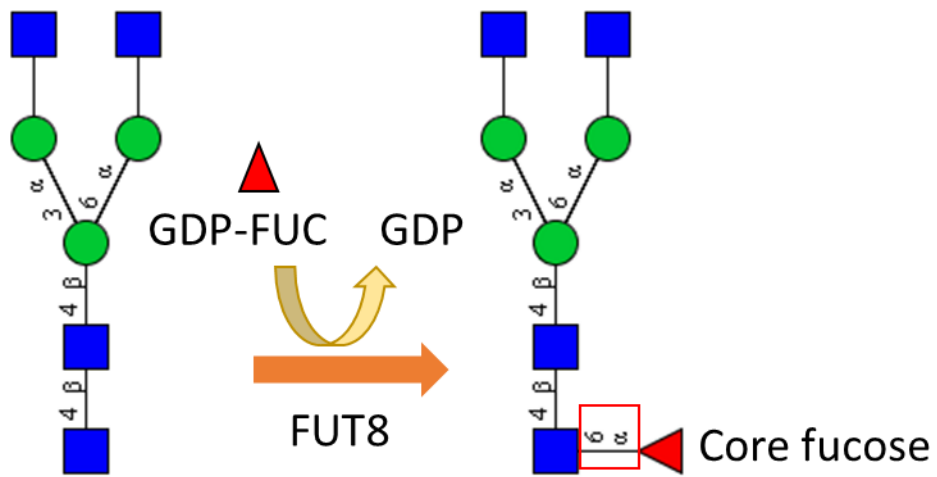
Figure 35. Fucosyltransferase 8 (FUT8) linkage. FUT8 catalyzes the addition of a fucose residue (red triangle) from donor substrate GDP-FUC to the α1,6 linkage on the core of the glycan, resulting in a core fucosylated N-glycan. Graphic adapted from (256)

7.2 Current in vivo *Fut8* model

The current in vivo model for Fut8 is a constitutive *Fut8* whole-body knockout

mouse model (94,160). One of the first reports in this model was in the context of liver

disease, *Fut8* knockout (*Fut8* KO) mice were treated with hepatotoxins

Diethylnitrosamine (DEN) and Pentobarbital (PB) to induce HCC tumors (94). *Fut8* KO

mice under carcinogen treatments demonstrated a reduction in liver tumor incidence,

tumor formation, and decreased cell proliferation. The proposed mechanism for this was

elucidated in HepG2 cells where the deletion of *Fut8* affects epidermal growth factor

(EGF) and hepatocyte growth factor (HGF) receptors, interferes with receptor and ligand

interaction, and downregulates their respective downstream signaling pathway. From

this study, *Fut8* was suggested to have roles in oncogenesis, liver disease progression,

and poor survival. Similarly, it was proposed as a prognostic marker and therapeutic

target for HCC by the same group (94). No other in vivo studies with this mouse model

or other *Fut8* deleted mouse models in the context of liver disease have been reported

to our knowledge.

7.2.1 Limitations of *Fut8* whole-body deletion

While the current mouse model has provided valuable information in terms of

possible mechanisms of Fut8 in liver cancer, this same model also suffered from

significant phenotypic alterations including post-natal lethality by day 3, growth

retardation in survivors, schizophrenia-like behaviors, and emphysema (139,160). An

important reason for these off-target effects is that *Fut8* has numerous roles in multiple

cellular processes and is crucial for mouse development. Similarly, the brain, placenta,

lung, stomach, and small intestine tissues have a high *Fut8* expression, suggesting its

importance in the respective tissue (259). Due to the adverse effects of *Fut8* whole-body

deletion, this model lacks to clarify liver-specific (let alone hepatocyte) impacts of fucosylation. This supports the need for tissue-specific mouse models to avoid off-target effects. Here, we address these limitations by the generation of a hepatocyte-specific *Fut8* knockout mouse model.

7.3 Hepatocyte-specific *Fut8* knockout model

A genetically engineered mouse model (GEMM) is an animal modeling tool generated by introducing genetic mutations that are normally associated with the disease of interest. GEMMs have demonstrated great value in many fields, these models can provide information in terms of specific mechanisms, and elucidation of specific genes' roles in disease establishment, maintenance, and progression (146). The mouse model can be designed with a gain-of-function or loss-of-function strategy for the gene of interest. Similarly, GEMMs can either have a constitutive, conditional, or inducible gene expression. Since the use of homozygous null mutation of the *Fut8* model (constitutive system) affects the tissue specificity by inducing abnormalities in other tissues and possibly activating compensatory molecular pathways, we exploited the use of a conditional gene expression system that allows for the spatial control of the gene activity by the use of a tissue-specific promoter. A liver-specific *Fut8* deleted mouse model allows us to elucidate liver-specific roles in different liver processes and diseases.

7.3.1 Mouse model design overview

The targeted gene is Fucosyltransferase 8 (Fut8) (*Mus musculus* (house mouse) in a C57BL/6J strain by a CRISPR/Cas- mediated genome engineering. *The Fut8* gene (Gene ID: 53618) is located on mouse chromosome 12 (Figure 36A and Table 14). Exon 9 was selected as the conditional knockout region by two different single guide RNAs (sgRNA) (Figure 36B and 36C), deletion of this region results in a frameshift and the loss

of function of the mouse *Fut8* gene. A LoxP strategy was used to target exon 9 where

sgRNAs were present (Figure 36B and 36C). For 5' LoxP site insertion, the size of intron

8 is 18874 bp and for 3' LoxP site insertion, the size of intron 9 is 35523 bp. The size of

the conditional knockout region is approximately 658 bp with no other known gene

present in this region.

Table 14. Fut8 flox/flox mouse model design details

Model design details	
Gene and ID	fucosyltransferase 8 (Fut8)/ Gene ID: 53618
Chromosome	12
Transcript	Fut8=203
Number of exons in transcript	13 exons in the longest coding protein ATG start codon in exon 5 and the TGA stop codon in exon 13
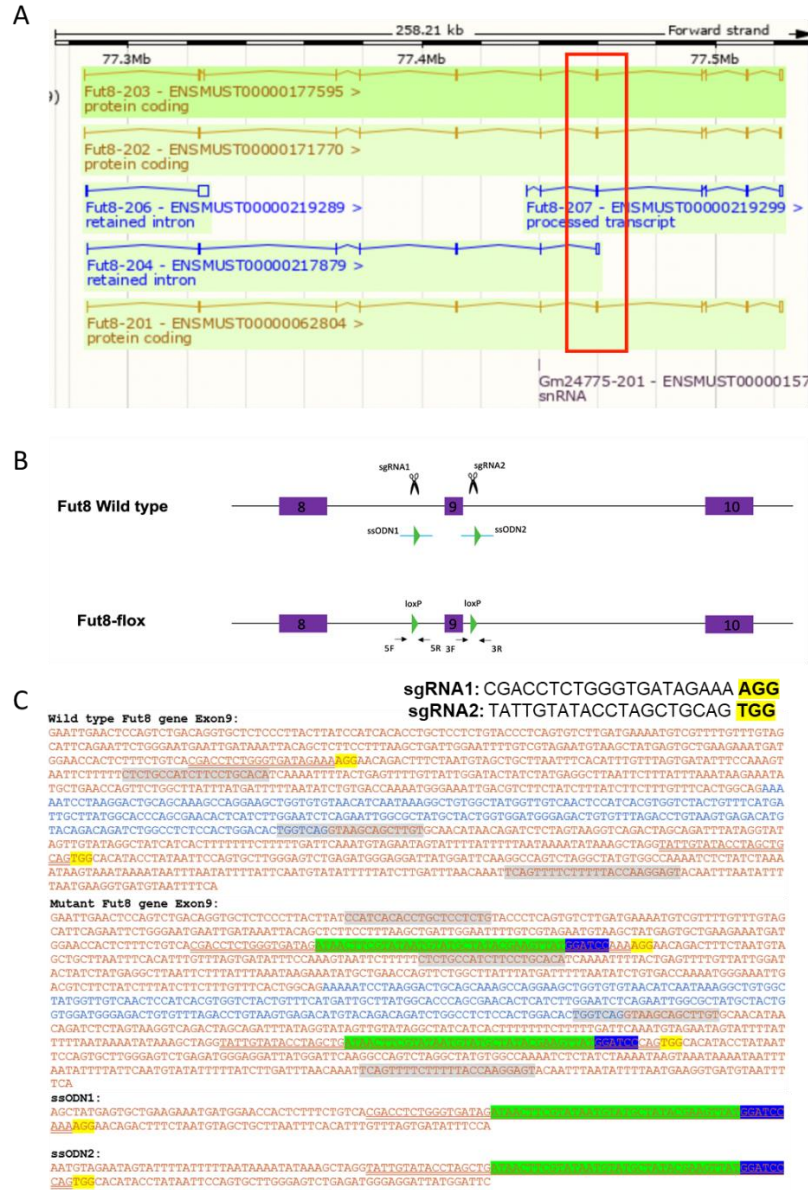
Figure 36. Strategy and design for the generation of a Fut8 Cre-LoxP model. (A) Transcript summary of Fut8, Fut8-203 is the longest protein-coding transcript and was selected for the design. (**B**) Outline of Fut8 wild type demonstrating the positions of sgRNA1 and sgRNA2 around targeted exon 9, where the two LoxP sites were placed (Fut8-flox). (**C**) Sequence of fut8 wild type (top) and mutant (bottom) for exon 9, highlighting sequences targeted.

7.3.1.1 Setback of mouse model generation

After the first round of injection following the model in Figure 36, founder mice were identified with only a 3' LoxP insertion and lacking a 5' LoxP insertion. To circumvent this, an 11-base pair (bp) deletion was identified surrounding the 5' sgRNA binding site. A re-targeting strategy was designed based on the 11-bp sequence to insert the 5' LoxP into the 11-bp deletion region (Figures 37A and 37B).

Figure 37. Re-targeting design for the generation of a Fut8 Cre-LoxP model. (**A**) Outline of FUT8 3' LoxP for insertion of 5' LoxP site at the 11bp deletion region (top), resulting in the placement of both LoxP sites around exon 9. (**B**) Sequences of fut8 wild type (top) and mutant (bottom) for exon 9, highlighting sequences targeted.

7.3.2 Fut8 flox/flox model

After completing the retargeting strategy, founder (F0) mice with genotype Flox/WT were identified with both LoxP sites. Intending to expand the Flox/WT line, Flox/WT and WT/WT breeding resulted in F1 mice with Flox/WT, also with both LoxP sites present, confirming a germline transmission of both LoxP sites (Figure 38A). DNA Sequencing was done to confirm the correct insertion of both 5' and 3' LoxP sites (Figure 38B). Finally, F1 pups were interbred to generate F2 mice with the desired genotype Flox/Flox (Figure 38C) and confirmed by DNA sequencing (Figure 38D).
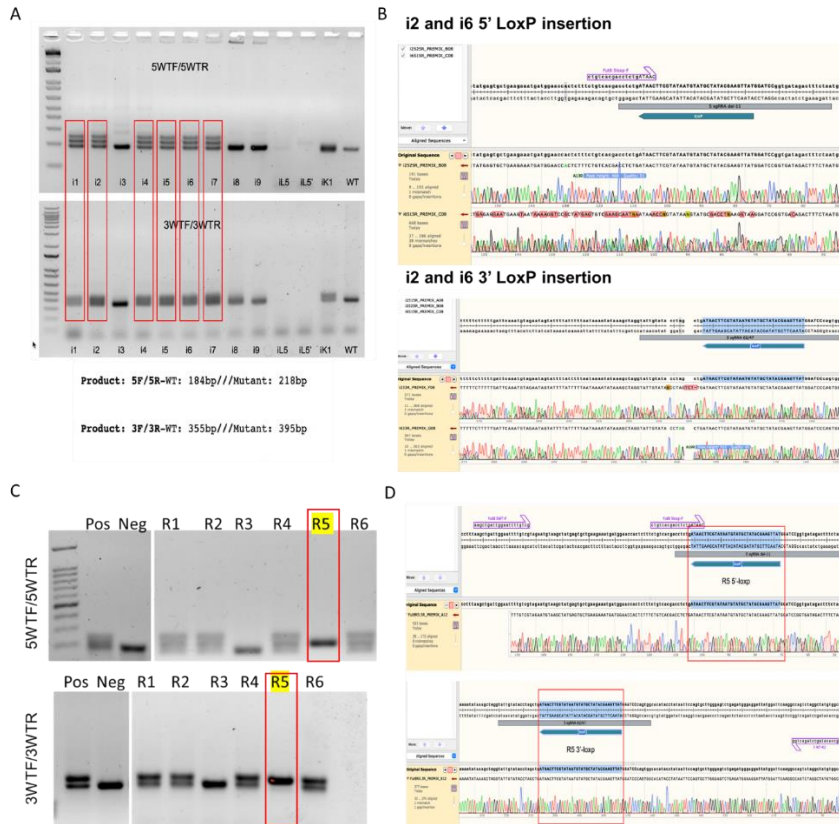
Figure 38. Confirmation of Fu8 flox/flox model genotype. (A) PCR gel from litter "i1-i9" where red boxes indicate pups with 5'LoxP (top) and 3' LoxP (bottom) insertion, resulting in genotype Fut8 flox/WT. (B) DNA sequencing snapshots of pups i2 and i6. (C) PCR gel from litter R1-R6, the red box indicates R5 pup with both LoxP sites, resulting in genotype Fut8 flox/flox. (D) DNA sequencing snapshots of pup R5.

7.3.3 Albumin-Cre transgenic mouse

Albumin (Alb) is the most abundant circulating protein in plasma, in healthy patients, it represents half of the total protein content in plasma. Alb is synthesized in hepatocytes and secreted into the bloodstream, this is considered a fast process where very little albumin is left and stored in the liver. Serum Alb works as a modulator of colloid osmotic pressure in plasma and a transporter of endogenous and exogenous ligands, including hormones, vitamins, and drugs throughout the body (260).

The endogenous *alb* gene is expressed exclusively in hepatocytes and is induced at differentiation in mice (261). *Alb*-Cre transgene has been used for decades for the generation of a transgenic mouse model in a Cre/lox system, that is driven by the serum albumin (alb) gene promoter (261). A Cre/Lox system offers a great advantage to control gene expression profiles: from the location (tissue-specific) to the timing (inducible system) of gene expression. The protein Cre recombinase recognizes the LoxP sites and according to the location and orientations, Cre expression will result in an inversion, deletion, or translocation of the gene of interest. An Alb-Cre model is a transgenic line that expresses Cre recombinase under the control of the mouse alb enhancer/promoter. This model is one of the most well-established models for Cre-dependent excision of LoxP ("floxed") sequences in adult hepatocytes for over a decade (261,262).

While alb mRNA levels are first detected at embryonic day 10.5, it is known that gene deletion by alb-Cre is progressive with the mouse age, at birth the efficiency recombination is 40%, at P30 is 60%, at 3 weeks of age (weaning) 75%, and is complete by 6 weeks of age (263). It is hypothesized that the lack of complete recombination at birth is due to a low expression of Cre since endogenous albumin is expressed early in

mouse development (stage 7- 8 somites). A separate study interested in the reliability of alb Cre in fetal and juvenile mice analyzed Cre activity (in situ patterns) during this time and reported that albCre transgene is functionally expressed close to the activation of endogenous alb gene (at differentiation), suggesting that hepatocytes at any stage of development expressed Cre and recombined the albumin reporter (261). Overall, this model has been accepted as a well-suited model for any stage: late fetal, neonatal, and adult stages.

7.3.4 Albumin-Cre *Fut8* mouse model

Based on the design previously presented (Figure 36) where the LoxP sites are facing the same direction, the crossing of a FUT8 flox/flox mouse with an Alb-cre mouse results in a deletion of the sequence between the Loxp sites being excised by Cre as a circular piece of DNA. The generation of this model was achieved following the breeding plan described in Figure 39A. Fut8 flox/flox and Alb-Cre genotypes were confirmed by PCR (Figure 39B). This breeding resulted in the generation of a FUT8fl/fl; Alb-Cre mouse (*Fut8 Alb-Cre knockout).* Relative to littermates there are not any phenotypic or behavioral differences to report up to 2 months of age (Figure 39C).

Overall, the generation of the desired mouse model, hepatocyte-specific Fut8 deleted (Fut8 Alb-Cre knockout) was successful. This newly generated transgenic mouse model will address current limitations with other models and allow for a complete characterization specifically in the liver.

190
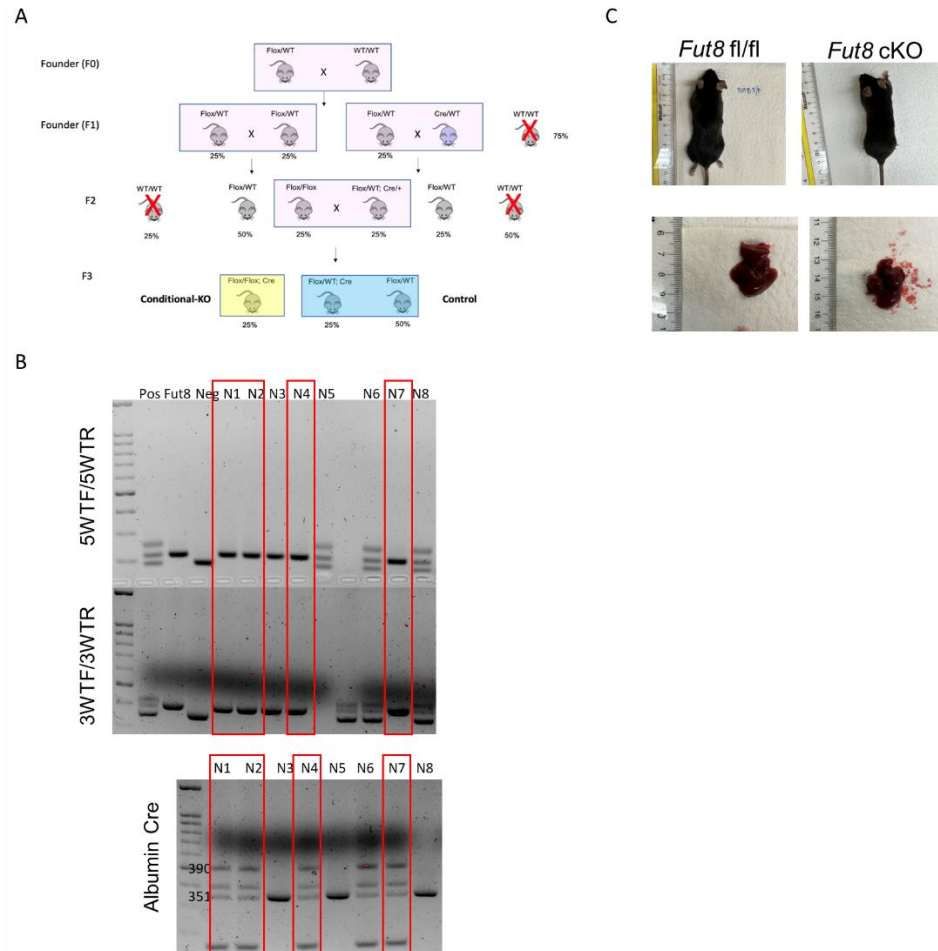
Figure 39. Generation of a hepatocyte-specific Fut8 knockout mouse model. (A)
Overview of the breeding plan from Founder 0 (F0) to Founder 3 (F3) for the generation
of conditional knockout genotype (yellow box). **(B)** PCR gel for confirmation of Fu8
flox/flox (top and middle panels) and Albumin Cre (bottom panel) model genotype. Red
boxes for mice (N1, N2, N4, and N7) with Fut8 flox/flox; Alb-Cre genotype. **(C)**
Representative images of F3 litter at weaning age (from the breeding of F2 mice)
including genotypes flox/flox; Alb-Cre, Flox/WT; Cre, and Flox/WT.

Chapter 8: Conclusions, Limitations, and Future Directions

8.1 Overall findings

Modifications in N-linked glycosylation have been established to be a sign of liver malignancy, however, the timing of when these modifications initiate, the identification of specific N-glycan signatures in different liver diseases, and the role of the N-glycan-related genes responsible for these modifications were all research gaps in the literature. The work presented in this dissertation addresses these gaps and identifies the area to advance in the discovery of N-glycan-related targets for diagnostic and therapeutic strategies. A crucial part of the findings reported here is being able to utilize mouse models for the induction of liver diseases that span the different stages of liver disease observed in humans. Chapter 3 provides an overview of a wide variety of mouse models available to study liver diseases with emphasis on their importance and relevance for human liver disease. Aim 1 addressed the need for understanding if and how N-glycosylation is altered at an early stage of liver disease in mice and humans while also providing information on the histological origin of the alterations identified (Chapter 4), Aim 2 takes this analysis further and provides a temporal understanding of N-glycan modifications in progressive liver diseases, from early liver disease to the early neoplastic condition (Chapter 5 and 6). Finally, Aim 3 addressed the lack of in vivo models for liver-specific N-glycan investigation and generates a mouse model based on the modifications observed in Aims 1 and 2 (Chapter 7). This chapter will focus on the conclusion, limitations, and future directions of each of the Aims presented above.

8.2 Non-alcoholic steatohepatitis (NASH) N-glycan profiling in mouse and humans tissues

8.2.1 Conclusions

Aim 1 focused on the N-linked glycan characterization of NASH in mouse and human tissue samples. Previous studies related to N-glycan modifications in NAFL/NASH had focused on the use of human serum samples exclusively (89,90,97,204,264). Here, we proposed Aim 1 intending to answer the following questions: (1) can a diet-inducible mouse model recapitulate the characteristics of NASH and also be useful for N-glycan-related studies; (2) Can N-glycome alterations in tissue be driven mainly by the consumption of high caloric diets, like those used to induce NASH disease; and (3) Are the N-glycan modifications previously reported in human NASH serum also observed in NASH liver tissue?. These questions were addressed by using diet-inducible mouse models for NAFL (use of a high-fat diet), and NASH (use of a western diet) and correlating the findings to NASH human liver biopsies. Mouse and human tissues were used for N-glycan profiling by MALDI-IMS. As outlined in Chapter 4, we were able to answer the proposed questions and confirm that NASH mouse models are reliable tools for N-glycan studies, since they had a similar N-glycan profile to what was observed in humans, we prove that alterations in the N-glycome can be driven by diet since mice in an HFD had a different N-glycan profile from those in a western diet. A major N-glycan modification that has been previously reported as increased in NASH serum is fucosylation, here we confirmed that this modification is also observed in liver tissue and identified the type of fucosylation. In addition, we found a correlation between core fucosylation and the level of fibrosis in NASH patients. These findings suggest that the alterations in fucosylation previously reported in serum could be originating from the

194

tissue and should be considered relevant for biomarker use. Additionally, we found that

high mannose N-glycans also had a strong correlation with steatotic areas in liver tissue.

Overall, the identification of these modifications suggests their promising use for

biomarker strategies, and the possibility of using them to diagnose specific stages of the

disease progression like NAFL (high levels of steatosis), and NASH (high levels of

fibrosis). The completion of this aim provides novel information for biomarker-related

research in specific N-glycan modifications that could be used for screening a specific

stage of disease and confirming that modifications observed in serum are in a way

consistent with what is observed in tissue. This was the first study to characterize a full

N-glycan profile in whole tissue based on histopathological characteristics of the disease

in mouse and human NAFL/NASH models/patients.

8.2.2 Limitations

A major limitation of this study was the challenge of inducing a complete

spectrum of NAFLD. As previously mentioned, one of the main characteristics of this

disease is the high level of fibrosis. Induction of fibrosis at the same level that is

observed in human disease has always been a challenge when using mouse models.

This limitation did not allow us to do a correlation between fibrosis and N-glycan

modification in our mouse models. However, mice fed a WD did develop a higher level of

fibrosis relative to the mice fed a normal or a high-fat diet (NAFL model), this was

confirmed by two different immunohistochemical stainings. In addition, we believe we

were able to induce NASH disease based on mice exhibiting other characteristics of the

disease and based on the similar N-glycan profiles between "NASH" mice and NASH

human samples. To circumvent this challenge in future studies, as described in Chapter

3, the use of hepatotoxic agents like Carbon Tetrachloride (CCL_4) could be used to

induce a similar level of fibrosis as in humans that in combination with diet could make this model a more clinically relevant model.

A second limitation of this study was the limited clinical information that we were able to obtain from human NASH liver biopsies. We were provided liver biopsies that were diagnosed as NASH patients with a fibrosis score given by an independent pathologist at the moment of diagnosis. However, more clinical information like the scoring of steatosis and inflammation would have provided our study a greater value for the characterization of the complete spectrum of disease instead of only fibrosis.

8.2.3 Future directions

An important direction that would address fundamental biological questions is investigating the cell origin of these N-glycan modifications. Aim 1 provided N-glycan origin in terms of histopathological alterations within the tissue like fibrotic or steatotic areas. However, knowledge of the cells expressing these modifications would be useful for elucidating specific cell types involved in the disease and targeting these cells for further investigation including their role and/or mechanism in the progression of the disease.

Another major direction of the potential of these modifications for clinical use is the use of NASH serum samples in combination with tissue samples. Since N-glycan modifications characterized in Aim 1 do correlate with some of the modifications reported in serum and many of the serum glycoproteins are known to have a liver origin, it would be ideal to validate these findings in a tissue-serum-matched cohort. Our group has developed a biofluid N-glycan profiling approach (also used in chapter 5) that would characterize N-glycans in serum with high throughput. Using this strategy in NAFL/NASH we would be able to confidently answer if the same N-glycan profile

present in tissue can also be identified in serum. Ideally, clinical information like liver enzyme levels, steatosis, inflammation, and fibrosis scores would be available. Overall, the use of clinical information coupled with N-glycan structure modifications could help establish promising biomarker strategies for clinical diagnostics.

8.3 Characterization of N-glycosylation changes in liver disease progression: from early liver disease to primary liver cancers.

8.3.1 Conclusions

NALFD is one of the most common etiologies for chronic liver diseases (CLDs) and is also known to progress to a wide variety of liver malignancies. NAFLD itself includes liver damage like steatosis, inflammation, and fibrosis, and the progression of these can result in cirrhosis and primary liver cancers including Hepatocellular Carcinoma (HCC) and Cholangiocarcinoma (CCA). Aim 2 focused on understanding the N-glycan modifications at each of the stages of the disease that can progress from NAFLD, to address the variety of liver injuries and malignancies in a reliable model system we divided this aim (sub-aims 2.1 and 2.2) based on the disease of interest.

Based on findings by our group and others, there is a good understanding of the most common N-glycan modifications observed in liver disease (48,76,153,158,265). However, a major gap is the lack of information on the timing of these N-glycan modifications during the progression of these liver diseases, more specifically, we need to understand when these modifications are initiated and how are they changing from the early stages of NAFLD to the development of liver cancer. *Aim 2.1 characterized the temporal changes in N-glycosylation in a NAFLD and HCC-induced mouse models.* This sub-aim focused on addressing limitations in previous studies focused on a specific liver disease, and not on a liver disease progression, and was proposed in part due to some

of the limitations of aim 1 related to not being able to characterize the full spectrum of disease due to limited accessibility of clinical information. To circumvent these limitations we utilized established liver disease inducible mouse models to induce NAFL (steatosis), NASH (steatosis, fibrosis, and inflammation), advanced-liver disease (severe fibrosis, and inflammation), and nodule formation (HCC-like characteristics). We divided our studies between NAFL to NASH progression (induced only by diet) and NAFL/NASH to HCC-related characteristics progression (induced by diet and carcinogen). The different stages of diseases were based on specific time points proposed based on the literature. As is described in Chapter 5, we elucidated specific N-glycan signatures present in specific stages of the disease. In addition, we were able to track specific modifications like fucosylated N-glycans from very early liver damage to the nodular formation stage. Importantly, we reported that N-glycan modification could be observed even before any histological liver disease characteristics were identified. This finding suggests, that N-glycosylation, specifically residues within these N-glycan structures like fucose residues, could be playing an important role in disease initiation and that the identification of these modifications is not dependent on histological information. In addition, as illustrated in Chapter 5, we coupled our N-glycan MALDI-IMS images with immunohistochemical images of important immune cell markers and identified a correlation between N-glycan modification with immune cell populations.

Overall, the completion of this aim, reveals specific N-glycan modifications that should be further studied for their biomarker capabilities. In addition, based on how early N-glycan alterations initiate in liver disease, we can confirm the value of further exploring the N-glycan-related genes that are responsible for those N-glycan modifications in an in vivo model to determine if these have roles in disease initiation and/or promotion.

Previous work by our group and others has identified the N-glycan modification present in HCC and within HCC subtypes (48,76). However, N-glycan modifications in CCA have not been addressed and it was not known if the same N-glycan modification observed in HCC would also be observed in CCA since they are both primary liver cancers. *Aim 2.2 Identified promising biomarker strategies based on N-glycosylation changes in Cholangiocarcinoma using human tissue and serum.* For this sub-aim, we were particularly interested in intrahepatic CCA (iCCA), which is the main subtype known to progress from NASH. While the first part of this aim was focused on utilizing mouse models for liver disease induction, mouse models for CCA are very limited, and replicating all the characteristics of CCA in the single mouse models has been challenging. Instead, we used human tissues for N-glycan characterization and took a biomarker discovery approach by also using serum samples. Tissue microarrays that included tissues from patients with HCC, CCA, and other types of liver diseases were used to identify a full N-glycome profile for CCA. Importantly, we identified a bisected doubly fucosylated N-glycan modification to be present with high specificity to samples from CCA patients. We further explored this modification in an independent serum sample set, and also identified the same bisected doubly fucosylated N-glycan structure as in tissue. This N-glycan modification was able to distinguish CCA from other types of liver diseases including fatty liver, PSC, viral hepatitis, and healthy samples in tissue and serum. Finally, we were interested to determine if this modification could distinguish CCA from primary Sclerosing cholangitis (PSC), which is considered a more clinically relevant situation since PSC is a major etiology and bile duct disease and diagnosis between these diseases has been challenging. Interestingly, the combination of a

bisected doubly fucosylated N-glycan with other common N-glycans was able to identify iCCA and perform better than the gold standard for CCA diagnosis, CA19-9.

Overall Aim 2, focused on addressing the N-glycan-related research gaps in diseases that are known to progress from NALF/NASH with the idea of identifying hallmarks, promising drivers, and biomarkers of different liver diseases.

8.3.2 Limitations

An important limitation of aim 2.1 were the challenges in inducing a full stage of liver cancer, specifically HCC, at least histologically. Our selected end-point of 11.5 months was based on previous literature that had observed HCC disease at 9 months. However, histological analysis and annotations did not identify HCC in our models. Instead, pathology analysis identified cancer-related characteristics like nodular abnormalities, abnormal cells, and hyperplasia. Based on these disease characteristics that are very related to HCC, we believe that a longer time point of approximately 1 month could have resulted in the complete induction of HCC. Another reason for not inducing HCC completely, even with the use of a carcinogen could be due to a late start of the carcinogen treatment at 4 weeks old instead of 2 weeks as most studies had reported. Since 2 weeks is considered a very early age in mice and a vital stage for liver development, we preferred to postpone carcinogen treatment to 4 weeks old, to avoid interfering with normal liver development. We wanted to be certain that the N-glycan changes were characterizing was from the diet and/or carcinogen, and not associated with the disruption in processes for normal liver development. To address this limitation in future studies, as described in Chapter 3, other hepatotoxin combinations or strategies could be used to induce HCC phenotypes. Another important limitation of aim 2.1 was the low number of mice that were used for some of the time points. Even though our

group planned for an n of at least 6 animals per group, different situations related to mouse housing and health issues with some of the mice during the study resulted in a low number of animals. Future studies should plan for a higher number of mice, taking into consideration that some mice might need to be removed from the study due to health concerns.

A limitation of aim 2.2 was the use of an independent serum sample set to link N-glycan modifications identified in tissue. While the serum samples used in our study revealed consistent N-glycan modifications to those observed in tissue, the ideal set for biomarker studies would be a matched tissue and serum patient set, meaning that serum samples were from the same patients that the tissue was collected from. This limitation is one of the most common limitations in doing translational research, where the accessibility to these types of sample sets is very challenging. The use of a matched sample set would allow answering with certainty to the relationship between N-glycan observed in tissue and those in serum and the value of N-glycan modifications in tissue to be proposed for clinical use.

8.3.3 Future directions

An important direction for aim 2.1, is the elucidation of the type of fucosylation observed in timepoint studies. Fucosylation is catalyzed by different fucosyltransferases where each will add a fucose residue to the N-glycan at different linkages, the different types of fucosylation can be divided into a core or outer-arm fucosylation. This elucidation could be addressed using an exoglycosidase enzyme (Endo F3) during the MALDI-IMS protocol. EndoF3 has a high specificity for core fucosylated N-glycan, and it will cleave them between two GlcNac residues at the core of the N-glycan. Core

fucosylated N-glycans can be identified during analysis by a mass shift of the expected m/z value for the specific N-glycan structure.

A second important direction that could be taken into consideration from aim 2.1 is to target specific glycosyltransferases responsible for N-glycan modifications observed. Since the main goal would be to help improve early detection strategies for liver diseases, N-glycan structure modifications that were observed at early stages of the disease as fucosylated N-glycans would be of interest. Fucosylated N-glycan modifications were observed before significant histological modifications were observed, this suggests that fucose N-glycan-related genes, like fucosyltransferases, would be ideal for targeting in a mouse model to elucidate their role in early liver disease.

A major direction that our group will follow up on for aim 2.2 is identifying the glycoproteins which these N-glycan modifications are originating from. Due to the specific bisected fucosylated N-glycan modification to only iCCA tissue and serum, we are interested in elucidating if this specificity is due to a specific glycoprotein that could be proposed for the biomarker tool. This study would add value to our study by now combining information in the glycoprotein in addition to the specific N-glycan modification that can identify patients with iCCA.

8.4 Generation of an in vivo model to study core fucosylation

8.4.1 Conclusions

Core fucosylation is one of the major N-linked glycosylation modifications in many cancers, including liver cancer. Core fucosylation is catalyzed by Fucosyltransferase 8 (FUT8), and its alteration with HCC has been exploited for biomarker use. Alpha-fetoprotein (AFP) is a serum glycoprotein and a diagnostic serum marker for HCC. However, the low specificity of AFP has led to the use of AFP-L3, the

core fucosylated form of AFP as the preferred marker for predicting HCC (266). AFP-L3 levels have been associated with HCC early detection, aggressiveness, and poor prognosis (229). Similarly, studies have correlated FUT8/core fucosylated expression with HCC survival and progression in many system models. In addition, the current studies outlined in this dissertation confirm the importance of FUT8/core fucosylation in primary liver cancers and also contribute to the field by reporting the importance of core fucosylation in early liver diseases like NAFL/NASH. However, FUT8/Core fucosylation thus far has been studied for the characterization of its expression under different experimental conditions, but the elucidation of the role and mechanism of FUT8/core fucosylation in the liver *in vivo* has not been addressed. One of the main limitations in addressing this gap is a lack of *in vivo* study models. The only current model is a whole-body FUT8 knockout mouse model that demonstrated to have a decrease in liver tumor incidence when exposed to a carcinogen. However, this model poses questions due to suffering from severe phenotypic complications like post-natal lethality, emphysema, and seizures. Aim 3 addressed the need for liver-specific study models and generates a hepatocyte-specific deletion of *Fut8* mouse model. The generation of this model followed different strategies, including CRISPR-Cas9 and a liver-specific (Albumin) Cre-LoxP strategy. It is important to note that this aim included the generation of two different mouse models, one is the Fut8 fl/fl model which has the insertion of the two LoxP sites, and the second is the Albumin-Cre *Fut8* model (which is the result from the breeding of Fut8fl/fl with the Albumin Cre). The generation of this model was in collaboration with the Transgenic and Genomic core at MUSC, they were in charge of the model design and our lab took charge of the colony maintenance, breeding, and genotyping. As expected due to the low expression of *Fut8* in the liver under normal conditions, this model does

not show any phenotype alterations at 10 weeks of age. To our knowledge, this is the first and only *Fut8* liver-specific mouse model and its use will be an exponential advancement for liver disease clinical strategies.

8.4.2 Limitations

Thus far this model does not present any specific limitations. Future limitations in the model itself could be due to the type of Cre used in the design. The Albumin Cre model used will excise the LoxP sites from fetal stages, this will limit the studies related to *Fut8* at different stages of development or in disease progression (261). To address this limitation, the Fut8 fl/fl model could be bred with an inducible Cre system to select for a specific time point when Fut8 deletion is needed. By using an inducible Cre system, gene deletion will occur by the exogenous inducer tamoxifen or tetracyclin treatment, providing temporal control of gene deletion.

Another potential problem of the model is that during characterization experiments, *Fut8* expression is still present in the liver, specifically in hepatocytes. A possible explanation for this could be due to the *Fut8* fl/fl model not having active LoxP sites or due to an issue with Albumin Cre excision. Resolving this problem would need in vitro studies to first confirm the activity of the LoxP sites and reanalyze the model design to possibly insert new LoxP sites.

Another potential, but unexpected limitations, that could present itself when we start using the model in a disease context include that the loss of Fut8/core fucosylation does not induce any phenotypic differences in liver diseases, possibly due to a compensatory mechanism when hepatic *Fut*8 is absent. While this information would still be valuable for the field. We would analyze this further by comparing the N-glycan profiles

between genotypes and determine if other glycosyltransferases are compensating for *Fut8's* role.

8.4.3 Future directions

The first direction that our lab will address is the characterization of this model, where we will confirm the deletion of *Fut8* in the liver at the transcript and protein levels. In addition, we will also take into consideration the gene expression of other N-glycan-related genes like glycosidases, transporters, and substrates to determine their response when *Fut8* is not present. We will also analyze the effects of *Fut8* deletion on liver morphology, liver histology, cell proliferation, migration, and apoptosis. Finally, we are interested in identifying how the loss of core fucosylation can affect the complete N-glycome profile in tissue and serum This would be interesting to determine if the loss of core fucosylation is compensated by other glycosyltransferases and results in the generation of different N-glycan structures. The reasoning for also characterizing the serum profiles is based on the idea that most of the serum glycoproteins are originating from the liver and the deletion of *Fut8* will most likely result in the loss of most of the core fucosylation normally observed in serum. The characterization of this model will be the start of the many directions this mouse model could be used in.

A second major direction is in the context of liver diseases, AlbCre-Fut8 mouse models with their respective controls will be used to induce liver diseases previously studied in this dissertation. We will use this model to elucidate the role of hepatic Fut8 in NAFLD, NASH, and HCC. A similar strategy for inducing diseases will be implemented to the ones used in Aims 1 and 2 to elucidate the role in liver disease initiation, progression, and metastasis. The main interest in this direction will be to determine if *Fut8* and core fucosylation are required for liver disease. Additionally, we would be

interested in identifying the mechanism by which Fut8 is involved in liver disease. Previous studies have reported that core fucosylation of EGFR, TGFB, and related molecular pathways leads to an increase in cell proliferation in disease conditions (94,95,255,267–270). Our goal would be to identify the main pathway in which *Fut8* is involved in liver disease and use this knowledge to propose therapeutic strategies.

The future directions mentioned here are more related to a liver context. However, the generation of this model sets the field of N-glycobiology and cancer biology to answer many underlying questions related to *Fut8*. Breast, colorectal, ovarian, lung, prostate, and pancreatic cancers have all been reported *Fut8* to have altered expression. All of these cancers can be studied by the use of *Fut8* fl/fl model bred with any tissue-specific Cre.

8.5 Final Thoughts

Collectively, the studies presented in this dissertation were proposed with the idea to address research gaps in the field of liver disease and N-linked glycosylation and to contribute to the translational field in improving diagnostic and therapeutic strategies. In brief, the completion of the Aims described in this dissertation elucidated the origin of N-glycan modifications in NAFLD mouse and human tissues, this was further explored in Aim 2 which took into consideration the variety of liver diseases that could progress from NAFLD and identified potential regulators of disease by studying the different stages of liver disease progression. In addition, we applied a biomarker approach to our N-glycan studies in CCA by proposing the use of N-glycan modification for biomarker strategies. Finally, the completion of Aim 3 contributes largely to the field by generating a mouse model that could answer essential biological questions about core fucosylation in any type of disease. Overall, this work establishes the importance of N-linked glycosylation

alterations for clinical use in early liver diseases and sets the field for the further

exploration of N-glycan-related genes mechanisms and structures that could help

develop applicable therapeutic strategies.

REFERENCES

1. Arjun Kalra1, Yetiskul E, Wehrle CJ, Faiz Tuma. Physiology, Liver. Treasure Island, fl: StatPearls Publishing; 2022. Available from: https://www.ncbi.nlm.nih.gov/books/NBK535438/#_article-24376_s1_

2. Moon AM, Singal AG, Tapper EB. Contemporary Epidemiology of Chronic Liver Disease and Cirrhosis. 2020;18:2650–66. Available from: https://doi.org/10.1016/ S2468-1253(19)30349-8

3. Collaborators C. The global , regional , and national burden of cirrhosis by cause in 195 countries and territories , 1990 – 2017 : a systematic analysis for the Global Burden of Disease Study 2017. 2020;245–66. Available from: https://doi.org/10.1016/ S2468-1253(19)30349-8

4. Sharma A, Nagalli S. Chronic Liver Disease. StatPearls Publishing; 2022. Available from: https://www.ncbi.nlm.nih.gov/books/NBK554597/#_NBK554597_pubdet_

5. Kozlitina J. Genetic Risk Factors and Disease Modifiers of Nonalcoholic Steatohepatitis. Gastroenterol Clin North Am [Internet]. Elsevier Inc; 2020;49:25–44. Available from: https://doi.org/10.1016/j.gtc.2019.09.001

6. Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. J Hepatol [Internet]. European Association for the Study of the Liver; 2019;70:151–71. Available from: https://doi.org/10.1016/j.jhep.2018.09.014

7. Parthasarathy G, Revelo X, Malhi H. Pathogenesis of Nonalcoholic Steatohepatitis: An Overview. Hepatol Commun. 2020;4:478–92.

8. Muthiah MD, Sanyal AJ. Burden of Disease due to Nonalcoholic Fatty Liver Disease. Gastroenterol Clin North Am [Internet]. Elsevier Inc; 2020;49:1–23. Available from: https://doi.org/10.1016/j.gtc.2019.09.007

9. Maurice J, Manousou P. Non-alcoholic Fatty Liver Disease. Clin Med (Northfield ll). 2018;18:245–50.

10. Hardy T, Anstee QM, Day CP. Nonalcoholic fatty liver disease: New treatments. Curr Opin Gastroenterol. 2015;31:175–83.

11. Ahmed A, Wong RJ, Harrison SA. Nonalcoholic Fatty Liver Disease Review: Diagnosis, Treatment, and Outcomes. Clin Gastroenterol Hepatol [Internet]. Elsevier, Inc; 2015;13:2062–70. Available from: http://dx.doi.org/10.1016/j.cgh.2015.07.029

12. Hirsova P. Modeling NASH and NASH-Induced Hepatocellular Carcinoma: Faster and Better. Cmgh [Internet]. The Author; 2021;12:1149–50. Available from: https://doi.org/10.1016/j.jcmgh.2021.05.015

13. Hashimoto E, Tokushige K. Prevalence, gender, ethnic variations, and prognosis

of NASH. 2011;46:63–9.

14. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64:73–84.

15. Natarajan Y, Kramer JR, Yu X, Li L, Thrift AP, El-Serag HB, et al. Risk of Cirrhosis and Hepatocellular Cancer in Patients With NAFLD and Normal Liver Enzymes. Hepatology. 2020;72:1242–52.

16. Wong RJ, Aguilar M, Cheung R, Perumpail RB, Harrison SA, Younossi ZM, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology [Internet]. Elsevier, Inc; 2015;148:547–55. Available from: http://dx.doi.org/10.1053/j.gastro.2014.11.039

17. Mikolasevic I, Filipec-Kanizaj T, Mijic M, Jakopcic I, Milic S, Hrstic I, et al. Nonalcoholic fatty liver disease and liver transplantation - Where do we stand? World J Gastroenterol. 2018;24:1491–506.

18. Benedict M, Zhang X. Non-alcoholic fatty liver disease: An expanded review. World J Hepatol. 2017;9:715–32.

19. Lonardo A, Nascimbeni F, Ballestri S, Fairweather DL, Win S, Than TA, et al. Sex Differences in Nonalcoholic Fatty Liver Disease: State of the Art and Identification of Research Gaps. Hepatology. 2019;70:1457–69.

20. Eslam M, Valenti L, Romeo S. Genetics and epigenetics of NAFLD and NASH : Clinical impact. J Hepatol [Internet]. European Association for the Study of the Liver; 2018;68:268–79. Available from: https://doi.org/10.1016/j.jhep.2017.09.003

21. Mann JP, Pietzner M, Wittemans LB, Rolfe EDL, Kerrison ND, Imamura F, et al. Insights into genetic variants associated with NASH-fibrosis from metabolite profiling. Hum Mol Genet. 2020;29:3451–63.

22. Riazi K, Raman M, Taylor L, Swain MG, Shaheen AA. Dietary patterns and components in nonalcoholic fatty liver disease (NAFLD): What key messages can health care providers offer? Nutrients. 2019;11:1–17.

23. Wongjarupong N, Assavapongpaiboon B, Susantitaphong P, Cheungpasitporn W, Treeprasertsuk S, Rerknimitr R, et al. Non-alcoholic fatty liver disease as a risk factor for cholangiocarcinoma: a systematic review and meta-analysis. BMC Gastroenterol. BMC Gastroenterology; 2017;17:149.

24. Fedchuk L, Nascimbeni F, Pais R, Charlotte F, Housset C, Ratziu V. Performance and limitations of steatosis biomarkers in patients with nonalcoholic fatty liver disease. Aliment Pharmacol Ther. 2014;40:1209–22.

25. Kuchay MS, Martinez-Montoro JI, Choudhary NS, Fernandez-Garcia JC, Ramos-molina B. Non-Alcoholic Fatty Liver Disease in Lean and Non-Obese Individuals : Current and Future Challenges. Biomedicines. 2021;9:1–21.

26. Ajmera V, Loomba R. Imaging biomarkers of NAFLD , NASH , and fi brosis. Mol Metab. 2020;1–7.

27. Dietrich CG, Rau M, Geier A. Screening for nonalcoholic fatty liver disease-when, who and how? World J Hepatol. 2021;27:5803–21.

28. Alsaleh M, Leftley Z, Barbera TA, Sithithaworn P, Khuntikeo N, Loilome W, et al. Cholangiocarcinoma: A guide for the nonspecialist. Int J Gen Med. 2019;12:13–23.

29. Tischendorf JJW, Hartmut H, Kruger M, Manns MP, Meier PN. Characterization , Outcome , and Prognosis in 273 Patients with Primary Sclerosing Cholangitis : A Single Center Study. Am J Gastroenterol. 2007;107–14.

30. Bambha K, Kim RW, Talwalkar J, Torgerson H, Benson JT, Therneau TM, et al. Incidence, Clinical Spectrum, and Outcomes of Primary Sclerosing Cholangitis in a United States Community. Gastroenterology. 2003;125:1364–9.

31. Muñoz-Martínez S, Rimola J, Londoño MC, Cárdenas A, Forner A. Cholangiocarcinoma: early detection and screening in high-risk population. Hepatoma Res [Internet]. 2022;8:30. Available from: https://hrjournal.net/article/view/5072

32. KARLSEN TH, FRANKE A, MELUM E, KASER A, HOV JR, BALSCHUN T, et al. Genome-Wide Association Analysis in Primary Sclerosing Cholangitis. Gastroenterology [Internet]. Elsevier Inc.; 2010;138:1102–11. Available from: http://dx.doi.org/10.1053/j.gastro.2009.11.046

33. Rawla P, Samant H. Primary Sclerosing Cholangitis. Treasure Island, fl: StatPearls Publishing; 2022.

34. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis : A systematic review. J Hepatol [Internet]. European Association for the Study of the Liver; 2012;56:1181–8. Available from: http://dx.doi.org/10.1016/j.jhep.2011.10.025

35. Molodecky NA, Kareemi H, Parab R, Barkema HW, Quan H, Myers RP, et al. Incidence of Primary Sclerosing Cholangitis: a Systematic Review and Meta-Analysis. Hepatology. 2011;53:1417–788.

36. Lazaridis KN, LaRusso NF. The Cholangiopathies. Mayo Clin Proc. 2015;90:791–800.

37. Rawla P, Sunkara T, Raj JP. Role of biologics and biosimilars in inflammatory bowel disease: current trends and future perspectives. J Inflamm Res. 2018;11:215–26.

38. Buettner S, Van Vugt JLA, Ijzermans JNM, Koerkamp BG. Intrahepatic cholangiocarcinoma: Current perspectives. Onco Targets Ther. 2017;10:1131–42.

39. Fujiwara N, Liu P-H, Athuluri-Divakar SK, Zhu S, Hoshida Y. Risk Factors of

Hepatocellular Carcinoma for Precision Personalized Care. 2019;3–25.

40. Ren Z, Ma X, Duan Z, Chen X. Diagnosis, Therapy, and Prognosis for Hepatocellular Carcinoma. Anal Cell Pathol. 2020;2020:3–4.

41. Nelson Hayes C, Zhang P, Chayama K. The Role of Lipids in Hepatocellular Carcinoma. In: Tirnits-Parker J, editor. Hepatocell Carcinoma. Codon Publications; 2019. page 95–110.

42. Natri HM, Wilson MA, Buetow KH. Distinct molecular etiologies of male and female hepatocellular carcinoma. BMC Cancer. BMC Cancer; 2019;19:1–12.

43. Wu EM, Wong LL, Hernandez BY, Ji J, Jia W, Sandi A, et al. Gender differences in hepatocellular cancer: disparities in nonalcoholic fatty liver disease/steatohepatitis and liver transplantation. Hepatoma Res. 2018;4:1–17.

44. Janevska D, Chaloska-ivanova V, Janevski V. Hepatocellular Carcinoma: Risk Factors, Diagnosis and Treatment. 2015;3:732–6.

45. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention. Nat Rev Gastroenterol Hepatol. 2019;16:589–604.

46. Raza A, Sood GK. Hepatocellular carcinoma review : Current treatment , and evidence-based medicine. World J Gastroenterol. 2014;20:4115–27.

47. Schneller D, Angel P. Cellular Origin of Hepatocellular Carcinoma. In: Tirnitz-Parker JEE, editor. Hepatocell Carcinoma. Codon Publications; 2019. page 1–28.

48. DelaCourt A, Black A, Angel P, Drake R, Hoshida Y, Singal A, et al. N-Glycosylation Patterns Correlate with Hepatocellular Carcinoma Genetic Subtypes. Mol Cancer Res. 2021;1–11.

49. Sandhu DS, Roberts LR. Diagnosis and management of cholangiocarcinoma. Curr. Gastroenterol. Rep. 2008.

50. Sarcognato S, Sacchi D, Fassan M, Fabris L, Cadamuro M, Zanus G, et al. Cholangiocarcinoma. Pathologica. 2021;113:158–69.

51. Brindley PJ, Bachini M, Ilyas SI, Khan SA, Loukas A, Sirica AE, et al. Cholangiocarcinoma. Nat Rev Dis Prim. 2021;7.

52. Blechacz B. Cholangiocarcinoma : Current Knowledge and New Developments. Gut Liver. 2017;11:13–26.

53. Razumilava N, Gores GJ. Cholangiocarcinoma. Lancet. 2014;383:2168–79.

54. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol [Internet]. Springer US; 2020;17:557–88. Available from: http://dx.doi.org/10.1038/s41575-020-0310-z

55.   Kirstein MM, Vogel A. Epidemiology and Risk Factors of Cholangiocarcinoma. Visc Med. 2016;32:395–400.

56.   Scara S, Bottoni P, Scatena R. CA 19-9: Biochemical and Clinical Aspects. In: Scatena R, editor. Adv Cancer Biomarkers. Springer US; 2015. page 247–60.

57.   Sriwanitchrak P, Viyanant V, Chaijaroenkul W, Srivatanakul P, Gram HR, Eursiddhichai V, et al. Proteomics analysis and evaluation of biomarkers for detection of cholangiocarcinoma. Asian Pacific J Cancer Prev. 2011;12:1503–10.

58.   Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. Biochim Biophys Acta - Gen Subj. Elsevier B.V.; 2012;1820:1347–53.

59.   Schwarz F, Aebi M. Mechanisms and principles of N-linked protein glycosylation. Curr Opin Struct Biol [Internet]. Elsevier Ltd; 2011;21:576–82. Available from: http://dx.doi.org/10.1016/j.sbi.2011.08.005

60.   Apweiler R, Henning H, Sharon N. On the frequency of protein glycosylation , as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta. 1999;4–8.

61.   Pinho SS, Reis CA. Glycosylation in cancer: Mechanisms and clinical implications. Nat Rev Cancer. Nature Publishing Group; 2015;15:540–55.

62.   Gagneux P, Hennet T, Varki A. Biological Functions of Glycans. In: Glycobiology TC of, editor. Essentials Glycobiol. 4th ed. La Jolla, California: Cold Spring Harbor Laboratory Press; 2022.

63.   Lütteke T, Frank M, editors. Glycoinformatics. 2015th ed. Humana Press; 2015.

64.   Lowe JB, Marth JD. A Genetic Approach to Mammalian Glyan Function. Annu Rev Biochem. 2003;72:643–91.

65.   Cummings R, Esko JD, Stanley P, Hart GW, Aebi M, Mohnen D, et al., editors. Essentials of Glycobiology. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022.

66.   Tian Y, Zhang H. Characterization of disease-associated N-linked glycoproteins. Proteomics. 2013;

67.   Chen B, Liu W, Li Y, Ma B, Shang S, Tan Z. Impact of N-Linked Glycosylation on Therapeutic Proteins. Molecules. 2022;27:1–10.

68.   Blomme B, Van Steenkiste C, Callewaert N, Van Vlierberghe H. Alteration of protein glycosylation in liver diseases. J Hepatol [Internet]. European Association for the Study of the Liver; 2009;50:592–603. Available from: http://dx.doi.org/10.1016/j.jhep.2008.12.010

69.   Kang X, Wang N, Pei C, Sun L, Sun R, Chen J, et al. Glycan-related gene expression signatures in human metastatic hepatocellular carcinoma cells. Exp Ther Med. 2012;3:415–22.

70. Stanley P, Taniguchi N, Aebi M. N-Glycans. Essentials Glycobiol [Internet]. 3rd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK453020/

71. Rudd PM, Karlsson NG, Khoo K-H, Thaysen-Andersen M, Wells L, Packer NH. Glycomics and Glycoproteomics. Essentials Glycobiol [Internet]. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579904/?report=reader#_NBK579904_pubdet_

72. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. Nat Rev Nephrol [Internet]. Springer US; 2019;15:346–66. Available from: http://dx.doi.org/10.1038/s41581-019-0129-4

73. Kam RKT, Poon TCW. The potentials of glycomics in biomarker discovery. Clin Proteomics. 2008;4:67–79.

74. Drake RR. Glycosylation and cancer: Moving glycomics to the forefront [Internet]. 1st ed. Adv. Cancer Res. Elsevier Inc.; 2015. Available from: http://dx.doi.org/10.1016/bs.acr.2014.12.002

75. McDowell CT, Lu X, Mehta AS, Angel PM, Drake RR. Applications and continued evolution of glycan imaging mass spectrometry. Mass Spectrom Rev. 2021;

76. West CA, Wang M, Herrera H, Liang H, Black A, Angel PM, et al. N-Linked Glycan Branching and Fucosylation Are Increased Directly in Hcc Tissue As Determined through in Situ Glycan Imaging. J Proteome Res. 2018;17:3454–62.

77. Ochoa-Rios S, O'Connor IP, Kent LN, Clouse JM, Hadjiyannis Y, Koivisto C, et al. Imaging Mass Spectrometry Reveals Alterations in N-linked glycosylation that are Associated with Histopathological Changes in Non-alcoholic Steatohepatitis in Mouse and Human. Mol Cell Proteomics [Internet]. The Authors; 2022;21:1–12. Available from: https://doi.org/10.1016/j.mcpro.2022.100225

78. Liu XE, Desmyter L, Gao CF, Laroy W, Dewaele S, Vanhooren V, et al. N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus. Hepatology. 2007;46:1426–35.

79. Betesh L, Comunale MA, Wang M, Liang H, Hafner J, Karabudak A, et al. Identification of fucosylated Fetuin-A as a potential biomarker for cholangiocarcinoma. Proteomics - Clin Appl. 2017;11:9–10.

80. Talabnin K, Talabnin C, Ishihara M, Azadi P. Increased expression of the high-mannose M6N2 and NeuAc3H3N3M3N2F tri-antennary N-glycansin cholangiocarcinoma. Oncol Lett. 2018;15:1030–6.

81. Indramanee S, Silsirivanit A, Pairojkul C, Wongkham C, Wongkham S. Aberrant glycosylation in cholangiocarcinoma demonstrated by lectin-histochemistry. Asian Pacific J Cancer Prev. 2012;13:119–24.

82. Cheng L, Gao S, Song X, Dong W, Zhou H, Zhao L, et al. Comprehensive N-

glycan profiles of hepatocellular carcinoma reveal association of fucosylation with tumor progression and regulation of FUT8 by microRNAs. Oncotarget. 2016;7:61199–214.

83.   Lannoo N, Van Damme EJM. Review/N-glycans: The making of a varied toolbox. Plant Sci [Internet]. Elsevier Ireland Ltd; 2015;239:67–83. Available from: http://dx.doi.org/10.1016/j.plantsci.2015.06.023

84.   M. RJ, Esko JD. Glycosyltransferases and Glycan-Processing Enzymes. Essentials Glycobiol [Internet]. 3rd editio. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK453021/ doi: 10.1101/glycobiology.3e.006

85.   West CA, Black AP, Mehta AS. Analysis of Hepatocellular Carcinoma Tissue for Biomarker Discovery. 2019;93–107.

86.   Comunale MA, Lowman M, Long RE, Krakover J, Philip R, Seeholzer S, et al. Proteomic analysis of serum associated fucosylated glycoproteins in the development of primary hepatocellular carcinoma. J Proteome Res. 2006;5:308–15.

87.   Wang M, Long RE, Comunale MA, Junaidi O, Marrero J, Di Bisceglie AM, et al. Novel fucosylated biomarkers for the early detection of hepatocellular carcinoma. Cancer Epidemiol Biomarkers Prev. 2009;18:1914–21.

88.   Kizuka Y, Taniguchi N. Enzymes for N-Glycan branching and their genetic and nongenetic regulation in cancer. Biomolecules. 2016;6:1–21.

89.   Ogawa K, Kobayashi T, Furukawa J ichi, Hanamatsu H, Nakamura A, Suzuki K, et al. Tri-antennary tri-sialylated mono-fucosylated glycan of alpha-1 antitrypsin as a non-invasive biomarker for non-alcoholic steatohepatitis: a novel glycobiomarker for non-alcoholic steatohepatitis. Sci Rep. 2020;10:1–10.

90.   Yamasaki Y, Nouso K, Miyahara K, Wada N, Dohi C, Morimoto Y, et al. Use of non-invasive serum glycan markers to distinguish non-alcoholic steatohepatitis from simple steatosis. J Gastroenterol Hepatol. 2015;30:528–34.

91.   Wang M, Fang M, Zhu J, Feng H, Warner E, Yi C, et al. Serum N-glycans outperform CA19-9 in diagnosis of extrahepatic cholangiocarcinoma. Electrophoresis. 2017;38:2749–56.

92.   Betesh L, Comunale MA, Wang M, Liang H, Hafner J, Karabudak A, et al. Identification of fucosylated Fetuin-A as a potential biomarker for cholangiocarcinoma. Proteomics - Clin Appl. 2017;11:9–10.

93.   Wang Y, Fukuda T, Isaji T, Lu J, Gu W, Lee HH, et al. Loss of α1,6-fucosyltransferase suppressed liver regeneration: Implication of core fucose in the regulation of growth factor receptor-mediated cellular signaling. Sci Rep. 2015;5:4–10.

94. Wang Y, Fukuda T, Isaji T, Lu J, Im S, Hang Q, et al. Loss of α1,6-fucosyltransferase inhibits chemical-induced hepatocellular carcinoma and tumorigenesis by down-regulating several cell signaling pathways. FASEB J. 2015;29:3217–27.

95. Wang X, Gu J, Ihara H, Miyoshi E, Honke K, Taniguchi N. Core fucosylation regulates epidermal growth factor receptor-mediated intracellular signaling. J Biol Chem. 2006;281:2572–7.

96. Clarke JD, Novak P, Lake AD, Hardwick RN, Cherrington NJ. Impaired N-linked glycosylation of uptake and efflux transporters in human non-alcoholic fatty liver disease. Liver Int. 2017;37:1074–81.

97. Chen C, Schmilovitz-Weiss H, Liu XE, Pappo O, Halpern M, Sulkes J, et al. Serum protein N-glycans profiling for the discovery of potential biomarkers for nonalcoholic steatohepatitis. J Proteome Res. 2009;8:463–70.

98. A V, R K, B T, Pamela S. Glycosylation Changes in Cancer. In: A V, RD C, JD E, editors. Essentials Glycobiol [Internet]. 3rd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2017. Available from: https://www.ncbi.nlm.nih.gov/books/NBK453023/ doi: 10.1101/glycobiology.3e.047

99. Taniguchi N, Ohkawa Y, Maeda K, Harada Y, Nagae M, Kizuka Y, et al. True significance of N-acetylglucosaminyltransferases GnT-III, V and α1,6 fucosyltransferase in epithelial-mesenchymal transition and cancer. Mol Aspects Med [Internet]. Elsevier Ltd; 2020;100905. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33010941

100. Cheng L, Cao L, Wu Y, Xie W, Li J, Guan F, et al. Bisecting N-Acetylglucosamine on EGFR Inhibits Malignant Phenotype of Breast Cancer via Down-Regulation of EGFR/Erk Signaling. Front Oncol. 2020;10:1–14.

101. Yang X, Bhaumik M, Bhattacharyya R, Gong S, Rogler CE, Stanley P. New evidence for an extra-hepatic role of N- acetylglucosaminyltransferase III in the progression of diethylnitrosamine- induced liver tumors in mice. Cancer Res. 2000;60:3313–9.

102. Bhaumik M, Harris T, Sundaram S, Johnson L, Guttenplan J, Rogler C, et al. Progression of hepatic neoplasms is severely retarded in mice lacking the bisecting N-acetylglucosamine on N-glycans: Evidence for a glycoprotein factor that facilitates hepatic tumor progression. Cancer Res. 1998;58:2881–7.

103. Yang X, Tang J, Rogler CE, Stanley P. Reduced Hepatocyte Proliferation is the Basis of Retarded Liver Tumor Progression and Liver Regeneration in Mice Lacking N-Acetylglucosaminyltransferase III. Cancer Res. 2003;63:7753–9.

104. Ekuni A, Miyoshi E, Ko JH, Noda K, Kitada T, Ihara S, et al. A Glycomic Approach to Hepatic Tumors in N -acetylglucosaminyltransferase III (GnT-III) Transgenic Mice Induced by Diethylnitrosamine (DEN): Identification of Haptoglobin as a

Target Molecule of GnT-III. Free Radic Res. 2002;

105. Yinghui Song, Jason A. Aglipay, Joshua D. Bernstein, Sumanta Goswami and PS. The Bisecting GlcNAc on N-Glycans Inhibits Growth Factor Signaling and Retards Mammary Tumor Progression. Cancer Epidemiol Biomarkers Prev. 2010;70.

106. Yoshimura M, Taniguchi N. Suppression of lung metastasis of B16 mouse melanoma cells by introduction of N-acetylglucosaminyltransferase III gene. Nippon rinsho Japanese J Clin Med. 1995;53:1786–90.

107. Yoshimura M, Ihara Y, Matsuzawa Y, Taniguchi N. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. J Biol Chem [Internet]. Â© 1996 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology.; 1996;271:13811–5. Available from: http://dx.doi.org/10.1074/jbc.271.23.13811

108. Pinho SS, Reis CA, Paredes J, Magalhães AM, Ferreira AC, Figueiredo J, et al. The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. Hum Mol Genet. 2009;18:2599–608.

109. Liu J, Liu H, Zhang W, Wu Q, Liu W, Liu Y, et al. N-acetylglucosaminyltransferase V confers hepatoma cells with resistance to anoikis through EGFR/PAK1 activation. Glycobiology. 2013;23:1097–109.

110. Liu T, Zhang S, Chen J, Jiang K, Zhang Q, Guo K, et al. The transcriptional profiling of glycogenes associated with hepatocellular carcinoma metastasis. PLoS One. 2014;9.

111. Liu H, Wu Q, Liu Y, Liu W, Zhang W, Pan D, et al. Prognostic significance of β1,6-N-acetylglucosaminyltransferase V expression in patients with hepatocellular carcinoma. Jpn J Clin Oncol. 2015;45:844–53.

112. Cui J, Huang W, Wu B, Jin J, Jing L, Shi WP, et al. N-glycosylation by N-acetylglucosaminyltransferase V enhances the interaction of CD147/basigin with integrin β1 and promotes HCC metastasis. J Pathol. 2018;245:41–52.

113. Terao M, Ishikawa A, Nakahara S, Kimura A, Kato A, Moriwaki K, et al. Enhanced epithelial-mesenchymal transition-like phenotype in N-acetylglucosaminyltransferase V transgenic mouse skin promotes wound healing. J Biol Chem. 2011;286:28303–11.

114. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW. Suppression of tumor growth and metastasis in Mgat5-deficient mice. Nat Med. 2000;6:306–12.

115. Kamada Y, Mori K, Matsumoto H, Kiso S, Yoshida Y, Shinzaki S, et al. N-Acetylglucosaminyltransferase v regulates TGF-β response in hepatic stellate cells and the progression of steatohepatitis. Glycobiology. 2012;22:778–87.

116. Abdel Rahman AM, Ryczko M, Nakano M, Pawling J, Rodrigues T, Johswich A, et al. Golgi N-glycan branching N-acetylglucosaminyltransferases I, v and VI

promote nutrient uptake and metabolism. Glycobiology. 2015;25:225–40.

117. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry : an emerging technology for microbial identification and diagnosis. Front Microbiol. 2015;6:1–16.

118. Caprioli RM, Farmer TB, Gile J. Molecular Imaging of Biological Samples : Localization of Peptides and Proteins Using MALDI-TOF MS. Anal Chem. 1997;69:4751–60.

119. Smolira A, Wessely-szponder J. Importance of the Matrix and the Matrix / Sample Ratio in MALDI-TOF-MS Analysis of Cathelicidins Obtained from Porcine Neutrophils. Appl Biochem Biotechnol. 2015;175:2050–65.

120. Leopold J, Popkova Y, Engel KM, Schiller J. Recent Developments of Useful MALDI Matrices for the Mass Spectrometric Characterization of Lipids. Biomolecules. 2018;173.

121. Duncan M, Demarco ML. Clinical Mass Spectrometry MALDI-MS : Emerging roles in pathology and laboratory medicine. Clin Mass Spectrom [Internet]. The Association for Mass Spectrometry: Applications to the Clinical Lab (MSACL); 2019;13:1–4. Available from: https://doi.org/10.1016/j.clinms.2019.05.003

122. Morelle W, Faid V, Chirat F, Michalski J-C. Analysis of N- and O-Linked Glycans from Glycoproteins Using MALDI-TOF Mass Spectrometry. In: Packer NH, Karlsson NG, editors. Glycomics. Methods in. Humana Press; page 5–21.

123. Seeley EH, Caprioli RM. 3D imaging by mass spectrometry: A new frontier. Anal Chem. 2012;84:2105–10.

124. Powers TW, Jones EE, Betesh LR, Romano PR, Gao P, Copland JA, et al. Matrix assisted laser desorption ionization imaging mass spectrometry workflow for spatial profiling analysis of N-linked Glycan expression in tissues. Anal Chem. 2013;85:9799–806.

125. Black AP, Liang H, West CA, Wang M, Herrera HP, Haab BB, et al. A novel mass spectrometry platform for multiplexed N-glycoprotein biomarker discovery from patient biofluids by antibody panel based N-glycan imaging. Anal Chem. American Chemical Society; 2019;91:8429–35.

126. Blaschke CRK, Black AP, Mehta AS, Angel PM, Drake RR. Rapid N-Glycan Profiling of Serum and Plasma by a Novel Slide-Based Imaging Mass Spectrometry Workflow. J Am Soc Mass Spectrom. 2020;31:2511–20.

127. Drake RR, Powers TW, Norris-Caneda K, Mehta AS, Angel PM. In Situ Imaging of N-Glycans by MALDI Imaging Mass Spectrometry of Fresh or Formalin-Fixed Paraffin-Embedded Tissue. Curr Protoc Protein Sci. 2018;94:1–21.

128. West CA, Liang H, Drake RR, Mehta AS. New Enzymatic Approach to Distinguish Fucosylation Isomers of N-Linked Glycans in Tissues Using MALDI Imaging Mass Spectrometry. J Proteome Res. 2020;19:2989–96.

129. Drake RR, Powers TW, Jones EE, Bruner E, Mehta AS, Angel PM. MALDI Mass Spectrometry Imaging of N-Linked Glycans in Cancer Tissues [Internet]. 1st ed. Adv. Cancer Res. Elsevier Inc.; 2017. Available from: http://dx.doi.org/10.1016/bs.acr.2016.11.009

130. Bryda EC. The Mighty Mouse: The impact of Rodents on Advances in Biomedical Research. Mo Med. 2013;110:207–11.

131. Mukherjee P, Roy S, Ghosh D, Nandi SK. Role of animal models in biomedical research : a review. Lab Anim Res [Internet]. BioMed Central; 2022;38:1–17. Available from: https://doi.org/10.1186/s42826-022-00128-1

132. Rogers AB, Dintzis RZ. Liver and Gallbladder [Internet]. First Edit. Comp. Anat. Histol. Elsevier Inc.; 2012. Available from: http://dx.doi.org/10.1016/B978-0-12-381361-9.00013-5

133. Kruepunga N, Hakvoort TBM, Hikspoors JPJM, S KE, H LW. Anatomy of rodent and human livers : What are the di ff erences. BBA-Molecular Basis Dis. 2019;1865:869–78.

134. Antonopoulos A, North SJ, Haslam SM, Dell A. Glycosylation of mouse and human immune cells: insights emerging from N-glycomics analyses. Biochem Soc Trans. 2011;39:1334–40.

135. Cabral J, Hanley SA, Gerlach JQ, Leary NO, Cunningham S, Ritter T, et al. Distinctive surface glycosylation Patterns associated With Mouse and human cD4 + regulatory T cells and Their suppressive Function. Front Immunol. 2017;8.

136. Lee J, Ha S, Kim M, Kim S, Yun J, Ozcan S, et al. Spatial and temporal diversity of glycome expression in mammalian brain. PNAS. 2020;117:28743–53.

137. Stanley P. What have we learned from glycosyltransferase knockouts in mice? J Mol Biol. 2016;428:3166–82.

138. Cheung P, Pawling J, Partridge EA, Sukhu B, Grynpas M, Dennis JW. Metabolic homeostasis and tissue renewal are dependent on b 1 , 6GlcNAc-branched N - glycans. Glycobiology. 2007;17:828–37.

139. Fukuda T, Hashimoto H, Okayasu N, Kameyama A, Onogi H, Nakagawasai O, et al. α1,6-fucosyltransferase-deficient mice exhibit multiple behavioral abnormalities associated with a schizophrenia-like phenotype: Importance of the balance between the dopamine and serotonin systems. J Biol Chem. 2011;286:18434–43.

140. Ellies LG, Ditto D, Levy GG, Wahrenbrock M, Ginsburg D, Varki A, et al. Sialyltransferase ST3Gal-IV operates as a dominant modifier of hemostasis by concealing asialoglycoprotein receptor ligands. PNAS. 2002;99:10042–7.

141. Imai-Nishiya H, Mori K, Inoue M, Wakitani M, Iida S, Shitara K, et al. Double knockdown of α1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: A new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADCC. BMC

Biotechnol. 2007;7:1–13.

142. Nairn A V, York WS, Harris K, Hall EM, Pierce JM, Moremen KW. Regulation of Glycan Structures in Animal Tissues. 2008;283:17298–313.

143. Varki A. Loss of N-Glycolylneuraminic Acid in Humans : Mechanisms , Consequences , and Implications for. Yearb Phys Anthropol. 2001;44:54–69.

144. Hancock SM, Vaughan MD, Withers SG. Engineering of glycosidases and glycosyltransferases. Curr Opin Chem Biol. 2006;10:509–19.

145. Kim H, Kim M, Im S-K, Fang S. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. Lab Anim Res. 2018;34:147.

146. Kersten K, Visser KE, Miltenburg MH, Jonkers J. Genetically engineered mouse models in oncology research and cancer medicine. EMBO Mol Med. 2017;9:137–53.

147. Bastian K, Scott E, Elliott DJ, Munkley J. FUT8 Alpha- (1 , 6) -Fucosyltransferase in Cancer. Int J Mol Med Sci. 2021;1–22.

148. Anstee QM, Reeves HL, Kotsiliti E, Govaere O, Heikenwalder M. From NASH to HCC: current concepts and future challenges. Nat Rev Gastroenterol Hepatol. 2019;16:411–28.

149. Peng C, Stewart AG, Woodman OL, Ritchie RH, Qin Xue C. Non-Alcoholic Steatohepatitis : A Review of Its Mechanism, Models and Medical Treatments. Front Pharmacol. 2020;11.

150. De Lorenzo S, Tovoli F, Mazzotta A, Vasuri F, Edeline J, Malvi D, et al. Non-alcoholic steatohepatitis as a risk factor for intrahepatic cholangiocarcinoma and its prognostic role. Cancers (Basel). 2020;12:1–14.

151. AOYAGI Y, SUZUKI Y, ISEMURA M, NOMOTO M, SEKINE Chu, IGARASHI K, et al. Alpha-fetoprotein and its fucosylation index in the early diagnosis of hepatocellular carcinoma. Cancer. 1988;61:769–74.

152. Johnson PJ, Poon TCW, Hjelm NM, Ho CS, Ho SKW, Welby C, et al. Glycan composition of serum alpha-fetoprotein in patients with hepatocellular carcinoma and non-seminomatous germ cell tumour. Br J Cancer. 1999;81:1188–95.

153. Mehta A, Herrera H, Block T. Glycosylationa and Liver Cancer. Physiol Behav. 2014;63:1–18.

154. Zhu J, Lin Z, Wu J, Yin H, Dai J, Feng Z, et al. Analysis of serum haptoglobin fucosylation in hepatocellular carcinoma and liver cirrhosis of different etiologies. J Proteome Res. 2014;13:2986–97.

155. Kimawaha P, Thanan R, Jusakul A, Jamnongkan W, Silsirivanit A, Sa-Ngaimwibool P, et al. Serum α2,6-sialylated glycoform of serotransferrin as a glycobiomarker for diagnosis and prediction of clinical severity in

cholangiocarcinoma. Clin Chim Acta [Internet]. Elsevier B.V.; 2022;536:142–54. Available from: https://doi.org/10.1016/j.cca.2022.09.012

156. Kinoshita M, Kubo S, Tanaka S, Shigekazu T, Nishioka T, Hamano G, et al. The Association Between Non-Alcoholic Steatohepatitis and Intrahepatic Cholangiocarcinoma: A Hospital Based Case-Control Study. J Surg Oncol. 2016;779–83.

157. Welzel TM, Graubard BI, El-serag HB, Shaib YH, Ann W, Davila JA, et al. Risk factors for intra- and extrahepatic cholangiocarcinoma in the United States: a population based case-control study. Clin Gastroenterol Hepatol. 2008;5:1221–8.

158. Nie H, Liu X, Zhang Y, Li T, Zhan C, Huo W, et al. Specific N-glycans of Hepatocellular Carcinoma Cell Surface and the Abnormal Increase of Core-α-1, 6-fucosylated Triantennary Glycan via N-acetylglucosaminyltransferases-IVa Regulation. Sci Rep. 2015;5:1–11.

159. Noda K, Miyoshi E, Uozumi N, Yanagidani S, Ikeda Y, Gao CX, et al. Gene expression of α1-6 fucosyltransferase in human hepatoma tissues: A possible implication for increased fucosylation of α-fetoprotein. Hepatology. 1998;28:944–52.

160. Wang X, Inoue S, Gu J, Miyoshi E, Noda K, Li W, et al. Dysregulation of TGF-β1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. Proc Natl Acad Sci U S A. 2005;102:15791–6.

161. Surwit RS, Kuhn CM, Cochrane C, McCubbin JA, Feinglos MN. Diet-induced type II diabetes in C57BL/6J mice. Diabetes. 1988;37:1163–7.

162. Farrell G, Schattenberg JM, Leclercq I, Yeh MM, Goldin R, Teoh N, et al. Mouse Models of Nonalcoholic Steatohepatitis: Toward Optimization of Their Relevance to Human Nonalcoholic Steatohepatitis. Hepatology. 2019;69:2241–57.

163. Emilien Loeuillard, Fischbach SR, Gregory J. Gores, Rizvi S. Animal Models of Cholangiocarcinoma. Biochim Biophys Acta Mol Basis Dis. 2018;982–92.

164. Crespo Yanguas S, Cogliati B, Willebrords J, Maes M, Colle I, Van den Bossche B, et al. Experimental models of liver fibrosis [Internet]. J. Hepatol. 2016. Available from: http://dx.doi.org/10.1016/j.cgh.2014.09.046%0Ahttp://dx.doi.org/10.1038/nm.2807%0Ahttps://doi.org/10.1016/j.addr.2017.05.007%0Ahttp://dx.doi.org/10.1016/j.matbio.2015.01.004%0Ahttp://dx.doi.org/10.1016/j.jhep.2016.04.023%0Ahttp://dx.doi.org/10.1038/nrgast

165. Hansen HH, Feigh M, Veidal SS, Rigbolt KT, Vrang N, Fosgerau K. Mouse models of nonalcoholic steatohepatitis in preclinical drug development. Drug Discov Today [Internet]. Elsevier Ltd; 2017;22:1707–18. Available from: https://doi.org/10.1016/j.drudis.2017.06.007

166. Fujii M, Shibazaki Y, Wakamatsu K, Honda Y, Kawauchi Y, Suzuki K, et al. A murine model for non-alcoholic steatohepatitis showing evidence of association between diabetes and hepatocellular carcinoma. Med Mol Morphol. 2013;46:141–52.

167. Tsuchida T, Lee YA, Fujiwara N, Maria Y, Allen B, Sebastiao M, et al. A Simple Diet- and Chemical-Induced Murine NASH Model with Rapid Progression of Steatohepatitis, Fibrosis and Liver Cancer. J Hepatol. 2018;176:139–48.

168. Nevzorova YA, Boyer-Diaz Z, Cubero FJ, Gracia-Sancho J. Animal models for liver disease – A practical approach for translational research. J Hepatol [Internet]. Elsevier B.V; 2020;73:423–40. Available from: https://doi.org/10.1016/j.jhep.2020.04.011

169. Bakiri L, Wagner EF. Mouse models for liver cancer. Mol Oncol. 2013;7:206–23.

170. Brown ZJ, Heinrich B, Greten TF. Mouse models of hepatocellular carcinoma: an overview and highlights for immunotherapy research. Nat Rev Gastroenterol Hepatol [Internet]. Springer US; 2018;15:536–54. Available from: http://dx.doi.org/10.1038/s41575-018-0033-6

171. Jilkova ZM, Kurma K, Decaens T. Animal Models of Hepatocellular Carcinoma : The Role of Immune System. Cancers (Basel). 2019;11:1–12.

172. Leenders MWH, Nijkamp MW, Borel Rinkes IHM. Mouse models in liver cancer research: A review of current literature. World J Gastroenterol. 2008;14:6915–23.

173. Charlton M, Krishnan A, Viker K, Sanderson S, Cazanave S, McConico A, et al. Fast food diet mouse: Novel small animal model of NASH with ballooning, progressive fibrosis, and high physiological fidelity to the human condition. Am J Physiol - Gastrointest Liver Physiol. 2011;301.

174. Sellmann C, Priebs J, Landmann M, Degen C, Engstler AJ, Jin CJ, et al. Diets rich in fructose, fat or fructose and fat alter intestinal barrier function and lead to the development of nonalcoholic fatty liver disease over time. J Nutr Biochem [Internet]. Elsevier Inc.; 2015;26:1183–92. Available from: http://dx.doi.org/10.1016/j.jnutbio.2015.05.011

175. Kohli R, Kirby M, Xanthakos SA, Softic S, Feldstein AE, Saxena V, et al. High-Fructose Medium-Chain-Trans-Fat Diet Induces Liver Fibrosis & Elevates Plasma Coenzyme Q9 in a Novel Murine Model of Obesity and NASH. Hepatology. 2010;52:934–44.

176. Clapper JR, Hendricks MD, Gu G, Wittmer C, Dolman CS, Herich J, et al. Diet-induced mouse model of fatty liver disease and nonalcoholic steatohepatitis reflecting clinical disease progression and methods of assessment. Am J Physiol - Gastrointest Liver Physiol. 2013;305:483–95.

177. Febbraio MA, Reibe S, Shalapour S, Ooi GJ, Watt MJ, Karin M. Preclinical Models for Studying NASH-Driven HCC: How Useful Are They? Cell Metab

[Internet]. Elsevier Inc.; 2019;29:18–26. Available from: https://doi.org/10.1016/j.cmet.2018.10.012

178. Asgharpour A, Cazanave SC, Pacana T, Seneshaw M, Vincent R, Banini BA, et al. A diet-induced animal model of non-alcoholic fatty liver disease and hepatocellular cancer. 2016;65:579–88.

179. Jha P, Knopf A, Koefeler H, Mueller M, Lackner C, Hoefler G, et al. Role of adipose tissue in methionine-choline-deficient model of non-alcoholic steatohepatitis (NASH). Biochim Biophys Acta - Mol Basis Dis. 2014;1842:959–70.

180. Klair JS, Yang JD, Abdelmalek MF, Guy CD, Gill RM, Yates K, et al. A longer duration of estrogen deficiency increases fibrosis risk among postmenopausal women with nonalcoholic fatty liver disease. Hepatology. 2016;64:85–91.

181. Lonardo A, Nascimbeni F, Ballestri S, Fairweather D, Win S, Than TA, et al. Sex Differences in NAFLD: State of the Art and Identification of Research Gaps. Hepatology [Internet]. 2019;1457–69. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6766425/pdf/nihms-1020123.pdf

182. Balakrishnan M, Patel P, Dunn-valadez S, Dao C, Khan V, Ali H, et al. Women have Lower Risk of Nonalcoholic Fatty Liver Disease but Higher Risk of Progression vs Men: A Systematic Review and Meta-analysis. 2022;19:61–71.

183. Van Herck MA, Vonghia L, Francque SM. Animal models of nonalcoholic fatty liver disease—a starter's guide. Nutrients. 2017;9:1–13.

184. Kubota N, Kado S, Kano M, Masuoka N, Nagata Y, Kobayashi T, et al. A high-fat diet and multiple administration of carbon tetrachloride induces liver injury and pathological features associated with non-alcoholic steatohepatitis in mice. Clin Exp Pharmacol Physiol. 2013;40:422–30.

185. Scholten D, Trebicka J, Liedtke C, Weiskirchen R. The carbon tetrachloride model in mice. Lab Anim. 2015;49:4–11.

186. Fujii T, Fuchs BC, Yamada S, Lauwers GY, Kulu Y, Goodwin JM, et al. Mouse model of carbon tetrachloride induced liver fibrosis : Histopathological changes and expression of CD133 and epidermal growth factor. BMC Gastroenterol [Internet]. 2010;10. Available from: http://www.biomedcentral.com/1471-230X/10/79%0APage

187. Graham ML, Janecek JL, Kittredge JA, Hering BJ, Schuurman HJ. The streptozotocin-induced diabetic nude mouse model: Differences between animals from different sources. Comp Med. 2011;61:356–60.

188. Schulien I, Hasselblatt P. Diethylnitrosamine-induced liver tumorigenesis in mice [Internet]. 1st ed. Methods Cell Biol. Elsevier Inc.; 2021. Available from: http://dx.doi.org/10.1016/bs.mcb.2020.08.006

189. Qi Y, Chen X, Chan CY, Li D, Yuan C, Yu F, et al. Two-dimensional differential

gel electrophoresis/analysis of diethylnitrosamine induced rat hepatocellular carcinoma. Int J Cancer. 2008;122:2682–8.

190. Schulien I, Hockenjos B, Van Marck V, Ayata CK, Follo M, Thimme R, et al. Extracellular ATP and purinergic P2Y2 receptor signaling promote liver tumorigenesis in mice by exacerbating DNA damage. Cancer Res. 2020;80:699–708.

191. Hoshida Y, Nijman SMB, Kobayashi M, Chan JA, Brunet J, Chiang DY, et al. Integrative Transcriptome Analysis Reveals Common Molecular Subclasses of Human Hepatocellular Carcinoma. Cancer Res. 2009;69:7385–92.

192. Maronpot RR. Biological basis of differential susceptibility to hepatocarcinogenesis among mouse strains. J Toxicol Pathol. 2009;22:11–33.

193. He G, Karin M. NF-κB and STAT3- key players in liver inflammation and cancer. Cell Res. Nature Publishing Group; 2011;21:159–68.

194. Park EJ, Lee JH, Yu GY, He G, Ali SR, Holzer RG, et al. Dietary and Genetic Obesity Promote Liver Inflammation and Tumorigenesis by Enhancing IL-6 and TNF Expression. Cell. 2010;140:197–208.

195. Naugler WE, Sakurai T, Kim S, Maeda S, Kim KH, Elsharkawy AM, et al. Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. Science (80-). 2007;317:121–4.

196. Li M, Zhou X, Wang W, Ji B, Shao Y, Du Q, et al. Selecting an Appropriate Experimental Animal Model for Cholangiocarcinoma Research. J Clin Transl Hepatol [Internet]. 2022;10:700–10. Available from: https://www.doi.org/10.14218/JCTH.2021.00374

197. Waxman DJ, Azaroff L. Phenobarbital induction of cytochrome P-450 gene expression. Biochem J. 1992;281:577–92.

198. He L, Tian DA, Li PY, He XX. Mouse models of liver cancer: Progress and recommendations. Oncotarget. 2015;6:23306–22.

199. Eferl R, Walker JM. Mouse Models of Cancer Series Editor. 2015.

200. Kirsch R, Clarkson V, Shephard EG, Marais DA, Jaffer MA, Woodburne VE, et al. Rodent nutritional model of non-alcoholic steatohepatitis: Species, strain and sex difference studies. J Gastroenterol Hepatol. 2003;18:1272–82.

201. Rao KVN, Vesselinovitch SD. Age- and Sex-associated Diethylnitrosamine Dealkylation Activity of the Mouse Liver and Hepatocarcinogenesis. Cancer Res. 1973;33:1625–7.

202. Brown GT, Kleiner DE. Histopathology of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Metabolism [Internet]. Elsevier B.V.; 2016;65:1080–6. Available from: http://dx.doi.org/10.1016/j.metabol.2015.11.008

203. Alkhouri N, Tincopa M, Loomba R, Harrison SA. What Does the Future Hold for Patients With Nonalcoholic Steatohepatitis: Diagnostic Strategies and Treatment Options in 2021 and Beyond? Hepatol Commun. 2021;0:1–14.

204. Blomme B, Francque S, Trépo E, Libbrecht L, Vanderschaeghe D, Verrijken A, et al. N-glycan based biomarker distinguishing non-alcoholic steatohepatitis from steatosis independently of fibrosis. Dig Liver Dis [Internet]. Editrice Gastroenterologica Italiana; 2012;44:315–22. Available from: http://dx.doi.org/10.1016/j.dld.2011.10.015

205. Verhelst X, Dias AM, Colombel JF, Vermeire S, Van Vlierberghe H, Callewaert N, et al. Protein Glycosylation as a Diagnostic and Prognostic Marker of Chronic Inflammatory Gastrointestinal and Liver Diseases. Gastroenterology [Internet]. Elsevier, Inc; 2020;158:95–110. Available from: https://doi.org/10.1053/j.gastro.2019.08.060

206. Joo Ann H, Froehlich J, Lebrilla CB. Determination of Glycosylation Sites and Site-specific Heterogeneity in Glycoproteins. Russ Chem Bull. 2009;13:421–6.

207. Dang L, Jia L, Zhi Y, Li P, Zhao T, Zhu B, et al. Mapping human N-linked glycoproteins and glycosylation sites using mass spectrometry. TrAC - Trends Anal Chem [Internet]. Elsevier Ltd; 2019;114:143–50. Available from: https://doi.org/10.1016/j.trac.2019.02.009

208. Thomas D, Rathinavel AK, Radhakrishnan P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. Biochim Biophys Acta - Rev Cancer [Internet]. Elsevier B.V.; 2020;1875:188464. Available from: https://doi.org/10.1016/j.bbcan.2020.188464

209. Cuello HA, Ferreira GM, Gulino CA, Toledo AG, Segatori VI, Gabri MR. Terminally sialylated and fucosylated complex N-glycans are involved in the malignant behavior of high-grade glioma. Oncotarget. 2020;11:4822–35.

210. Kamada Y, Akita M, Takeda Y, Yamada S, Fujii H, Sawai Y, et al. Serum Fucosylated Haptoglobin as a Novel Diagnostic Biomarker for Predicting Hepatocyte Ballooning and Nonalcoholic Steatohepatitis. PLoS One. 2013;8.

211. Basaranoglu M, Basaranoglu G, Bugianesi E. Carbohydrate intake and nonalcoholic fatty liver disease: fructose as a weapon of mass destruction. Hepatobiliary Surg Nutr. 2015;4:109–16.

212. Jegatheesan P, De Bandt JP. Fructose and NAFLD: The multifaceted aspects of fructose metabolism. Nutrients. 2017;9:1–13.

213. Spruss A, Kanuri G, Wagnerberger S, Haub S, Bischoff SC, Bergheim I. Toll-like receptor 4 is involved in the development of fructose-induced hepatic steatosis in mice. Hepatology. 2009;50:1094–104.

214. Tiniakos D, Vos M, Brunt EM. Nonalcoholic Fatty Liver Disease: Pathology and Pathogenesis. Annu Rev Pathol Mech Dis [Internet]. 2010;5:145–71. Available

from: https://doi.org/10.1146/annurev-pathol-121808-102132

215. Shi Z, Rockey DC. Upregulation of actin cytoskeleton via myocardin leads to increased expression of type 1 collagen. Physiol Behav. 2017;176:139–48.

216. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with imageJ. Biophotonics Int. 2004;11:36–41.

217. Drake RR, Powers TW, Jones EE, Bruner E, Mehta AS, Angel PM. MALDI Mass Spectrometry Imaging of N-Linked Glycans in Cancer Tissues. Adv Cancer Res. 2017;134:85–116.

218. Angel PM, Mehta A, Norris-Caneda K, Richard R. Drake. MALDI Imaging Mass Spectrometry of N-glycans and Tryptic Peptides from the Same Formalin-Fixed, Paraffin-Embedded Tissue Section. Physiol Behav. 2016;176:139–48.

219. McDowell CT, Klamer Z, Hall J, West CA, Wisniewski L, Powers TW, et al. Imaging mass spectrometry and lectin analysis of n-linked glycans in carbohydrate antigen-defined pancreatic cancer tissues. Mol Cell Proteomics [Internet]. The Authors; 2021;20:100012. Available from: https://doi.org/10.1074/mcp.RA120.002256

220. Neuschwander-Tetri BA, Ford DA, Sahaja A, Gilkey G, Basaranoglu M, Tetri L, et al. Dietary trans-Fatty Acid Induced NASH is Normalized Following Loss of trans-Fatty Acids from Hepatic Lipid Pools. Lipids. 2012;47:941–50.

221. Rogers AB, Dintzis RZ. Hepatobiliary System [Internet]. Comp. Anat. Histol. Elsevier Inc.; 2018. Available from: http://dx.doi.org/10.1016/B978-0-12-802900-8.00013-0

222. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005;41:1313–21.

223. Schwenger KJP, Chen L, Chelliah AE, Da Silva HE, Teterina A, Comelli EM, et al. Markers of activated inflammatory cells are associated with disease severity and intestinal microbiota in adults with non-alcoholic fatty liver disease. Int J Mol Med. 2018;42:1857–64.

224. Rudman N, Gornik O, Lauc G. Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes. FEBS Lett. 2019;593:1598–615.

225. Kim T, Xie Y, Li Q, Artegoitia VM, Lebrilla CB, Keim NL, et al. Diet affects glycosylation of serum proteins in women at risk for cardiometabolic disease. Eur J Nutr [Internet]. Springer Berlin Heidelberg; 2021; Available from: https://doi.org/10.1007/s00394-021-02539-7

226. Mehta A, Comunale MA, Rawat S, Casciano JC, Lamontagne J, Herrera H, et al. Intrinsic hepatocyte dedifferentiation is accompanied by upregulation of mesenchymal markers, protein sialylation and core alpha 1,6 linked fucosylation. Sci Rep [Internet]. Nature Publishing Group; 2016;6:1–14. Available from:

http://dx.doi.org/10.1038/srep27965

227. Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. Cell surface protein glycosylation in cancer. Proteomics. 2014;14:525–46.

228. Drake RR, Jones EE, Powers TW, Nyalwidhe JO. Altered glycosylation in prostate cancer [Internet]. 1st ed. Adv. Cancer Res. Elsevier Inc.; 2015. Available from: http://dx.doi.org/10.1016/bs.acr.2014.12.001

229. Li D, Mallory T, Satomura S. AFP-L3: A new generation of tumor marker for hepatocellular carcinoma. Clin Chim Acta. 2001;313:15–9.

230. Nakagawa T, Uozumi N, Nakano M, Mizuno-Horikawa Y, Okuyama N, Taguchi T, et al. Fucosylation of N-glycans regulates the secretion of hepatic glycoproteins into bile ducts. J Biol Chem [Internet]. Â© 2006 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology.; 2006;281:29797–806. Available from: http://dx.doi.org/10.1074/jbc.M605697200

231. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology. 2015;149:389-397.e10.

232. Katherine T. Brunner, Henneberg CJ, Wilechansky RM, Long MT. Nonalcoholic Fatty Liver Disease and Obesity Treatment. Curr Obes Rep. 2019;8:220–8.

233. Blechacz B, Komuta M, Roskams T, Gores GJ. Clinical diagnosis and staging of cholangiocarcinoma. Nat Rev Gastroenterol Hepatol. 2011;8:512–22.

234. Cheemerla S, Balakrishnan M. Global Epidemiology of Chronic Liver Disease. 2020;17:365–70.

235. Liang W, Menke AL, Driessen A, Koek GH, Lindeman JH, Stoop R, et al. Establishment of a general NAFLD scoring system for rodent models and comparison to human liver pathology. PLoS One. 2014;9:1–17.

236. Powers TW, Jones EE, Betesh LR, Copland A, Mehta AS, Drake RR. A MALDI Imaging Mass Spectrometry Workflow for Spatial Profiling Analysis of N-linked Glycan Expression in Tissues. Anal Chem [Internet]. 2013;85:9799–806. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3969840/

237. Kishida N, Matsuda S, Itano O, Shinoda M, Kitago M, Yagi H, et al. Development of a novel mouse model of hepatocellular carcinoma with nonalcoholic steatohepatitis using a high-fat, choline-deficient diet and intraperitoneal injection of diethylnitrosamine. BMC Gastroenterol [Internet]. BMC Gastroenterology; 2016;16:1–13. Available from: http://dx.doi.org/10.1186/s12876-016-0477-5

238. Tolba R, Kraus T, Liedtke C, Schwarz M, Weiskirchen R. Diethylnitrosamine (DEN)-induced carcinogenic liver injury in mice. Lab Anim. 2015;49:59–69.

239. Basturk O, Farris AB, Adsay NV. Immunohistology of the Pancreas, Biliary Tract, and Liver. In: Dabbs DJ, editor. Diagnostic Immunohistochem [Internet]. Elsevier; 2011. page 541–92. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9781416057666000194

240. Chang TT, Cheng JH, Tsai HW, Young KC, Hsieh SY, Ho CH. Plasma proteome plus site-specific N-glycoprofiling for hepatobiliary carcinomas. J Pathol Clin Res. 2019;5:199–212.

241. Arnold JN, Saldova R, Abd Hamid UM, Rudd PM. Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. Proteomics. 2008;8:3284–93.

242. Hakomori S. Glycosylation defining cancer malignancy: New wine in an old bottle. Proc Natl Acad Sci U S A. 2002;99:10231–3.

243. Keeley TS, Yang S, Lau E. The diverse contributions of fucose linkages in cancer. Cancers (Basel). 2019;11:1–25.

244. Debruyne EN, Vanderschaeghe D, Van Vlierberghe H, Vanhecke A, Callewaert N, Delanghe JR. Diagnostic value of the hemopexin n-glycan profile in hepatocellular carcinoma patients. Clin Chem. 2010;56:823–31.

245. Zhang Y, Zhu J, Yin H, Marrero J, Zhang XX, Lubman DM. ESI-LC-MS Method for Haptoglobin Fucosylation Analysis in Hepatocellular Carcinoma and Liver Cirrhosis. J Proteome Res. 2015;14:5388–95.

246. Zhu J, Wu J, Yin H, Marrero J, Lubman DM. Mass spectrometric N-glycan analysis of haptoglobin from patient serum samples using a 96-well plate format. J Proteome Res. 2015;14:4932–9.

247. Yin H, Lin Z, Nie S, Wu J, Tan Z, Zhu J, et al. Mass-selected site-specific core-fucosylation of ceruloplasmin in alcohol-related hepatocellular carcinoma. J Proteome Res. 2014;13:2887–96.

248. Blaschke CRK, McDowell CT, Black AP, Mehta AS, Angel PM, Drake RR. Glycan Imaging Mass Spectrometry: Progress in Developing Clinical Diagnostic Assays for Tissues, Biofluids, and Cells. Clin Lab Med. 2021;41:247–66.

249. Kuhn M, Johnson K. Applied Predictive Modeling. 2013.

250. de Vienne DM. Tanglegrams are misleading for visual evaluation of tree congruencele. HAL open Sci [Internet]. 2018; Available from: https://hal.archives-ouvertes.fr/hal-01807411

251. Borcard D, Gillet F, Legendre P. Numerical Ecology with R. 2018.

252. Block TM, Comunale MA, Lowman M, Steel LF, Romano PR, Fimmel C, et al. Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. Proc Natl Acad Sci U S A. 2005;102:779–84.

253. Pj J. Structures of disease-specific serum alpha-fetoprotein isoforms. 2000;83:1330–7.

254. Huang C, Xu X, Wang M, Xiao X, Cheng C, Ji J, et al. Serum N-glycan fingerprint helps to discriminate intrahepatic cholangiocarcinoma from hepatocellular carcinoma. Electrophoresis. 2021;1–9.

255. Tu CF, Wu MY, Lin YC, Kannagi R, Yang RB. FUT8 promotes breast cancer cell invasiveness by remodeling TGF-β receptor core fucosylation. Breast Cancer Res. Breast Cancer Research; 2017;19:1–15.

256. Calderon AD, Li L, Wang PG. FUT8: From biochemistry to synthesis of core-fucosylated N-glycans. Pure Appl Chem. 2017;89:911–20.

257. López-Cortés R, Muinelo-Romay L, Fernandez-Briera A, Gil-martin E. Inhibition of α (1 , 6) fucosyltransferase : Effects on Cell Proliferation , Migration , and Adhesion in an SW480 / SW620 Syngeneic Colorectal Cancer Model. Int J Mol Sci. 2022;

258. Wang X, Chen J, Li QK, Peskoe SB, Zhang B, Choi C, et al. Overexpression of α (1,6) fucosyltransferase associated with aggressive prostate cancer. Glycobiology. 2014;24:935–44.

259. Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, et al. A comparative encyclopedia of DNA elements in the mouse genome. Nature. Nature Publishing Group; 2014;515:355–64.

260. Moman RN, Gupta2 N, Varacallo M. Physiology, Albumin. reasure Island (FL); 2022.

261. Weisend CM, Kundert JA, Suvorova ES, Prigge JR, E. ES. CRE ACTIVITY IN FETAL albCre MOUSE HEPATOCYTES: UTILITY FOR DEVELOPMENTAL STUDIES. Genesis [Internet]. 2009;47:789–92. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2828742/pdf/nihms-178705.pdf

262. Kellendonk C, Opherk C, Anlag K, Schütz G, Tronche F. Hepatocyte-specific expression of Cre recombinase. Genesis. 2000;26:151–3.

263. Postic C, Magnuson MA. DNA excision in liver by an albumin-Cre transgene occurs progressively with age. Genesis. 2000;26:149–50.

264. Zhu J, Huang J, Zhang J, Chen Z, Lin Y, Grigorean G, et al. Glycopeptide Biomarkers in Serum Haptoglobin for Hepatocellular Carcinoma Detection in Patients with Nonalcoholic Steatohepatitis. J Proteome Res. 2020;19:3452–66.

265. Zhao T, Jia L, Li J, Ma C, Wu J, Shen J, et al. Heterogeneities of Site-Specific N-Glycosylation in HCC Tumors With Low and High AFP Concentrations. Front Oncol. 2020;10:1–12.

266. Johnson PJ, Pirrie SJ, Cox TF, Berhane S, Teng M, Palmer D, et al. The detection of hepatocellular carcinoma using a prospectively developed and

validated model based on serological biomarkers. Cancer Epidemiol Biomarkers Prev. 2014;23:144–53.

267. Liang C, Fukuda T, Isaji T, Duan C, Song W, Wang Y, et al. a1 , 6-Fucosyltransferase contributes to cell migration and proliferation as well as to cancer stemness features in pancreatic carcinoma. BBA - Gen Subj [Internet]. Elsevier B.V.; 2021;1–10. Available from: https://doi.org/10.1016/j.bbagen.2021.129870

268. Miyoshi E, Noda K, Ko JH, Ekuni A, Kitada T, Uozumi N, et al. Overexpression of α1-6 fucosyltransferase in hepatoma cells suppresses intrahepatic metastasis after splenic injection in athymic mice. Cancer Res. 1999;59:2237–43.

269. Kuang M, Wu H, Hu L, Guo X, He D, Liu B, et al. Up - regulation of FUT8 inhibits TGF- β 1-induced activation of hepatic stellate cells during liver fibrogenesis. Glycoconjugate Journal; 2021;

270. Li F, Zhao S, Cui Y, Guo T, Qiang J, Xie Q, et al. cancer-promoting capacity of cancer-associated fibroblasts (CAFs) by modifying EGFR core fucosylation (CF) in non-small cell lung cancer (NSCLC). 2020;10:816–37.