January 2019

A Genomic Approach To Idiopathic Liver Disease In Adults

Aaron Hakim

Follow this and additional works at: https://elischolar.library.yale.edu/ymtdl

Recommended Citation
https://elischolar.library.yale.edu/ymtdl/3501

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
A GENOMIC APPROACH TO IDIOPATHIC LIVER DISEASE IN ADULTS

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

by

Aaron Hakim

2019
ABSTRACT

Adult patients suffering from liver disease of unknown cause represent an understudied and underserved population. Over the past 15 years, next-generation sequencing technologies have matured into an inexpensive, effective, and widely available set of tools to do genomic analysis. One of these technologies, whole-exome sequencing (WES), allows for high throughput sequencing of all of the genome’s protein coding regions (exons). In pediatric cohorts, WES combined with deep clinical phenotyping has been shown to be an effective and unbiased method of identifying rare protein-altering coding variants in individual genes. WES has also contributed to the diagnosis and individualization of medical care in oncologic patients. The use of WES for the study of a broader spectrum of non-oncological diseases, among adults, remains poorly understood. We assessed the utility of WES in the diagnosis and management of adults with unexplained liver disease despite a comprehensive conventional workup and with no history of alcohol overuse.

We performed exome sequencing and deep phenotyping in two independent adult cohorts with unexplained liver disease. In the first cohort, we analyzed nineteen unrelated adult patients with idiopathic liver disease recruited at Yale New Haven Hospital. In a second cohort from Bridgeport Hospital, four unrelated adult patients presenting with fatty liver disease, hypertriglyceridemia, insulin resistance, and physical exam findings suggestive of lipodystrophy were recruited for genomic analysis.
In cohort 1, analysis of the exome in nineteen cases identified four monogenic disorders in five unrelated adults. Patient 1 suffered for 18 years from devastating complications of undiagnosed Type 3 Familial Partial Lipodystrophy due to a deleterious heterozygous variant in PPARG. Molecular diagnosis enabled initiation of leptin replacement therapy with subsequent normalization of liver transaminases, and amelioration of dyslipidemia. Patients 2 and 3 were diagnosed with MDR3 deficiency (also known as PFIC3, progressive intrahepatic familial cholestasis type 3) due to recessive mutations in ABCB4. Patient 4 with a prior diagnosis of non-alcoholic steatohepatitis was found to harbor a mitochondrial disorder due to a homozygous pathogenic variant in NDUFB3; subsequent muscle biopsy revealed a deficiency of rotenone sensitive I+III activity consistent with a mitochondrial disorder. This finding enabled initiation of disease-preventative measures including supplementation with antioxidants. Patient 5 is a lean patient with hepatic steatosis of unknown etiology who was found to have a damaging heterozygous variant in APOB, consistent with familial hypobetalipoproteinemia. In cohort 2, we identified a potential genetic diagnosis in all four cases of suspected lipodystrophy, including a patient with an LMNA mutation, a patient with two pathogenic heterozygous mutations in APOE, a patient with a homozygous deleterious mutation in the leptin receptor (LEPR), and a patient with a pathogenic heterozygous variant in PPARG.

In conclusion, WES provided a diagnosis with impact on clinical management in a significant number of adults suffering from liver disease of unknown cause, gaining insight into disease pathogenesis and identifying new therapeutic and
preventive medicine interventions. This study supports the use of WES in the evaluation and management of adults with idiopathic liver disease in clinical practice.
Published in part:


Presented in part:

ACKNOWLEDGEMENTS

The work presented in this thesis is a direct result of the incredible support and phenomenal mentorship of my supervisor, Dr. Vilarinho. She has imparted a true passion for bench to bedside translational research. I would also like to thank all the patients and their families whose contribution to this study led to advancing our understanding of liver disease, Dr. Michael Nathanson and Dr. Sachin K. Majumdar for their efforts to refer patients to this study, and the staff of the Yale Center for Genome Analysis. I am also indebted to my family for their unending love and support.

Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award Number K08DK113109. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was also supported in part by Yale Liver Center P30DK034989, and AASLD Sheila Sherlock Clinical and Translational Research Award in Liver Disease (to S.V.).
TABLE OF CONTENTS

TABLE OF CONTENTS……………………………………………………………………………………..i
LIST OF FIGURES…………………………………………………………………………………………1
LIST OF TABLES……………………………………………………………………………………………..2
INTRODUCTION………………………………………………………………………………………………3
STATEMENT OF PURPOSE………………………………………………………………………………9
PATIENTS AND METHODS………………………………………………………………………………10
RESULTS………………………………………………………………………………………………………..17
DISCUSSION…………………………………………………………………………………………………48
REFERENCES………………………………………………………………………………………………..54
LIST OF FIGURES

Figure 1: The incidence of cryptogenic cirrhosis has been steadily declining…4

Figure 2: Overview of whole-exome sequencing pipeline…………………………………7

Figure 3. Representative flowchart of genetic variant filtering strategy in this study………………………………………………………………………………13

Figure 4: Example of principal component analysis to determine ethnicity clustering………………………………………………………………………………21

Figure 5: Liver histology findings in patient 1, cohort 1……………………………………23

Figure 6: Representative plot read of disease-causing mutation identified in Patient 1, cohort 1………………………………………………………………………25

Figure 7: Genetic findings in patient 1, cohort 1……………………………………………27

Figure 8: Illustrative representation of the role of PPARG, peroxisome proliferator-activated receptor-gamma, in adipocyte differentiation and relationship to serum leptin………………………………………………………………………………28

Figure 9: Laboratory findings in patient 1, cohort 1…………………………………………29

Figure 10: Liver histology findings in patient 2, cohort 1…………………………………31

Figure 11: Genetic findings in patient 2, cohort 1…………………………………………32

Figure 12: Conservation findings in patient 2, cohort 1…………………………………33

Figure 13: Liver histology findings in patient 3, cohort 1…………………………………34

Figure 14. Liver biopsies of patient 4, cohort 1……………………………………………36

Figure 15. Liver biopsy of patient 5, cohort 1………………………………………………38

Figure 16: Schematic representation of multidisciplinary Genome Rounds in Adult Hepatology……………………………………………………………………………52
LIST OF TABLES

Table 1. Gene name, accession number, and forward and reverse primer sequences used for Sanger sequencing .................................................................15

Table 2: Summary of study population characteristics and demographics in cohort 1 (n = 19) ........................................................................................................18

Table 3: Sequencing coverage and quality metrics for patient cohort 1 (n = 19) ..................................................................................................................19

Table 4: Demographics, clinical features, and genetic diagnosis identified in five subjects in cohort 1, and its clinical implications .............................................20

Table 5: Diagnostic genetic variants identified in five subjects in cohort 1 ..........26

Table 6. Summary of outcomes for fourteen patients who remain unexplained after whole exome sequencing analysis ......................................................40

Table 7: Sequencing coverage and quality metrics for patient cohort 2 (n = 4) ..........................................................................................................................43

Table 8: Demographics, clinical features, and genetic diagnosis identified in four subjects in cohort 2, and its clinical implications .............................................43

Table 9: Diagnostic genetic variants identified in cohort 2 ..............................47
INTRODUCTION

Liver Disease of Unknown Etiology: A Historical Perspective

Liver disease is a major public health problem that affects approximately 30 million people and leads to over 40,000 deaths annually in the United States.¹ Chronic liver disease (CLD) is often silent unless there is awareness of subtle clinical signs, behavioral risk factors and/or investigation of abnormal liver function tests. Untreated liver disease may progress to end-stage liver disease (cirrhosis) and further decompensation with ascites, hepatic encephalopathy, esophageal variceal hemorrhage, jaundice, and/or hepatocellular carcinoma, leading to liver failure and death.² Advances in our understanding of liver disease have led to a marked decline in the attribution of CLD to “unknown etiology” (Figure 1). Prior to 1965, cryptogenic cirrhosis, defined as cirrhosis of unknown etiology after extensive clinical, laboratory, and histological analysis, accounted for >50% of all cases of cirrhosis.³ The discovery of hepatitis B virus (1965)⁴, hepatitis D virus (1977)⁵, and hepatitis C virus (1989)⁶ eventually led to the recognition of their contributions to cirrhosis worldwide. The description of non-alcoholic steatohepatitis as a clinical entity in 1980⁷, and improved diagnostic criteria for autoimmune hepatitis, first published in 1998⁸, further reduced the diagnosis of cryptogenic cirrhosis, as did improved diagnosis of iron overload syndromes (i.e. HFE mutation)⁹, alpha-1-antitrypsin deficiency (A1ATD)¹⁰, and Wilson’s disease¹¹ (Figure 1). However, it is currently estimated that up to 30% of cases of cirrhosis and up to 14% of adults awaiting liver transplantation suffer from liver disease of unknown etiology.¹²,¹³ In tertiary medical centers, the
incidence of cirrhosis of unknown etiology has been estimated at 5-10%. These patients often undergo a long and costly odyssey of diagnostic tests, interventions and medical opinions, and represent an understudied and underserved population. Understanding the etiology of CLD is essential to halt the progression of liver dysfunction, as illustrated by the development of a vaccine and anti-viral therapy for hepatitis B, and the highly effective, safe and curative anti-viral therapies for hepatitis C.\textsuperscript{15}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{The incidence of cryptogenic cirrhosis has been steadily declining as new etiologies are being recognized, as outlined above. A1ATD, alpha-1-antitrypsin deficiency; HBV, hepatitis B virus; PFIC, progressive familial intrahepatic cholestasis (Byler disease); NASH, non-alcoholic steatohepatitis; HCV, hepatitis C virus; HFE, human hemochromatosis protein; AIH, autoimmune hepatitis; WES, whole-exome sequencing.}
\end{figure}
Current State of Genetic Analysis in Hepatology Clinical Practice

The taxonomy of CLD in clinical practice is based broadly on categories of etiology such as exposure to toxins, viral infections, cholestatic, autoimmune, metabolic and select genetic disorders. A significant limitation of this approach is that it precludes consideration of a wider array of underlying genetic disorders masquerading within these broad phenotypes. Indeed, the current state of genetic analysis of adult liver disease in practice only involves the exclusion of a limited number of inherited conditions through single gene tests, including Wilson’s Disease (ATP7B), Hemochromatosis (HFE, HJV, HAMP, TFR2, SCL40A1), and A1ATD (SERPINA1). In some circumstances, commercial gene panel tests including the Jaundice Chip or EGL Cholestasis Panel are used, which include up to 72 genes. However, these panel tests represent only a small fraction of the ~20,000 protein coding genes of the human genome reference sequence, completed in 2003.

Whole-Exome Sequencing

Advances in human genetics and genomics have created an unprecedented opportunity for gene discovery and diagnosis in the clinic. Over the past 15 years, next-generation sequencing technologies have matured into an inexpensive, effective and widely available set of tools. One of these technologies, whole-exome sequencing (WES), consists of targeted capture and sequencing of the ~20,000 human protein-coding genes (Figure 2). Although WES excludes the 99% of the genome that does not code for proteins, it is estimated that approximately 85% of all mutations with large effects on disease-
related traits are located within exomes. Furthermore, WES has traditionally been described as having excellent sensitivity and specificity (~98-99%), especially when the mean per-base coverage is over 20x and with a minimal local read depth of 13x. Importantly, various bioinformatics programs have been optimized to translate raw WES data into manageable and intelligible datasets: performing base calling (translating the raw signals from the sequencers into A, C, T or G, with an accompanying quality score), alignment of the reads to the human reference genome (searching for the best matching segment), removal of PCR duplicates (generated during preparation of the genomic library), variant calling (recording all positions that differ from the reference genome), and variant annotation (using data from various public databases such as ClinVar, Online Mendelian Inheritance in Man [OMIM], Exome Aggregation Consortium database [ExAC], gnomAD, 1000 Genomes, National Heart, Lung and Blood Institute’s [NHLBI] Exome Variant Server, HapMap, and others to provide information about minor allele frequency, function of the gene, degree of inter-species amino acid conservation, etc). Furthermore, there are existing computational tools for *in silico* assessment of a given variant’s pathogenic potential. As such, WES currently represents a remarkable balance between cost (<$300 for a research exome without analysis, ~$3000 for a clinical exome), time of analysis, and information collected, making it attractive and suitable for clinical use and translational research studies. Nearly 3,000 genes underlying over 4,000 Mendelian phenotypes have been discovered, and next-
generation sequencing approaches including WES account for more than three times as many discoveries as conventional methods.\textsuperscript{20}

Figure 2: Overview of whole-exome sequencing pipeline. SNV, single nucleotide variant; Indel, insertion/deletion. Oligonucleotide probes designed to specifically hybridize to all exons in the genome are in solution. The probes are linked to magnetic beads. Other exome capture systems have probes attached to a microarray. Adapted from Gerald Goh and Muri Choi.\textsuperscript{21}

\textit{Diagnostic Utility of Whole-Exome Sequencing}

WES combined with deep clinical phenotyping has been increasingly applied as a first-line diagnostic tool in clinical medicine, particularly for the diagnosis of
inborn metabolic and neurodevelopmental disorders, unexplained liver failure in children\textsuperscript{22-26}, as well as for the detection of causal mutations in cancer\textsuperscript{27,28}. In these contexts, exome sequencing can inform medical management, including prognosis, choice of therapy, and accurate reproductive counselling. However, to date, most studies that investigate the use of next generation sequencing technologies in the diagnosis and individualization of medical care have been performed in either pediatric or cancer patients. There is a paucity of information on the clinical utility of these approaches for a broader spectrum of diseases among adults. A number of small studies and one study in a large cohort support the usefulness of exome sequencing for the diagnosis of early onset or familial nephropathy\textsuperscript{29-31}, sporadic chronic kidney disease\textsuperscript{32}, and inherited cardiovascular diseases\textsuperscript{33}, however to date the utility of this approach in chronic liver disease has not been elucidated. By using unbiased genomic analysis, we may begin to understand parameters of adult clinical presentations that harbor an underlying monogenic cause, and to develop a more comprehensive category of ‘genetic’ liver diseases in adults beyond the traditionally considered disorders such as Wilson’s disease, A1ATD, or hemochromatosis. Here, we provide data to support the utility of WES in the diagnosis and management of adults with liver disease of unknown cause, with or without involvement of other diseases and/or unusual clinical findings. We also extend our analysis to an independent cohort of patients with fatty liver and physical exam findings suggestive of lipodystrophy, a group of heterogeneous disorders characterized by the absence or reduction of subcutaneous adipose tissue.
STATEMENT OF PURPOSE

To assess the utility of whole-exome sequencing in the diagnosis and management of adults with unexplained liver disease.

MAIN OUTCOMES AND MEASURES

To obtain the diagnostic yield of WES and its direct impact in providing new therapeutic options, targeted preventive medicine interventions, and adequate family counselling.
PATIENTS AND METHODS

Human Subjects

Study protocol was approved by the Yale Human Investigational Committee, and informed consent was obtained in accordance with institutional review board standards. In cohort 1, nineteen adults with unexplained liver disease despite a comprehensive evaluation at Yale New Haven Hospital (unrevealing hepatitis viral serologies including negative HBsAg and anti-HBc, ferritin, iron studies, ceruloplasmin, ANA, alpha-1-antitrypsin phenotype, abdominal imaging, liver biopsy, etc) underwent further investigation using whole-exome sequencing. Patients may have had other medical co-morbidities but did not have a history of alcohol overuse. For some patients in the cohort, we questioned prior diagnosis such as non-alcoholic fatty liver disease (NAFLD) in absence of typical metabolic or body habitus features. Where possible, samples from available family members were also obtained for segregation studies. In cohort 2, we recruited four unrelated adult patients with fatty liver disease, insulin resistance/diabetes, hypertriglyceridemia, and physical exam findings suggestive of lipodystrophy per evaluation by an endocrinologist at Bridgeport Hospital.

DNA isolation, exome capture and sequencing

Genomic DNA was isolated from peripheral blood mononuclear cells or buccal swabs using standard procedures. DNA fragments contained in exonic sequences were captured and sequenced on the Illumina HiSeq platform.
Exome Sequencing Analysis

Exome sequencing data were mapped and aligned to the reference human genome (reference sequence hg19) using BWA. Variants were called using GATK\textsuperscript{34,35} and annotated using Annovar.\textsuperscript{36} Since the goal of this project was to identify genetic causes for rare Mendelian conditions, we focused on variants that are uncommon in the general population. In other words, the higher the frequency of the variant, the lower the probability to be causal of a rare disease. Variants were selected for minor allele frequency (MAF) <0.01 for homozygous and compound heterozygous variants (recessive inheritance pattern) or <2\times10^{-5} for heterozygous variants (dominant inheritance pattern). Variants with MAF >1% are unlikely to cause recessive disorders with full penetrance (prevalence 1:10,000 or less) in the general population. If autosomal dominance is the suspected pattern of inheritance, the favored MAF cutoff is more stringent because a single allele is sufficient to cause disease. Allele frequencies were determined using the genome aggregation database (gnomAD) databases,\textsuperscript{37} including the Exome Aggregation Consortium database (ExAC), 1000 Genomes, and the National Heart, Lung and Blood Institute's (NHLBI) Exome Variant Server. After filtering out common variants, variants located in intronic and intergenic segments of the genome were removed. Subsequently, protein-altering variants were selected and prioritized based on their predicted deleteriousness. Deleterious prediction methods might filter, for example, coding variants that do not result in an amino acid change, substitutions that do not alter the physicochemical properties of protein product despite the mutated amino
acid, or amino acid variants that are not well conserved across orthologues. MetaSVM\textsuperscript{38} was used to infer the impact of missense mutations. Rare protein-altering variants predicted to be deleterious were then selected if they occurred as pathogenic variants described in NCBI Clin Var, and/or in genes previously associated with liver-related diseases listed in the Online Mendelian Inheritance in Man (OMIM) database. BLAT, a local alignment software embedded within the UCSC Human Genome Browser\textsuperscript{39}, was used to verify that a pathogenic variant and its surrounding sequences mapped specifically to the target gene. Figure 3 outlines the genetic variant filtering strategy used in this study.
Figure 3. Representative flowchart of genetic variant filtering strategy in this study. Minor allele frequencies were determined using the gnomAD database.
Principal Component Analysis

Principal component analysis (PCA) was performed to determine the ancestry of the patients in our cohort. All tag SNP genotypes (genotype of a subset of single nucleotide polymorphisms within a linkage disequilibrium block) were obtained from WES data and used as inputs, along with the same SNPs from subjects in the HapMap project, to perform PCA with EIGENSTRAT software.\textsuperscript{40}

Sanger Sequencing

Sanger sequencing of the identified *PPARG* variant (p.Gly161Val) in patient 1 was performed by PCR amplification of genomic DNA of the proband and her parents. Sanger sequencing of the identified *ABCB4* variants (p.Arg549Cys and p.Ala934Thr) in patient 2 was performed by PCR amplification of genomic DNA of the proband, her mother and her son. Sanger sequencing of the identified *ABCB4* variant (p.Ter1280Arg) in patient 3 was performed by PCR amplification of genomic DNA of the proband. Sanger sequencing of the identified *NDFUB3* variant (p.Trp22Arg) in patient 4 was performed by PCR amplification of genomic DNA of the proband and her parents. Sanger sequencing of the heterozygous splice-site variant (c.2067+1G>A) in *APOB* in patient 5 was confirmed by PCR amplification of genomic DNA of the proband. Forward and reverse primers for each variant are described in Table 1.
Table 1. Gene name, accession number, and forward and reverse primer sequences used for Sanger sequencing.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>NCBI reference sequence</th>
<th>Amino acid change</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPARG</td>
<td>NM_015869</td>
<td>p.Gly161Val</td>
<td>5'-CAGGCCAGTATACC TTTCGC-3'</td>
<td>5'-GGATCCGACAGTT AAGATCACA-3'</td>
</tr>
<tr>
<td>ABCB4</td>
<td>NM_000443</td>
<td>p.Gly161Val</td>
<td>5'-ATGTGGTGGTCTTT CAGCT-3'</td>
<td>5'-CTTCAAGAGCTGAT CCATGT-3'</td>
</tr>
<tr>
<td>ABCB4</td>
<td>NM_000443</td>
<td>p.Ala934Thr</td>
<td>5'-ACCAAATCGAAAAC AACCGGCA-3'</td>
<td>5'-AGGAGGCTGAAGA GATGGTTACA-3'</td>
</tr>
<tr>
<td>ABCB4</td>
<td>NM_000443</td>
<td>p.Ter1280Arg</td>
<td>5'-ATCAAGACAGGTTG CACTTCTAACT -3'</td>
<td>5'-GAATGGGAGAGTC AAGGACAT -3'</td>
</tr>
<tr>
<td>NDFUB3</td>
<td>NM_002491</td>
<td>p.Trp22Arg</td>
<td>5'-GTGTTATCTTTTCC TTACAGACATG-3'</td>
<td>5'-CATTGAAAAGCAAC ATAGACACTTG-3'</td>
</tr>
<tr>
<td>APOB</td>
<td>NM_000384</td>
<td>c.2067+1G&gt;A</td>
<td>5'-GGAAGTGCTGGTG GTTCTT-3'</td>
<td>5'-TTCCATCACCTTGAC CCAGCC-3'</td>
</tr>
</tbody>
</table>
Orthologues

Full-length orthologous protein sequences from both vertebrate and invertebrates were obtained from GenBank. Protein sequences were aligned using the ClustaW or Clustal Omega algorithm.

Author Contributions

A.D., E.O., D.A., M.S., J.B., D.J., P.K.M., and S.V. participated in patient recruitment and/or patient’s ascertainment and management; A.H. performed the exome sequencing analysis for all patients in the study; D.D., K.D., and A.B. assisted with DNA extraction, next-generation and Sanger sequencing analysis; X.Z., and D.J. analyzed pathologic specimens. A.H. wrote the thesis.
RESULTS

Study population characteristics and whole-exome sequencing in cohort 1

Nineteen adults with unexplained liver disease and no history of alcohol overuse were recruited from Yale New Haven Health after an unrevealing conventional work-up performed by a hepatologist. These individuals presented between the ages of 22 and 73 years-old with a variety of liver disorders (Table 2) with or without other co-morbidities. We performed individual whole-exome sequencing of germ line DNA isolated from each patient. Targeted bases were sequenced by a mean of 90 reads, with 94% of targeted bases having more than eight independent reads, and 92% having more than fifteen independent reads, conferring high confidence calling of homozygous and heterozygous variants across the exome (Table 3). Genomic analysis identified a monogenic disorder in five patients of this adult population cohort (~25%), gaining insight into liver disease pathogenesis and with direct impact on clinical management (Table 4). Ethnicity was determined using Principal Component Analysis (Figure 4).
Table 2: Summary of study population characteristics and demographics in cohort 1 (n = 19). *n*, number; HELLP, hemolysis, elevated liver enzymes and low platelet counts; *yo*, years-old.

<table>
<thead>
<tr>
<th>Clinical Category</th>
<th>Patients, <em>n</em></th>
<th>Mean age (range), <em>yo</em></th>
<th>Male, <em>n</em></th>
<th>Female, <em>n</em></th>
<th>Alive <em>n</em></th>
<th>Deceased or transplanted <em>n</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryptogenic cirrhosis</td>
<td>7</td>
<td>39 (29-70)</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Non-obese NAFLD +/- NASH</td>
<td>6</td>
<td>30 (24-34)</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Idiopathic cholestasis</td>
<td>4</td>
<td>46 (23-73)</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>HELLP and severe hyperammonemia</td>
<td>1</td>
<td>22</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Idiopathic noncirrhotic portal hypertension</td>
<td>1</td>
<td>60</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19</strong></td>
<td><strong>38 (22-73)</strong></td>
<td><strong>11</strong></td>
<td><strong>8</strong></td>
<td><strong>16</strong></td>
<td><strong>3</strong></td>
</tr>
</tbody>
</table>
Table 3: Sequencing coverage and quality metrics for patient cohort 1 (n = 19).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read length, bp</td>
<td>74-99</td>
</tr>
<tr>
<td>Mean independent reads per targeted base</td>
<td>90</td>
</tr>
<tr>
<td>% of targeted bases with ≥ 8 independent reads</td>
<td>94.0</td>
</tr>
<tr>
<td>Mean error rate</td>
<td>0.0027</td>
</tr>
</tbody>
</table>
Table 4: Demographics, clinical features, and genetic diagnosis identified in five subjects in cohort 1, and its clinical implications. Ethnicity was determined by principal component analysis as described in Methods section; yo, years-old, F, female; M, male.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Age (yo)</th>
<th>Ethnicity</th>
<th>Sex</th>
<th>Presenting Features</th>
<th>Genetic Diagnosis</th>
<th>Clinical Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33</td>
<td>European</td>
<td>F</td>
<td>Lean NASH, hypertriglyceridemia, recurrent pancreatitis</td>
<td>Familial Partial Lipodystrophy Type 3</td>
<td>Initiation of leptin therapy; preventative cardiovascular measures; family counseling</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>African</td>
<td>F</td>
<td>Cryptogenic cirrhosis decompensated by esophageal variceal hemorrhage</td>
<td>MDR3 Deficiency</td>
<td>Family counseling; transplant candidacy</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>European</td>
<td>M</td>
<td>Cryptogenic cirrhosis at 8 years-old, now status post liver transplantation</td>
<td>MDR3 Deficiency</td>
<td>Family counseling; re-transplant candidacy</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>European</td>
<td>M</td>
<td>Non-obese NAFLD, recurrent avascular necrosis, short stature</td>
<td>Mitochondrial complex I deficiency</td>
<td>Supplementation with co-enzyme Q10, vitamins B2 &amp; B5; preventive interventions; family counseling</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>Asian</td>
<td>M</td>
<td>Lean NAFLD</td>
<td>Heterozygous Familial Hypobetalipoproteinemia</td>
<td>Family counseling; consideration for vitamin E supplementation</td>
</tr>
</tbody>
</table>
Figure 4: Example of principal component analysis to determine ethnicity clustering. Tag SNPs from exome sequences of subjects in our cohort were combined with HapMAP SNP data and PCA was performed as described in Methods. This figure shows that the patient (red cross) strongly clusters with individuals of European ancestry.
Exome sequencing yields a diagnosis and initiation of therapy in a patient who suffered from devastating complications of undiagnosed Familial Partial Lipodystrophy Type 3 for 18 years

Patient 1 is a 33 year-old female with biopsy proven severe (80-90%) hepatic macrovesicular steatosis with periportal and pericentral fibrosis (Figure 5A, B). There was moderate portal inflammation with occasional hepatocyte ballooning, rare poorly-formed Mallory-Denk bodies, ceroid laden macrophages and marked Kupffer cell siderosis. Her past medical history is significant for early onset hyperlipidemia diagnosed in childhood, recurrent episodes of hypertriglyceridemia-induced pancreatitis complicated by total pancreatectomy, splenectomy and insulin dependence. She also has history of hypertension and of pre-eclampsia at the age of 29. Her social and family history is non-contributory. She had been seen and evaluated by many expert pediatric and adult physicians at several U.S. tertiary medical centers within the last 18 years and despite a comprehensive work-up, her operational diagnosis was hyperchylomicronemia syndrome although genetic deficiency of lipoprotein lipase or apolipoprotein CII could not be demonstrated.
Figure 5: Liver histology findings in patient 1, cohort 1. (A) Liver parenchyma shows marked steatosis with moderate steatohepatitis (H&E stain). (B) Trichrome stain of liver biopsy tissue shows portal, periportal and perisinusoidal fibrosis consistent with stage 2 fibrosis (Brunt grading and staging system).

We performed WES of germ line DNA to investigate a possible underlying genetic defect. Since her biological parents were unaffected, we analyzed her exome data considering both a recessive as well as a dominant pattern of inheritance. Consistent with an unrelated union, no rare homozygous genotypes were observed in the proband. However, she harbored one missense variant (chr3:12434114, G>T, NM_015869, c.482G>T, p.Gly161Val) in PPARG (Figure 6), which encodes peroxisome proliferator-activated receptor γ, and heterozygous pathogenic variants in this gene have been related to autosomal dominant Familial Partial Lipodystrophy Type 3 (FPLD3). This variant is predicted to be damaging by MetaSVM and it is absent among > 100,000 alleles in the gnomAD database, and therefore is likely to be pathogenic in this patient (Table 5). Sanger sequencing confirmed the heterozygous variant in the proband and
further showed that neither parent harbors the variant, revealing that it occurred 
*de novo* in the patient (Figure 7A). This variant is located in the DNA binding 
domain of the PPARG protein at a highly conserved position across orthologues 
(Figure 7B). Moreover, a single case of an older female with clinical features 
consistent with FPLD3 and harboring the same PPARG variant (p.Gly161Val) as 
patient 1 has been reported. At this point, in light of new genotype knowledge 
and presumed diagnosis, we re-evaluated her clinical and laboratory findings, 
which were consistent with FPLD3.
Figure 6: Representative plot read of disease-causing mutation identified in Patient 1, cohort 1, identified using next generation sequencing. This illustrates the concept of reads and coverage for a particular genomic segment.
Table 5: Diagnostic genetic variants identified in five subjects in cohort 1. AA, amino acid; gnomAD, Genome Aggregation Database (includes 123,136 exome and 15,496 whole-genome sequences); MAF, minor allele frequency; MetaSVM scores missense variants on a scale of -2 to 3, with scores <0 predicted to be tolerated (T) and scores >0 predicted to be damaging (D); yo, years-old; N/A, not applicable.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Gene</th>
<th>Inheritance and Effect of Variant</th>
<th>AA or cDNA change</th>
<th>Meta SVM score prediction</th>
<th>gnomAD (overall)</th>
<th>gnomAD (highest frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PPARG (NM_015869)</td>
<td>Heterozygous, Missense</td>
<td>p.Gly161Val (c.482G&gt;T)</td>
<td>1.104 (D)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>ABCB4 (NM_000443)</td>
<td>Compound Heterozygous, Missense</td>
<td>p.Arg549Cys (c.1645C&gt;T)</td>
<td>0.448 (D)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p.Ala934Thr (c.2800G&gt;A)</td>
<td>0.676 (D)</td>
<td>1.3e⁻³</td>
<td>1.4e⁻² (African)</td>
</tr>
<tr>
<td>3</td>
<td>ABCB4 (NM_000443)</td>
<td>Homozygous, Stop loss</td>
<td>p.Ter1280Arg (c.3838T&gt;A)</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>NDUFB3 (NM_002491)</td>
<td>Homozygous, Missense</td>
<td>p.Trp22Arg (c.64T&gt;C)</td>
<td>0.485 (D)</td>
<td>8.2e⁻⁴</td>
<td>1.3e⁻³ (European)</td>
</tr>
<tr>
<td>5</td>
<td>APOB (NM_000384)</td>
<td>Heterozygous, Splice Site</td>
<td>c.2067+1G&gt;A</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 7: Genetic findings in patient 1, cohort 1. (A) Pedigree depicts proband and unaffected subjects in black and white symbols, respectively, with Sanger sequencing chromatograms of the proband and her unaffected parents. **PPARG** alleles are denoted WT (wild-type) or Mut (p.Gly161Val). The **PPARG** p.Gly161Val variant is heterozygous in the proband and absent in both parents. (B) Location of Gly-161 in **PPARG** and its conservation across species. The patient's variant is in a highly conserved DNA-binding domain (DBD) of the **PPARG** protein. Amino acid positions identical to the human reference are highlighted in yellow. N-terminal transactivation domain (AF1), highly conserved DNA-binding domain (DBD), C-terminal ligand-binding domain (LBD).

**PPARG** is known to be a key transcription factor in adipocyte differentiation, which explains the lack of adipose tissue in the patient and the striking accumulation of triglycerides in the bloodstream as well as its accumulation in other organs, such as the liver (Figure 8). Given the diagnosis of FPLD3 and physical exam findings of decreased overall adiposity, we postulated that the patient's leptin levels should be low since leptin is mainly produced by adipocytes. As predicted, her leptin level was significantly low at 1.5 ng/mL (normal range for age and gender = 8.0-38.9 ng/mL) (Figure 9A). This finding not
only supported our genetic diagnosis, but led to a new therapeutic intervention. The patient was initiated on leptin replacement therapy in a compassionate use program. During the ensuing 13 months there was significant amelioration of dyslipidemia: total cholesterol fell from 238 to 130 mg/dL, triglyceride levels fell from 3532 to 267 mg/dL, and HDL cholesterol increased from 8 to 24 mg/dL (Figure 9A). Concomitantly, there was normalization of her liver transaminases: AST and ALT levels decreased from 114 to 19 U/L and from 94 to 17 U/L, respectively (Figure 9B). Furthermore, the patient’s daily insulin requirements decreased by approximately half compared to doses of insulin required prior to leptin replacement therapy.

![Diagram](image)

Figure 8: Illustrative representation of the role of *PPARG*, peroxisome proliferator-activated receptor-gamma, in adipocyte differentiation and relationship to serum leptin. Adapted from Tsoukas and Mantzoros.⁴²
Figure 9: Laboratory findings in patient 1, cohort 1. (A) Triglycerides and leptin levels before (no shadow) and after (depicted by a gray shadow) initiation of leptin replacement therapy. (B) ALT/AST levels before (no shadow) and after (depicted by a gray shadow) initiation of leptin replacement therapy. LLN, lower limit of normal; ULN, upper limit of normal.
**MDR3 deficiency diagnosed in adulthood in two unrelated patients**

Patient 2 is a 31-year-old female who presented with an acute esophageal variceal hemorrhage in her 22\textsuperscript{nd} week of her second gestation. Of note, the patient had symptomatic cholestasis during her first pregnancy and she was diagnosed with benign intrahepatic cholestasis of pregnancy at that time. On this presentation, she had severe cholestasis (with high $\gamma$-glutamyltransferase, GGT) and liver biopsy revealed cirrhosis and cholestasis with ductular proliferation (Figure 10A, B). WES revealed two rare variants predicted to be damaging by MetaSVM (chr7:87069069, C>T, NM_000443, c.1645C>T, p.Arg549Cys and chr7:87041333, G>A, NM_000443, c.2800G>A, p.Ala934Thr) in \textit{ABCB4}, which encodes ATP binding cassette subfamily B member 4, also known as multiple drug resistance protein 3 (MDR3) (Table 5). There were no other disease-causing variants in genes known to be associated with liver disease, using recessive or dominant models. Sanger sequencing confirmed both variants in the patient’s genomic DNA. Study of available family members revealed that her mother is a heterozygous carrier of p.Arg549Cys whereas her son is heterozygous for the other variant, p.Ala934Thr, supporting that the two variants found in our patient were located \textit{in trans} at the \textit{ABCB4} gene locus (Figure 11A, B), consistent with an autosomal recessive disorder. Both variants are conserved across orthologues (Figure 12). As expected, her total bile acids (cholic acid, deoxycholic acid and chenodeoxycholic acid) levels have been persistently elevated, ranging from 36 to 105$\mu$mol/L (normal is <6.8$\mu$mol/L).
Figure 10: Liver histology findings in patient 2, cohort 1. (A) Trichrome stain for patient 2 reveals fibrous septa with nodule formation consistent with cirrhosis. (B) Cytokeratin 7 immunohistochemical staining (depicted in brown) for patient 2 shows ductular proliferation.
Figure 11: Genetic findings in patient 2, cohort 1. (A) Pedigree depicts patient 2 and unaffected subjects shown in black and white symbols, respectively. (B) Sanger sequencing chromatograms of the proband (patient 2) and her unaffected mother and son. The patient harbors two variants in \textit{ABCB4}, p.Arg549Cys and p.Ala934Thr, whereas her mother and her son solely carry one of these variants each, p.Arg549Cys and p.Ala934Thr, respectively.
Figure 12: Conservation findings in patient 2, cohort 1. Conservation of Arg-549 and Ala-934 across species, respectively. Amino acid positions identical to the human reference are highlighted in yellow.

Patient 3 is a 29 year-old male with past medical history significant for liver transplantation at 8 years of age for cirrhosis of unclear etiology who presented with chronic graft rejection in the setting of medical non-adherence. Consistent with known consanguinity, patient 3 harbors twenty-seven rare homozygous variants. Twenty-six of these variants are missense variants in genes unrelated to liver disease by OMIM, and therefore are unlikely to be causal pathogenic variants in this patient. The other homozygous variant encoded a stop lost in ABCB4 (chr7:87031414, T>A, NM_000443, c.3838T>A, p.Ter1280Arg), which was confirmed by Sanger sequencing (data not shown). This variant has never been reported in general population by gnomAD. However, another substitution in the same nucleotide (NM_000443 c.3838T>C) resulting in an identical stop-loss variant in the p.Ter1280 codon has been previously associated with cholestasis and liver failure.43 Both the c.3838T>A and c.3838T>C nucleotide substitutions result in extension of the protein by 19 new amino acids. Given new knowledge of the patient’s genotype, we reviewed the patient’s liver explant.
Liver parenchyma showed cirrhosis with patchy macrovesicular steatosis, portal/septal chronic inflammation, marked cholestasis with ductular proliferation and increased hepatocytic copper deposition (Figure 13A, B, C, D), consistent with a genetic diagnosis of MDR3 deficiency.

Figure 13: Liver histology findings in patient 3, cohort 1. (A) Liver parenchyma (H&E stain) from patient 3 showing cirrhosis. (B, C) Liver parenchyma (H&E stain) from patient 3 at lower and higher magnification, respectively, revealing portal/septal chronic inflammation, ductular proliferation, and marked cholestasis. (D) Copper stain of liver parenchyma for patient 3 supporting marked cholestasis.
Recessive variant in NDFUB3 in a patient with elevated transaminases, hepatic steatosis, recurrent avascular necrosis and short stature

Patient 4 is a 32 year-old male with persistent elevation of transaminases (AST and ALT ~3 times upper limit of normal), recurrent avascular necrosis, and short stature (with both parents of average height). Liver biopsy showed minimal macrovesicular (small and large droplet fat) steatosis (<5%) (Figure 14A, B). He is the single child of unrelated European-descended parents. WES was remarkable for a very rare homozygous variant in NDUFB3 (chr2:201943669, T>C, NM_002491, c.64T>C, p.Trp22Arg), which encodes for NADH-dehydrogenase 1 beta complex 3 and consists of the first enzyme in the electron transport chain in mitochondria (Table 5). Both parents were found to be heterozygous for this variant by Sanger sequencing (data not shown). This variant is predicted to be damaging by MetaSVM and it was first reported in a single female infant with lethal complex I mitochondrial deficiency and more recently in 10 children from 8 families, seven of them of Irish ancestry. In this last cohort, all patients have short stature (<9th percentile) and similar facial dysmorphic features to patient 4, such as a prominent forehead, smooth philtrum and deep-set eyes. In contrast to the first patient reported with this homozygous variant, they have a good long-term prognosis, even though some patients presented with an acute metabolic crisis with evidence of an isolated complex I deficiency in muscle. In light of new genotype information, our patient 4 had mitochondrial ETC testing in skeletal muscle biopsy that revealed a deficiency of rotenone sensitive I+III activity despite normal citrate synthase activity, which
fulfills a minor criterion of the modified Walker criteria for diagnosis of a respiratory chain disorder (<30%). Liver electron microscopic findings were also suggestive of a mitochondrial abnormality, with hepatocytes showing different sized lipid droplets (Figure 14C). Moreover, the patient suffers from oculomotor dysfunction with optic nerve anomalies, episodes of lactic acidosis during surgical interventions, and progressive fatigue. He was started on mitochondrial cocktail supplement, including co-enzyme Q10, vitamins B2 and B6.

![Figure 14](image-url)

Figure 14. Liver biopsies of patient 4, cohort 1. (A, B) Patient 4 liver parenchyma shows near normal histology with minimal macrovesicular steatosis (small and large droplet fat) (H&E stain) at lower and higher magnification, respectively. (C) Electron microscopic findings for patient 4 are suggestive of a mitochondrial abnormality, with hepatocytes showing different sized lipid droplets.
Lean patient with hepatic steatosis of unknown etiology was found to have a novel damaging heterozygous variant in APOB

Patient 5 is a 23 year-old lean male who presented for evaluation of persistent elevated ALT (2 to 3 times upper limit of normal) for which he underwent a liver biopsy that showed mild macrovesicular steatosis (30%) with minimal lobular inflammation and minimal pericentral sinusoidal fibrosis (Figure 15). His work-up was also remarkable for a ceruloplasmin of 17mg/dL (normal range =18-36mg/dL). WES analysis revealed no rare variations in ATP7B and a pathogenic heterozygous splicing variant in APOB (chr2:21250699, c.2067+1G>A), which encodes apolipoprotein B (ApoB). This change was validated by Sanger sequencing (data not shown) and is predicted to abrogate the normal splicing of exon 14 in the APOB gene. Heterozygous carriers typically have decreased plasma levels of LDL-cholesterol and ApoB, and may be asymptomatic or have clinical manifestations such as liver steatosis. Consistent with the patient’s genotype, his circulating ApoB level was found to be half the lower limit of normal (26 mg/dL compared to normal range of 52 to 109 mg/dL), with low circulating LDL cholesterol (20mg/dL) and triglyceride (19 mg/dL) levels. Thus, this new genotype information explains both clinical and laboratory findings in this patient, and may have implications in his clinical management beyond family counseling. Specifically, fat-soluble vitamin E supplementation has been recommended for patients with homozygous hypobetalipoproteinemia to protect from neurological complications, and modest doses of Vitamin E have been proposed for treatment
of heterozygous patients,\textsuperscript{48} but further studies are required. Vitamin E has also been shown to be beneficial in those with non-alcoholic steatohepatitis.\textsuperscript{49}

Figure 15. Liver biopsy of patient 5, cohort 1. Hepatic parenchyma (H&E stain) from patient 5 shows predominantly macrovesicular steatosis with minimal steatohepatitis.
Outcomes of patients for patients in cohort 1 who remain unexplained after whole exome sequencing analysis

Fourteen patients in cohort 1 remain unexplained after genomic analysis. Our study aimed to demonstrate the utility of unbiased WES in attaining a definitive diagnosis in adults with idiopathic liver disease. Therefore, we utilized a conservative approach (to avoid false positives), and focused our analysis on variants previously reported as pathogenic in Clin Var and/or in genes associated with liver-related phenotypes by OMIM. Our study was not designed or powered to identify novel liver disease genes, to determine phenotypic impact of variants predicted to be benign by in silico methods (which would require dedicated functional analysis) or to search for disease-causing mutations with minor allele frequency above the thresholds discussed above.

Each of the 14 patients that remain undiagnosed in our cohort harbor rare variants in (1) genes not known to be related to liver disease and/or (2) in genes related to liver disease, but which do not fulfill the known genetic model of inheritance (i.e. patient harbors one heterozygous variant in a liver disease-related gene that is known to cause a recessive disorder where both alleles need to be mutated to cause disease, such as ATP7B) or are not supported by phenotype (a genetic diagnosis alone cannot establish the clinical diagnosis if not in accordance with clinical findings and should be strictly used in conjunction with the clinical data). Table 6 summarizes the outcomes of the 14 patients that remained unexplained.
Table 6. Summary of outcomes for fourteen patients who remain unexplained after whole exome sequencing analysis. Table shows the number of genes with rare deleterious variants (includes premature termination, splice-site, frameshift and missense variants predicted to be deleterious by MetaSVM) and with benign variants (includes missense variants predicted to be tolerated by MetaSVM). All rare deleterious variants in these fourteen patients occur in (1) genes not known to be related to liver disease and/or (2) in genes related to liver disease, but which do not fulfill known genetic model of inheritance or are not supported by phenotype. Rare deleterious variants found in liver disease-related genes are indicated below.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Recessive Pattern of Inheritance</th>
<th>Dominant Pattern of Inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Homozygous variants</td>
<td>Two heterozygous variants in the same gene</td>
</tr>
<tr>
<td></td>
<td>Deleterious</td>
<td>Benign</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<sup>a</sup>one heterozygous deleterious missense variation in *ABCB4*, chr7:87076396, NM_000443, c.959C>G, p.Ser320Cys. Laboratory tests and liver biopsy findings show no cholestasis.

<sup>b</sup>one heterozygous deleterious missense mutation in *SLC40A1*, chr2:190426791, NM_014585, c.1529T>C, p.Met510Thr; the patient’s phenotype is not consistent with autosomal dominant hereditary hemochromatosis (normal ferritin and normal-to-low transferrin saturation).

<sup>c</sup>one heterozygous splice site variation in *ABCB4*, chr7:87032450, NM_000443, c.3492+1G>A. The patient suffered from decompensated cirrhosis of unclear etiology with preserved bile ducts seen on biopsy, now status post liver transplantation. Additionally, at the time of transplant, the gallbladder showed no significant histopathologic abnormality.

<sup>d</sup>two heterozygous frameshift mutations in *ATP7B*: chr13:52548577, NM_000053, c.778T>TG, p.Gln260fs and 13:52548830, NM_000053, c.524CTT>C, p.Lys175fs. Segregation studies revealed that patient’s unaffected mother also carries both variants and therefore they are *in cis*. 
Patient 13 presented with cholestasis of unclear etiology. Whole exome sequencing revealed a rare, heterozygous deleterious mutation in peroxisome proliferator-activated receptor delta, or *PPARD*, 6:35393750, NM_001171820, c.926G->A, p.R309Q. *PPARD* is a nuclear receptor expressed in hepatocytes, cholangiocytes, Kupffer cells and hepatic stellate cells, and it controls genes involved in bile acid homeostasis. Selective agonists of the *PPARD* receptor have been shown to be potent anti-cholestatic agents in clinical trials,\(^5\) however mutations in *PPARD* have not been previously associated with cholestatic liver diseases. The primary therapeutic effects of the *PPARD* selective agonists in patients with primary biliary cholangitis result from decreasing hepatocellular bile acid concentrations by downregulation of *CYP7A1*, which encodes cholesterol 7-alpha-hydroxylase, the enzyme that hydroxylates cholesterol in the first step in the synthesis of bile acids.\(^6\)\(^,\)\(^7\) As a consequence, we hypothesized that patient 13 may have had increased activity of cholesterol 7-alpha-hydroxylase, leading to idiopathic cholestasis. To assess this, we measured serum C4 concentrations (7-alpha-hydroxy-4-cholesten-3-one) in peripheral blood, a reliable marker of the activity of hepatic 7-alpha-hydroxylase.\(^8\) It is known that serum C4 concentration in non-cirrhotic and cirrhotic primary biliary cholangitis patients is normal or low compared to healthy control subjects.\(^9\) The patient’s bile acids were elevated (deoxycholic acid 4.5 µmol/L [normal <2.4 µmol/L], chenodeoxycholic acid 4.6 µmol/L [normal <3.1 µmol/L], and total bile acids 10.8 µmol/L [normal <6.8 µmol/L]), consistent with known cholestasis despite ongoing ursodiol therapy. However, the results of the serum C4 concentration were within normal limits at
2.7 ng/mL (normal range 2.5-63.2 ng/mL), suggesting that increased activity of liver cholesterol 7-alpha-hydroxylase was not contributing to the patient’s phenotype. Further studies on other cholesterol synthesis intermediates known to be regulated by *PPARD* (squalene, lanosterol, desmosterol, lanthosterol, and 7-dehydrocholesterol), as well as intestinal cholesterol absorption markers known to be regulated by *PPARD* (β-sitosterol, campesterol, sigmasterol, coprostanol) are in progress.

*Study population characteristics and whole-exome sequencing in cohort 2*

Four adults were recruited from Bridgeport Hospital presenting with fatty liver disease, insulin resistance/diabetes, hypertriglyceridemia and physical exam findings suggestive of lipodystrophy per evaluation by an endocrinologist. Individual whole-exome sequencing of germline DNA isolated from each patient was performed. Targeted bases were sequenced by a mean of 75 reads, with 94% of targeted bases having more than eight independent reads, and 93% having more than fifteen independent reads (Table 7). We identified a pathogenic genetic alteration that could potentially explain the phenotype in all four patients in this cohort, with possible clinical implications of the diagnosis in each case (Table 8). Two patients had genetic alterations in genes previously associated with lipodystrophy (*LMNA* and *PPARG*) and which are included in comprehensive lipodystrophy sequencing panels, however two patients had alterations in genes not routinely found on lipodystrophy panels (*LEPR* and *APOE*).
Table 7: Sequencing coverage and quality metrics for patient cohort 2 (n = 4)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read length, bp</td>
<td>99</td>
</tr>
<tr>
<td>Mean independent reads per targeted base</td>
<td>75</td>
</tr>
<tr>
<td>% of targeted bases with ≥ 8 independent reads</td>
<td>94.3</td>
</tr>
<tr>
<td>Mean error rate</td>
<td>0.0024</td>
</tr>
</tbody>
</table>

Table 8: Demographics, clinical features, and genetic diagnosis identified in four subjects in cohort 2, and its clinical implications. Ethnicity was determined by principal component analysis as described in Methods section; yo, years-old, F, female; M, male

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Age (yo)</th>
<th>Ethnicity</th>
<th>Sex</th>
<th>Presenting Features</th>
<th>Genetic Diagnosis</th>
<th>Clinical Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62</td>
<td>European</td>
<td>F</td>
<td>Diabetes, hepatic steatosis</td>
<td>Familial Partial Lipodystrophy Type 2</td>
<td>Consideration for leptin therapy; preventative cardiovascular measures; family counseling</td>
</tr>
<tr>
<td>2</td>
<td>46</td>
<td>Mexican</td>
<td>F</td>
<td>Cirrhosis with steatohepatitis</td>
<td>Type 3 hyperlipoproteinemia</td>
<td>Preventative cardiovascular measures; family counseling</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>Mexican</td>
<td>F</td>
<td>Hypertriglyceridemic pancreatitis</td>
<td>Leptin receptor deficiency</td>
<td>Consideration for MCR4 agonist therapy, family counseling</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>European</td>
<td>F</td>
<td>Hypertriglyceridemic pancreatitis, diabetes</td>
<td>Familial Partial Lipodystrophy Type 3</td>
<td>Initiation of pioglitazone; preventative cardiovascular measures; family counseling</td>
</tr>
</tbody>
</table>
Genomic analysis in patients with fatty liver disease and lipodystrophy-like phenotype in cohort 2

Patient 1 in cohort 2 presented with hepatic steatosis and insulin resistance and was found to harbor a heterozygous missense mutation in *LMNA* (p.Leu92Val), encoding lamin A/C. This led to the diagnosis of Type 2 familial partial lipodystrophy (FLPD2), also called Dunnigan disease, a rare autosomal dominant lipodystrophic disorder (Table 9). FPLD2 disease manifestations comprise subcutaneous fat loss in the buttocks, trunk and limbs, resulting in peripheral muscular hypertrophy and cervicofacial fat accumulation. Metabolic complications, resulting from an inability to properly store lipids, are very common and are usually revealed by hypertriglyceridemia and diabetes. Patients are also at risk for cardiomyopathy, arrhythmia, coronary artery disease, infertility, and eclampsia. Knowledge of genotype therefore informs preventive medicine, reproductive counselling, and the surveillance for early occurrence of metabolic complications in young generations. Moreover, two independent studies have evaluated leptin replacement therapy (n = 6 and n = 24) with the Dunnigan-type familial partial lipodystrophy, demonstrating efficacy in decreasing circulating triglycerides and liver steatosis.

Patient 2 in cohort 2 presented with cirrhosis and steatohepatitis and had elevated LDL-C (240 mg/dL), and triglycerides (496 mg/dL). She was found to be heterozygous for two rare point mutations in *APOE* (p.Glu31Lys and p.Arg163Cys), reported to be pathogenic by Clin Var. Both mutations have been previously reported in a 24-year-old Puerto Rican white female with severe type
3 hyperlipoproteinemia, who presented with palmar xanthomas and
tuberoeruptive xanthomas on the elbows, knees, and buttocks, increased plasma
cholesterol and triglyceride levels, and the presence of cholesterol and apoE-
enriched VLDL and chylomicron remnants. It is well documented that the
absence of apolipoprotein E predisposes to hypercholesterolemia,
atherosclerosis, and obesity. Furthermore, ApoE−/− mice fed a high-fat-diet
mimic major characteristics of human NASH including steatosis, inflammation
and fibrosis. Of note, mutations in apolipoprotein E have not been previously
associated with the lipodystrophy phenotype.

Patient 3 presented with abdominal pain and was diagnosed with
hypertriglyceridemic pancreatitis (triglycerides of 2076 mg/dL). She had
prominent central obesity with a paucity of adipose tissue on the extremities and
gluteal region. Family history was significant for pancreatitis in the mother and
maternal grandfather, and NASH in her 14-year-old son and sister. Whole exome
sequencing uncovered a novel homozygous deleterious mutation in the leptin
receptor (p.Ser389Asn). Congenital deficiency of the leptin receptor in humans
results in extreme obesity and characteristic features of hyperphagia, obesity,
hypogonadism, and impaired T-cell-mediated immunity. It is also associated
with the selective deposition of fat mass, as seen in subjects with leptin
deficiency. Interestingly, patient 3 appears to have a milder phenotype (BMI of
27, mild hyperphagia, FSH, LH and TSH within normal limits) compared to
previously described cases of congenital deficiency of the leptin receptor.
Heterozygotes who are leptin receptor-deficient but not obese are known to have
increased fat mass.\textsuperscript{60,61} Further workup also showed low fasting leptin (4.2 ng/mL, reference range 8.0-38.9 ng/mL). Ongoing work will elucidate the functional status of the mutated receptor in this patient using T-cells, which are known to express the leptin receptor. We will also screen family members and search for potential genetic modifiers of the \textit{LEPR}-deficient state in this patient. Interestingly, MCR4 agonist therapy with setmelanotide has been suggested as a potential therapeutic option for severely obese patients with leptin receptor deficiency.\textsuperscript{62}

Patient 4 presented with hypertriglyceridemic pancreatitis (triglycerides of 4299 mg/dL), and was found to have a previously described mutation in \textit{PPARG}. The p.Arg194Trp mutation is at a position highly conserved across orthologues, leading to clinical manifestations of FPLD3 and associated metabolic disturbances.\textsuperscript{63} In vitro, this mutation was found to disrupt DNA binding activity and lead to haploinsufficiency, with no dominant negative activity. The patient was found to have a leptin level at the lower limit of normal and was started on pioglitazone, which selectively stimulates the nuclear receptor \textit{PPARG}. 
Table 9: Diagnostic genetic variants identified in cohort 2. *Pathogenic according to Clin Var. AA, amino acid; gnomAD, Genome Aggregation Database (includes 123,136 exome and 15,496 whole-genome sequences); MAF, minor allele frequency; MetaSVM scores missense variants on a scale of -2 to 3, with scores <0 predicted to be tolerated (T) and scores >0 predicted to be damaging (D); yo, years-old; N/A, not applicable.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Gene</th>
<th>Inheritance and Effect of Variant</th>
<th>AA or cDNA change</th>
<th>Meta SVM score prediction</th>
<th>gnomAD (overall)</th>
<th>gnomAD (highest frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LMNA (NM_001282626)</td>
<td>Heterozygous, Missense</td>
<td>p.Leu92Val (c.274C&gt;G)</td>
<td>0.547 (D)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>APOE (NM_000041)</td>
<td>Compound Heterozygous, Missense</td>
<td>p.Glu31Lys (c.91G&gt;A)</td>
<td>-0.751 (T)*</td>
<td>1.3e-4</td>
<td>3.4e-4 (Latino)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p.Arg163Cys (c.487C&gt;T)</td>
<td>0.147 (D)*</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>LEPR (NM_001198687)</td>
<td>Homozygous, Missense</td>
<td>p.Ser389Asn (c.1166G&gt;A)</td>
<td>0.882 (D)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>PPARG (NM_015869)</td>
<td>Heterozygous, Missense</td>
<td>p.Arg194Trp (c.580C&gt;T)</td>
<td>1.014 (D)</td>
<td>4.0e-6</td>
<td>8.0e-6 (European)</td>
</tr>
</tbody>
</table>
DISCUSSION

This study provides evidence that a subset of adult patients who suffer from liver disease of indeterminate etiology with or without other co-morbidities harbors an underlying Mendelian disorder, which may be unrecognized during their entire childhood until genetic testing is performed. These findings have several implications. First, by establishing a diagnosis for a substantial number of undiagnosed cases, we provided new insights into disease pathogenesis. Second, knowledge of genotype led to recognition of unappreciated phenotypic features as well as new therapeutic and preventive medicine interventions beyond family counseling. For example, genomic analysis in patient 1 (cohort 1) led to recognition of phenotypic aspects unappreciated by standard clinical examination by specialists not familiar with the FPLD3 phenotype. It also led to initiation of leptin replacement therapy with striking amelioration of metabolic dysregulation and liver disease. Correct diagnosis additionally allowed appropriate attention to monitoring and prevention of premature coronary artery disease and other cardiovascular risk factors. Third, our data highlight the importance of using WES in the investigation of liver disease of unknown cause so that we may start developing an understanding of what clinical presentations/diseases are genetic and may remain undiagnosed until adulthood. The genetic diagnosis of MDR3 deficiency in patient 2 (cohort 1) underscores the silent progression of inherited chronic liver disease to cirrhosis, portal hypertension and decompensation, remaining unrecognized for decades and with first presentation in adulthood. Fourth, reminiscent of the widely accepted
Radiology and Pathology Rounds in clinical practice, this study illustrates the potential clinical value of Genome Rounds in the individual assessment and medical care of adults suffering from liver disease of unknown cause (Figure 16). This approach perfectly exemplifies the mission of the Precision Medicine Initiative launched by U.S. President Barack Obama in January 2015. Physicians recognize that every patient is unique and have always tried to adjust their interventions as best they can to each individual. Now, we have the technology and knowledge to start translating this concept to routine clinical practice in tertiary medical centers across the world. Even in the contemporary taxonomy of liver diseases, there is little understanding of the heterogeneity of disease within each category and distinct subtypes based on their underlying genetic mutations and/or pathobiology. This concept is illustrated by the three cases in cohort 1 (patients 1, 4 and 5) and four cases in cohort 2, who each harbor distinct genetic defects affecting different molecular pathways leading to hepatic steatosis and presumed diagnosis of NAFLD, with direct implications in bedside therapeutic and preventive medicine interventions.

The four patients with metabolic abnormalities and physical exam findings of abnormal subcutaneous fat distribution in cohort 2 all harbored independent genetic alterations that may explain the phenotype. Interestingly, two of these patients harbored alterations in genes not included in lipodystrophy gene panels (LEPR and APOE), suggesting the utility of an unbiased genomic approach in cases of suspected lipodystrophy. Indeed, current guidelines recommend either "candidate gene sequencing, a panel of candidate genes, or whole-
exome/whole-genome sequencing”. Due to the clinical similarities of lipodystrophy with the common obesity-related metabolic syndrome, we suspect that many such patients are overlooked. Careful physical examination of patients with insulin resistance and hypertriglyceridemia combined with WES could help identify lipodystrophy patients and lead to novel therapeutic strategies. We suspect that the higher diagnostic yield in this cohort is the result of a more targeted clinical phenotyping used for enrollment, although further study is required.

Our study also provides support that a genomic approach to patients with undiagnosed liver diseases and/or unusual clinical presentations could lead to the discovery and development of new rational and targeted therapeutics. For example, loss of function mutations in PPARG and the concomitant phenotypic features of insulin resistance, hypertriglyceridemia, and fatty liver disease motivates the potential use of pharmacologic agonists to this nuclear receptor for the treatment of diabetes, hypertriglyceridemia, and NASH. Indeed, the thiazolidinediones, well-known PPARG agonists, are FDA-approved anti-diabetic agents, and recent randomized studies have shown utility for the treatment of NASH and dyslipidemia. Patient 5 in cohort 1 presented with hepatic steatosis and was found to have a mutation in APOB, leading to reduced LDL-C and apolipoprotein B levels. Notably, the FDA-approved drug mipomersen, an antisense oligonucleotide inhibitor of apolipoprotein B synthesis, has a black box warning for increasing hepatic steatosis, and elevated liver transaminases. Mipomersen reduced apolipoprotein B levels by ~22% and LDL-C by ~21% after
26 weeks, but also increased mean hepatic fat by ~10% and elevated liver transaminases to 2-3x ULN, thereby providing a pharmacologic correlate to the apo B phenotype.\textsuperscript{68} We expect genetic insights to drive further therapeutic advances, similar to how the discovery of loss-of-function PCSK9 variants causing hypocholesterolemia led to the development of monoclonal antibodies targeting PCSK9 to treat cardiovascular disease and hypercholesterolemia.\textsuperscript{69}

In our study, most of the diagnosed patients had seen a diverse array of medical and surgical specialists for several years prior to their diagnosis, such as primary care providers, internists, surgeons, endocrinologists and hepatologists, among others. This suggests that the investigation of unrecognized genetic disorders in adults would have clinical utility among a broad group of adult multispecialty clinical practices. Decades ago, Mendelian genetics mostly relied on family-based studies with very distinct and often severe phenotype(s). However, as illustrated in this study, the absence of family history of similar phenotype should not deter physicians from investigating a genetic cause for the unexplained liver disease since it might arise from a \textit{de novo} variant, which by definition is not inherited from any parent, or result from a recessive inheritance pattern when both parents are usually healthy carriers and 75% of siblings will be clinically unaffected.

One limitation of this study is a relatively small sample size and patient recruitment at a single tertiary care academic center and affiliated community hospital. Further studies are required to assess the generalizability of these findings in a broader liver disease population. Additionally, the patients in this
cohort whose phenotypes remain unexplained may have a pathogenic variant not detected by the methodology used, such as variants in the non-coding region of the genome (including promoters, enhancers, silencers, miRNA, etc), or in a gene not yet known to be associated with a human disease. In fact, approximately three quarters of human genes have not yet been linked to a human phenotype\textsuperscript{20}, and for this reason we will continue to re-analyze these patients’ WES data regularly. This study’s diagnostic yield is comparable to data recently reported in inherited cardiovascular diseases and chronic kidney disease in adults.\textsuperscript{29,32,33}

Figure 16: Schematic representation of multidisciplinary Genome Rounds in Adult Hepatology. It merges genotype-phenotype information with the goal of recognizing unappreciated phenotypic features by standard clinical examination, providing a diagnosis and new therapeutic options, and establishing adequate family counseling in adults. gDNA, genomic DNA.
Collectively, our data support the incorporation of WES in the diagnostic and management algorithms of adults suffering from idiopathic liver disease despite a comprehensive work-up, and underscore its value as a means of developing an understanding of the genetic basis of liver disease previously described as cryptogenic or “of unknown etiology”. A multidisciplinary Genome Rounds approach (Figure 16) is a likely forum from which to develop best practice guidelines for genomic medicine in a variety of non-oncological medical and surgical specialties, including hepatology. This strategy will shed further light on genetic contributions, and therefore underlying molecular pathogenesis, across different forms of liver disease that are clinically indistinguishable through conventional diagnostic approaches. We anticipate the diagnostic yield of WES to be highest in patients displaying chronic phenotypes from an early age, atypical clinical presentations such as lean NAFLD/NASH, the presence of congenital or syndromic features, a positive family history, the presence of consanguinity, or alternatively when WES is used to replace single gene and gene panel testing.
REFERENCES


50. Jones D, Boudes PF, Swain MG, et al. Seladelpar (MBX-8025), a selective PPAR-delta agonist, in patients with primary biliary cholangitis with an inadequate response to ursodeoxycholic acid: a double-blind, randomised,
57

51. Li T, Chiang JY. Regulation of bile acid and cholesterol metabolism by PPARs. PPAR research 2009;2009:501739.