Omics analyses of global dysregulation in Primary Biliary Cholangitis

by

Bruna Dutra

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Translational Medicine

Department of Medicine
University of Alberta

ABSTRACT

Primary Biliary Cholangitis (PBC) is a complex cholestatic liver disorder with both environmental and genetic factors that contribute to disease. While the etiology remains poorly understood, our lab continues to characterize the role of a human betaretrovirus (HBRV) in the disease process. Based on prior proteomic studies in biliary epithelial cells from patients with PBC, we have identified dysregulation in the splicing machinery. This process is necessary for the maturation of mRNA molecules by the removal of introns and reconnection of exons. Based on the knowledge that viruses systemically perturb cellular processes, we hypothesized that splicing might be globally disrupted and possibly related to patient outcomes. Using the relevant control samples, we performed additional transcriptomics and proteomics analyses using biliary epithelial cells (BEC) from patients with PBC and transcriptomics analysis on the peripheral blood of PBC patients. We found that splicing and other essential cellular pathways are deregulated in the blood and BEC of PBC patients, and the process is more advanced in PBC patients with worse prognoses. In addition, many of the proteins and genes dysregulated in the PBC patients have been previously linked with viral infection, supporting our hypothesis linking altered splicing with viral pathogenesis. These findings may be clinically significant, as they bring to light new evidence of pathogenic mechanisms implicated in the disease process. Based on these and other studies relating to the metabolic and other essential cellular processes that may be dysregulation by viral infection, we believe our data warrants further study investigating and validating the role of HBRV in generating these cellular changes *in vitro*

PREFACE

This thesis is an original work by Bruna Dutra. No part of this thesis has been previously published.

This document is a result of the work of several people. Filip Wysokinski was responsible for the extraction and culture of the Biliary Epithelial cells (BEC). Dr. Bo Meng, Dr. Wei Wang and Dr. Richard Fahlman completed the BEC protein extractions, LC-MS/MS and primary proteomics analysis. Tracy Jordan extracted RNA from the whole blood extracted from patients with primary biliary cholangitis (PBC) in a clinical trial studying the efficacy of obeticholic acid (OCA). These studies were performed in collaboration with Intercept Pharma, which provided the PBC patient samples and funding for the research. The RNA from these patients and comparison samples from patients with autoimmune hepatitis (AIH) and healthy subjects were sent to Novegene for Illumina library construction and next-generation sequencing. Dr. Juan Jovel oversaw the BEC RNA extraction, library construction, and Illumina NGS as the Applied Genomics Core manager. He also performed the differential informatic analysis for differential gene usage for the BEC and PBMC studies and the exon usage to determine splicing activity in PBMC. Hussain Syed further carried out the transcript's differential expression analysis and statistical analysis. I performed the pathway analysis and data interpretation.

These studies were approved by the University of Alberta, Human Research Ethics Board (Pro00005105 and Pro000085859)

ACKNOWLEDGMENTS

I am deeply thankful for everyone who has supported me in this journey, completing this master's degree.

Firstly, I would like to sincerely thank my supervisor, Dr. Andrew Mason, for his understanding, patience and support towards completing this degree. Under his mentorship, I gained not only scientific knowledge, but certainly learnt about perseverance and persistence. I will take his lessons with me for the rest of my life. I would also like to extend my gratitude to my committee members, Dr. Juan Abraldes and Dr. Allan Murray, for their time reviewing my work and valuable input.

I want to thank the Canadian Institute of Health Research for funding this work and the Mason lab.

I would also like to express my gratitude towards all the current and former members of the Mason lab, whom I had the pleasure to work with and learn from. I want to thank Hussain Syed, Kiandokht Bashiri, Hiatem Abofayed, Kerolous Messeha, Ning Sun, Bishoi Aziz, Dr. Steven Willows, Bavan Sangarasivam, Filip Wysokiniski and Tracy Jordan. I would like to especially thank Hussain Syed for all the help and assistance during this project and thesis development.

I want to thank my parents, Luciana and Francisco Dutra, for their unconditional love and support and my extended family, especially my grandmother, Cleci Steigleder (*in memorian*), who is always on my mind and heart.

Finally, I want to thank all my friends, the new and the old ones, for their continuous support and for helping me to go through the most challenging moments of my life.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Meaning
ALP	Alkaline phosphatase
AMA	Antimitochondrial antibody
ANA	Antinuclear antibody
BEC	Biliary epithelial cells
BLB	Bilirubin
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FDR	False discovery rate
FXR	Farnesoid X receptor
GO	Gene Ontology
GTF	General Transfer Format
HBRV	Human Betaretrovirus
HCl	Hydrochloric acid
HIV	Human Immunodeficiency Virus
HLA	Human leukocyte antigen
hnRNP	Heterogeneous nuclear ribonucleoprotein
ITAF	IRES trans-acting factor
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LPRr	Lopinavir-ritonavir
MMTV	Mouse Mammary Tumour Virus

mRNA	Messenger RNA
OCA	Obeticholic acid
OH	Hydroxyl
PBC	Primary Biliary Cholangitis
PBS	Phosphate-buffered saline
PDC-E2	Pyruvate dehydrogenase complex-E2
PMSF	Phenylmethylsulphonyl fluoride
POISE	Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis
Pre-mRNA	Precursor messenger RNA
PSC	Primary sclerosing cholangitis
RNA	Ribonucleic acid
RPM	Revolutions per minute
RPMI	Roswell Park Memorial Institute
snRNP	Small nuclear ribonucleoprotein
SR	Serine/arginine-rich protein
SS	Sjogren's syndrome
SSc	Systemic sclerosis
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TDF/FTC	Tenofovir-emtricitabine
TPM	Trannscripts per million
UDCA	Ursodeoxycholic acid
UTR	Untranslated region
UniProtKB	UniProt Knowledgebase

SUMMARY

1. INTRODUCTION

1.1 PRIMARY BILIARY CHOLANGITIS

Primary Biliary Cholangitis (PBC), previously known as Primary Biliary Cirrhosis, is a rare chronic cholestatic liver disorder that mainly affects middle-aged (40-60 years old) women, a predominance of up to 10 females to 1 male. The incidence worldwide is up to 58 cases per million, and the prevalence is up to 402 cases per million individuals annually (1-4). In Alberta, the prevalence of the disease in 2015 was 292 cases per million, and it accounts for 5% of the liver transplants in Canada (5).

PBC has an autoimmune component, characterized by the production of anti-mitochondrial antibodies (AMA), which target primarily the inner lipoic acid domain of the pyruvate dehydrogenase complex-E2 (PDC-E2), situated in the inner mitochondrial membrane (6, 7). AMAs are the serologic hallmark of the disorder and present in up to 95% (3). However, the magnitude of positivity does not seem not to be related to the severity of the disease (3, 8), and AMAs may persist after liver transplantation without the histological disease (9).

The disease is characterized by progressive nonsuppurative granulomatous destruction of the small intrahepatic bile ducts. There are at least two related processes that lead to liver damage. The first process involves the destruction of the small bile ducts, mediated by the activation of cytotoxic T lymphocytes. The second process consists of the chemical damage to hepatocytes caused by retention and accumulation of substances normally excreted into the bile (9). These processes lead to cholestasis and fibrosis and ultimately to liver failure, resulting in the need for a liver transplant (9, 10). The symptoms of this disorder include fatigue, pruritus, osteoporosis, oral

and ocular dryness (Sicca complex), abdominal pain, arthralgia, and, less commonly, malabsorption of lipids and lipid-soluble substances (9).

Most patients are diagnosed in the asymptomatic phase of the disease. The criteria for a positive diagnostic for PBC include at least two out of the three following conditions: in patients with normal biliary tract ultrasound result, an alkaline phosphatase (ALP) serum level increase during more than six months; AMA or antinuclear antibodies (ANA) positivity; and histologic evidence of nonsuppurative obstructive cholangitis compromising the bile ducts (10, 11).

PBC treatments aim to prevent liver complications and improve the symptoms associated with the disease (11, 12). The guidelines for the treatment of PBC establish ursodeoxycholic acid (UDCA) as the first-line therapy for the disorder, and, more recently, obeticholic acid (OCA) was approved as a second-line treatment for patients that do not respond adequately to UDCA (11).

1.1.1 PBC as a systemic disease

PBC is centred on the patient's liver, but individuals often present with extrahepatic symptoms and up to 80% of PBC patients present with non-liver-related manifestations (13, 14). Fatigue is the most common symptom present in up to 80% of the cases, and patients report that the malaise is mentally and physically debilitating daily (13, 15). Fatigue is present independent of stage and grade of disease. Patients derive no symptomatic relief with the standard of care UDCA or OCA, and the symptom can persist even after liver transplantation (15). Pruritus often occurs in PBC patients independent of disease stage, presents in up to 78% of the cases (16), and similar to fatigue, it does not improve with UDCA or OCA treatments (3, 17).

Other common manifestations present in PBC patients, in order of prevalence, are Sjogren's syndrome (SS), thyroid disorders, and systemic sclerosis (SSc) (14). SS is characterized by a progressive decrease of lacrimal and salivary glands excretions, resulting in ocular and oral dryness (14). This syndrome coexists with PBC in up to 73% of the patients (18), being present in only 0.2% - 3% of the general population (19). Thyroid diseases are present in up to 32% of PBC patients (20), in which the most common subtypes are Hashimoto's thyroiditis (the most frequent cause of hypothyroidism) in up to 12.5% of PBC patients and Grave's thyroiditis (the most common cause of hyperthyroidism) in around 2% of PBC patients (21). Finally, SSc, manifested by skin thickening and internal organs fibrosis, is associated with PBC in up to 12.3% of the patients (22).

1.1.2 Genetic and environmental factors in PBC

Although the etiology of PBC is not yet well elucidated, it is believed to be a multifactorial disorder caused by the association of genetic predisposition with environmental triggers (23). Genetics seems to play an important role in the development of the disease, which is evidenced by familial clustering and research showing the higher prevalence of the disease among first-degree relatives (24, 25). Twin studies, in which at least one individual has a positive diagnose for PBC, show one of the highest pairwise concordance rates in an autoimmune disorder, of up to 63% in monozygotic twins (26). Other studies show that PBC patients' siblings have a relative risk of 10.5% (27); and, a questionnaire-based study showed that there is a PBC occurrence of 5.9% in first-degree relatives, most frequently in sisters (4.3%) and mothers (1.7%) (28).

Research indicates that the human leukocyte antigen (HLA) class II genes also play a role in increasing susceptibility or conferring protection to PBC. More specifically, the DRB1*0801,

DRB1*0803, DRB1*14 and DPB1*0301 alleles are identified as susceptible; and DRB1*11 and DRB1*13 are identified as protective alleles (29).

Geographical clustering and disparity are strong indications of environmental triggers in PBC development (3, 25). There are several probable exterior causatives for the disorder. Amongst the non-genetic factor that may play a role, there are both xenobiotic and infectious agents (30).

Proposed xenobiotic agents are based on epidemiological, *in vitro,* and *in vivo* data. Studies report that the use of nail polish is linked with a higher PBC susceptibility (28). There are also indications that many chemicals commonly found in daily life items, such as cosmetics and chewing gum, can modify PDC-E2, generating immunogenic neoantigens, which may lead to a higher risk of developing the disease (24). Geographical clustering of PBC cases near toxic waste sites indicates that it might be a risk factor to PBC (31). Another indication of xenobiotic influence on PBC is the existence of murine models in which the disease was induced by xenobiotic immunization (24).

Regarding the role of infectious agents, several studies have associated different organisms as possible triggers for PBC, including bacteria and viruses. Numerous studies have focused on the role of infections caused by *Escherichia coli* and *Novosphingobium aromaticivorans*. Research has shown possible cross-reactivity between human PDC-E2 and antigens from these microorganisms, which could lead to the development of autoimmunity after a previous infection (32). Finally, there is also evidence that a Human Betaretrovirus (HBRV) might be causative of the disease (33).

1.2 HUMAN BETARETROVIRUS (HBRV)

A viral trigger for PBC was first suggested by a study showing seroreactivity of PBC patients against retroviral antigens (34). A Human Betaretrovirus was then identified in the biliary epithelium and lymph nodes of PBC patients through electron microscopy and cloning of exogenous retroviral nucleotide sequences. This virus strongly resembles the Mouse Mammary Tumour Virus (MMTV), with a 95 – 97% nucleotide sequence homology (33).

Subsequent studies showed frequent unique HBRV integration in biliary epithelial cells of PBC patients (Figure 1) but not in patients with non-cholestatic liver disorders (35). The full-length proviral HBRV DNA was cloned from patients with PBC (36). A recent study also revealed that 10% of PBC patients have serological reactivity to HBRV surface protein (37).

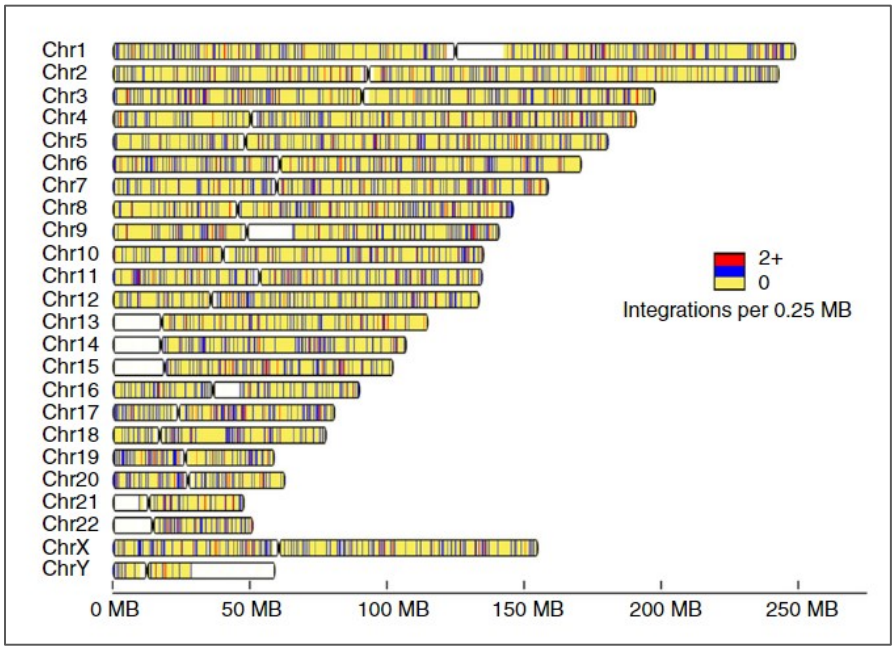

Figure 1: HBRV integration sites. Viral integrations were mapped to the human genome 19. Blue stripes mean single integrations, red stripes are more than two integrations per 250 kb, yellow indicates the known human chromosomal regions and white represents the unsequenced regions (35).

A double-blind, randomized controlled trial studied the efficacy of combination therapy with tenofovir-emtricitabine (TDF/FTC) and lopinavir-ritonavir (LPRr) drugs used for Human Immunodeficiency Virus (HIV) treatment, in PBC patients. Although enrolment was closed early due to lack of tolerance to LPr, treatment was linked with biochemical hepatic improvement and reduction of the viral load, indicating that patients benefit from HBRV suppression (38).

1.3 SPLICING

The process of maturation of a precursor messenger RNA (pre-mRNA) into a messenger RNA (mRNA) involves a process called RNA splicing, which removes non-coding intervening sequences (introns) and connects the coding regions (exons) (39, 40). This process happens in the nucleus of higher eukaryotic cells and occurs in large particles composed of small nuclear ribonucleoproteins (snRNPs), called spliceosomes, being essential for the production of proteins (40-42).

There are some highly conserved sites in the splicing process, with nucleotide sequences showing minimal variance. The sites where the exons and introns are joined are called splice sites (39, 41). In these regions, the introns usually follow the GU-AG rule, in which the 5’ end contains one guanine and one uracil, and the 3’ end has one adenosine and one guanine (39). Branch point A is another region that varies very little. It is an adenosine residue situated between 15 and 45 nucleotides upstream of the 3’ end. The recognition of these sequences is essential for starting the splicing process, but this is not well described (39).

The splicing of introns in a primary transcript requires two trans-esterification reactions, which happen successively and overall in a 5’ to 3’ order (41, 43). The first reaction involves

removing introns by cleavage at the splice sites, in which the 5' end is "attacked" by the branch site A's 2' OH (39, 41). This reaction cleaves the intron, forming an intron lariat, connecting its 5' end to the branch site A (41). In the second trans-esterification reaction, the 3'OH of the 5'end of one exon attaches to the 3' end of another exon, resulting in the mature mRNA molecule (39, 41). This complex process is summarized in Figure 2, extracted from Newman (1998) (41).

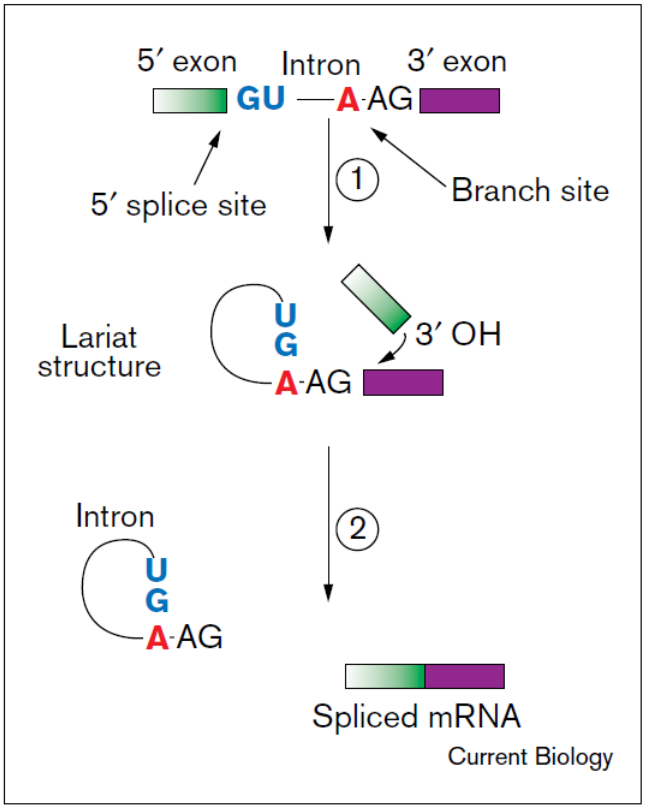

Figure 2: mRNA splicing process. The image shows the two trans-esterification reactions involved in the splicing of an immature mRNA molecule. The number 1 indicates the first reaction, in which the 2'OH of the adenosine (red) at the branch site near the 3' end of the intron attacks the 5' GU (blue) splice site, releasing the 5' exon (green) and leaving the 5' intron joined to the branch site adenosine, in the form of a lariat. The number 2 indicates the second reaction, in which the 3'OH of the 5'exon intermediate attacks the 3' intron splice site, splicing out the intron (41).

There are two types of splicing, constitutive and alternative. In the first one, most exons are ligated in the order they are in a gene. In the alternative splicing, some exons are skipped and

not joined, resulting in different mature mRNA molecules (44), which explains why humans have a much larger number of different proteins than protein-coding genes (44, 45).

1.4 HOW VIRUSES AFFECT SPLICING AND OTHER CELL PROCESSES

It is well established that viruses can modulate gene expression in host cells to enhance their own replication or escape/suppress antiviral responses (40). As such, viruses can impact several different processes, including protein expression, mRNA processing, cellular transcription and protein translation (40, 46-48). Various viruses that replicate both in the nucleus and in the cytoplasm have been shown to manipulate the host splicing machinery (40). The components of the splicing apparatus to be most targeted by viruses are serine/arginine-rich proteins (SR), small nuclear ribonucleoproteins (snRNPs), and heterogeneous nuclear ribonucleoproteins (hnRNPs) (40).

The mechanisms by which viruses modulate splicing vary. In general, the splicing alterations resulted from viral infection are caused by a combination of two complex processes: (i) direct effect, caused by viral manipulation of the splicing machinery, or (ii) indirect effect, triggered by virus-induced cellular damage (40) (Figure 3, modified from Ashraf et al., 2019 (40)). Examples of direct virus-induced splicing alterations include exon skipping (49) and intron retention, leading to the production of alternate protein isoforms, retention of RNA in the nucleus or cytoplasm, and nonsense-mediated decay (49, 50). In comparison, indirect splicing alterations caused by viruses may cause downstream effects such as apoptosis (51), immune responses to neoantigens (52), and DNA damage response (53). Accordingly, cellular perturbation to the

splicing machinery may have a critical effect and can contribute to disease development associated with viral infection (54).

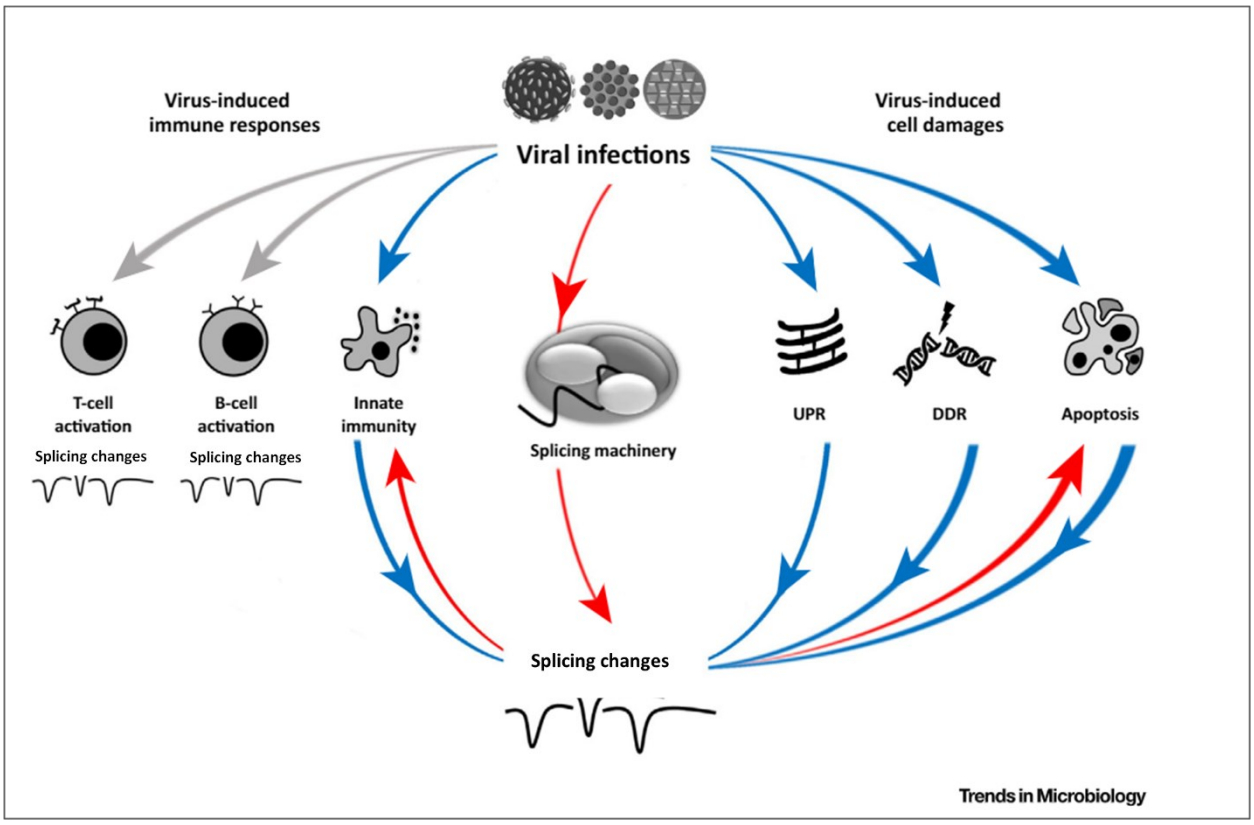

Figure 3: Mechanisms of virus-induced splicing modifications (40). Grey arrows represent virus-induced B cell and T cell immune responses. Red arrows indicate viral manipulation of the splicing machinery and the potential of these changes to modulate innate immunity. Blue arrows represent virus-induced cell damage and innate immune responses, and subsequent splicing changes. UPR: unfolded protein response; DDR: DNA damage response.

Besides interfering with splicing, viruses can disrupt other cellular processes. Viruses have developed diverse mechanisms as a strategy to use the host cell cycle in their favour. Some of the most common examples include the deactivation of a housekeeping protein function following viral infection with the hijacking of biosynthetic apparatus to produce viral proteins, which delays the host cell division (55).

1.5 OMICS STUDIES

Omics studies are helpful tools to identify and measure the diversity of different kinds of biomolecules in blood, tissues, and body fluids that may lead to hypothesis generation (56, 57). Multi-omics approaches include, but are not limited to, genomics, transcriptomics, proteomics and metabolomics (56). These methods have been broadly used in biomedical research for many years for identifying genes, proteins and low-weight molecules implicated in the development of diseases (58).

At the DNA level, genomics techniques study the entire genome of a specific organism (59). These techniques are powerful tools for discovering biomarkers, assessing the number of genes in a particular organism, and decoding DNA information and regulation (59, 60). Genomic methodologies have evolved since the application of DNA sequencing. Currently, next-generation sequencing technologies, including real-time single-molecule sequencing and pyrosequencing, are very popular and have the benefits of being more accurate, faster and cheaper than first-generation sequencing (59).

At the RNA level, transcriptomics technologies methodologies are divided into two main categories: one based on hybridization and another based on sequencing methods (56). These are used to analyze the whole set of transcripts for a specific cell line or tissue and reflect characteristics of the genome under specific physiological conditions (56, 59). Transcriptomics can provide insights into functional and transcriptional structural features of genes, splicing and pre-translational modifications, and differences in gene expression, helping understand the process of disease development (59).

Proteomics analysis is a complex high-throughput method to qualify and quantify protein abundance, variation and interaction in specific cells or tissues under sicken or healthy conditions

(56, 61, 62). These analyses are helpful for elucidating normal or dysfunctional splicing and post-translational modifications, assessing protein interaction and protein expression under different conditions (59). Proteomics techniques involve three primary stages: protein or peptide lysis and separation (classic unbiased methods or affinity purification methods), individual protein data acquisition by mass spectrometry and bioinformatics analysis (61, 62).

Finally, at the level of the metabolites, metabolomics studies are emerging tools for analyzing low molecular weight metabolites secreted in several metabolic pathways (56, 58, 63). These analyses usually use nuclear magnetic resonance, gas chromatography-mass spectrometry, and liquid chromatography-mass spectrometry, which can be used to understand a disease mechanism, assess biological pathways and discover new disease biomarkers (56, 58, 59).

Despite the importance and utility of the multi-omics approach, the use of this technique is limited. For example, there is a lack of omics studies in autoimmunity and primary biliary cholangitis. In contrast, in virology, omic studies are being used to provide a global view of virus-induced changes in different cellular processes to investigate how this might affect the disease pathogenesis and viral life cycle (40).

1.6 HYPOTHESIS

Our current understanding of PBC pathogenesis is still limited. Work in our lab has implicated a human betaretrovirus as a trigger for disease development as it alters mitochondrial protein expression that triggers autoimmunity (33). A previous study from our lab has indicated decreased expression of splicing proteins in the BEC of PBC patients. Based on this and on the knowledge that viruses can affect several cellular processes, we hypothesize that a systemic global dysregulation is present in Primary Biliary Cholangitis patients and potentially related to the

disease pathogenesis and viral infection. Thus, our aims are (i) to analyze dysregulated proteins and transcripts expression in the BEC of PBC patients; (ii) to analyze transcriptional differences in the peripheral blood of PBC patients and healthy volunteers; (iii) to analyze the differences of splicing in PBC and healthy individuals and its the role in disease progression.

1.7 IMPLICATIONS AND IMPORTANCE OF INVESTIGATION

To date, a limited number of hypothesis-generating, multi-omics studies describing diffuse changes in metabolism, splicing, protein translation and disruption of other housekeeping cellular processes have been described in patients with PBC. Studies such as this one serve to provide a molecular landscape that can be used to further build on the pathogenesis, viral or otherwise, in patients with PBC.

2. MATERIAL AND METHODS

2.1 PROTEOMICS

2.1.1 Sample collection and preparation

Biliary epithelial cells (BEC) samples from three PBC patients and four patients with other hepatic disorders (two cryptogenic cirrhosis, one autoimmune hepatitis and one alcohol-related cirrhosis) were obtained. Each patient gave written informed consent. The study protocol conformed with the ethical guidelines of the 1975 Declaration of Helsinki and was reviewed by the ethics review board at the University of Alberta (Pro00005105, Pro00085859).

2.1.1.1 BEC isolation

Using a protocol modified from established methods (64), BECs were extracted from liver tissue obtained from the explanted livers of end-stage liver disease transplant patients. 50-150 g of liver tissue was excised from the explanted whole liver and submerged in RPMI 1640 media (*Gibco, Waltham, MA, USA*). The tissue was either immediately used or stored at 4 °C for up to 24 hours before performing cell extractions. Liver tissue was diced into a paste with scalpels and then digested enzymatically with collagenase from *Clostridium histolyticum* (1-2 mg/mL, *Sigma-Aldrich, Oakville, ON, Canada*) for 20-30 minutes at 37 °C. The tissue slurry was then strained over a sterile mesh screen (*Sigma-Aldrich, Oakville, ON, Canada*) to isolate detached cells. The cell containing flow-through was then distributed to 50 mL conical tubes (*ThermoFisher Scientific, Waltham, MA, USA*) and centrifuged (2000 RPM, 5 minutes, 4 °C). Supernatants were then

decanted, and pellets were resuspended in PBS before combing two pellets into one tube. Tubes were then centrifuged (2000 RPM, 5 minutes, 4 °C), and the process was repeated until only two pellets remained. Cells were then semi-purified by density gradient centrifugation (2000 RPM, 30 minutes, no brake, room temperature) on a 33%/77% Percoll column (*ThermoFisher Scientific, Waltham, MA, USA*). Columns were made by underlaying 3-194 mL of 77% Percoll underneath 3-4 mL of 33% Percolland, then topped with 3-4 mL of cells resuspended in PBS. The BEC-containing interphase was collected, brought up to a volume of 50 mL with autoMACS running buffer (*Miletenyi Biotec, Auburn, CA, USA*) and centrifuged (2000 RPM, 5 minutes, 4 °C). The pellet was then resuspended in 500 µL of autoMACS running buffer supplemented with 60 µL of magnetically labelled antibodies that bind to the BEC specific cell surface marker, CD326 (*Miletenyi Biotec, Auburn, CA, USA*). Cells were left to incubate with the antibody at 4 °C for 20-30 minutes. During this incubation, MACS columns were prepared by running 3 mL of autoMACS running buffer over the column. The cell-antibody mixture was then put through a 30 µM pre-separation (*Miletenyi Biotec, Auburn, CA, USA*) or a 40 µM cell strainer (*ThermoFisher Scientific, Waltham, MA, USA*) before being added to the MACS column. The column was washed three times with autoMACS buffer. 5 mL of autoMACS buffer was added to the column, and the cells were plunged onto a second MACS column. The column was again washed three times with autoMACS buffer. Finally, 3 mL of BGM was added to the column and cells were plunged into a collagen-coated T-25 flask and stored in a cell incubator (37 °C, 5% CO2).

2.1.2 Shotgun Proteomics

Intrahepatic BECs from the seven patients were processed in duplicate for shotgun-proteomic analyses using liquid chromatography-tandem mass spectrometry (LC-MS/MS). Dr. B.

Meng performed lysate extractions. The BECs were washed four times with PBS before the addition of 400 µL of 2D lysis Buffer (7 M urea, 2 M thiourea, 4% chaps, 30 mM Trish-HCl pH 8.8, 1 M PMSF, 0.5 mM EDTA, 1 mM DTT, EDTA-free Halt protease inhibitor cocktail, (*ThermoFisher Scientific, Waltham, MA, USA*). Cells were scraped off the plate on ice, followed by rigorous pipetting to help facilitate lysis. Lysates were spun at 12000 xg for 10 minutes and stored at -80 °C. Protein concentration was quantified using a Bradford Assay. 25 µg of protein was loaded for PAGE separation. Gels were then given to the University of Alberta Proteomics Core (Dr. R. Fahlman's lab) for LC-MS/MS.

The Proteomics core facility performed the primary analysis of LC-MS/MS to determine peptide spectral counts as a quantitative measure of protein expression. Dr. W. Wang and Dr. B. Meng performed normalization and statistical analysis using the DanteR software package. For statistical analysis, normalized spectral counts were compared between PBC and non-PBC BEC. P-values less than 0.05 were considered significant. Proteins found to be differentially expressed were then separated into upregulated and downregulated lists of UniProt accession numbers.

2.1.3 Pathway analysis

The pathway analysis was made through STRING® (65) (available on string-db.org/). The input was a list of 46 UniProt Knowledgebase (UniProtKB) accession numbers (one for each protein) for the *Homo sapiens* database. The active interaction sources chosen were "Experiments" and "Databases", with a minimum required conference score of 0.7 (high confidence).

2.2 TRANSCRIPTOMICS

2.2.1 Whole Blood Sample collection and stratification

Serial samples from whole blood of 128 PBC patients enrolled in the POISE study, *A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis* (66), was collected in PAXgene® tubes (*Becton Dickinson, Franklin Lakes, NJ, USA)*. In total, 216 samples from PBC patients and 13 healthy controls were compared. Out of the total number of enrolled subjects, 86 had single timepoints sample collection, 36 patients had at least one baseline sample collection and one sample collection after the start of the treatment, and six patients did not have a sample collected at baseline.

Following the POISE study endpoint criteria (66), the patients' samples were identified as (i) "Responders" if the ALP level was <1.67 times the upper limit of the normal range, or with a 15% decrease in ALP and normalized BLB; or (ii) "Non-responders" to OCA treatment if ALP level was > 1.67 times the upper limit of normal at the end of the 12-months treatment period. The samples were annotated "Before" if baseline sample prior to treatment and "After" if collected at the appropriate time point after collection following 12 months of OCA treatment. Furthermore, we selected the 15 patients with the highest ALP levels in the Non-responders group (ALP > 400 U/L) and the 15 patients with the lowest ALP levels in the Responders group (ALP < 240 U/L) to increase the biochemical discrimination between each group.

2.2.2 RNA isolation

The RNA isolation of the 128 PBC patients and 13 healthy controls was made using the E.Z.N.A® PX Blood RNA Kit (*Omega Bio-Tek, Norcross, GA,* USA) and following its protocol. The PAXgene® Blood RNA tubes (*Becton Dickinson, Franklin Lakes, NJ, USA)* were centrifuged for 10 minutes at 4000 xg. After this, the supernatants were aspirated and discarded, and the pellets were resuspended with 4 mL of RNase-free water. The tubes were centrifuged again at the same speed and for the same duration. The supernatants were discarded, and the pellets were resuspended with 485 µL Nuclease-free water and transferred to a 1.5 mL microcentrifuge tube (*Corning Life Sciences*, Tewksbury, MA, USA). Then, 375 µL NTL Lysis Buffer and 40 µL Proteinase K Solution were added, vortex for 5 seconds and incubate for 15 minutes at 55 °C. The samples were transferred into homogenizer mini columns and centrifuged at 13000 xg for 3 minutes. The supernatants were then transferred to new 1.5 mL microcentrifuge tubes followed by the addition of 0.5 volumes of 100% ethanol and transfer into the HiBind® RNA mini columns and centrifuged at 13000 xg for 1 minute. The filtrates were discarded, and the columns were inserted into new 1.5 mL microcentrifuge tubes. 350 µL RWF buffer was added onto the columns and centrifuged at 13000 xg for 1 minute. The filtrates were discarded, and 75 µL of DNase I digestion reaction mix (DNase I digestion buffer and RNase-free DNase I) was added directly onto the surface of the HiBind® RNA mini columns for 15 minutes at room temperature. After this, 500 µL RWF buffer was added to the columns and then centrifuged at 13000 xg for 1 minute. After discarding the filtrates, 500 µL of RNA Wash Buffer II was added and centrifuged at 13000 xg for 1 minute, and the filtrates were discarded (this step was repeated twice). To dry out the HiBind® RNA mini columns, they were centrifuged empty at 13000 xg for 2 minutes. 60 µL of RNase-free

water was added onto the membranes and then spun at 13000 xg for 2 minutes to extract the RNA from the columns. The RNA was stored at -70 °C until further use.

2.2.3 Next-generation sequencing

After RNA isolation, the samples were sent to Novagene for NGS sequencing. First, ribosomal RNA and globin RNA was depleted from total RNA and then cloned into TruSeq libraries (*Illumina, San Diego, CA, USA*). Each library was processed by Illumina HiSeq (*Illumina, San Diego, CA, USA*) to generate 5 GB of data per library. Dr. Juan Joavel and Hussain Syed conducted differential expression analysis using the DESeq2 R package. Genes deregulated with a false discovery rate of ≤ 0.01 were considered differentially expressed. A differential expression score was calculated based on total increased or decreased reads per patient.

To create a score for correlating changes in splicing gene expression with disease prognosis, we selected the 46 genes present in the splicing pathway that were downregulated in the non-responder group. We compared the log-transformed TPM data of each transcript against the ALP and total BLB levels for all responder and non-responder patients in a linear regression analysis. Subsequently, the individual scores were compared for responders versus non-responders by both an unpaired t-test and a receiver operating characteristic (ROC) curve statistic. Then the aggregate log-transformed TPM scores of transcripts associated with prognosis were combined and re-assessed by an unpaired t-test and a receiver operating characteristic (ROC) curve statistic to create a predictive score.

2.2.4 Differential exome usage analysis

Alignments against the GRCh38 version of the human genome were conducted with TopHat2 (67). Initially, read counts mapping to genomic intervals defined in the human ensemble GTF (General Transfer Format) file were extracted with the script dexseq_count.py, which is included in the DEXSeq R package distribution (68) and internally uses HTSeq (69). Differential exon usage was tested with DEXSeq (68). Essentially, for each partially or entirely covered exon and for each sample, the number of reads mapping to each such feature was counted, as well as the number of reads mapping to any of the other exons in the same gene. The ratio between these two counts was compared between conditions by conducting a Fisher's exact test (68).

2.2.5 Biliary Epithelial Cell RNAseq

RNAseq libraries were constructed using the TruSeq RNA Library Prep Kit v2 (*Illumina, San Diego, CA, USA*) from three PBC and four non-cholestasis control samples using 100 ng of total BEC RNA. Each library was sequenced using a HiSeq 2500 system (*Illumina, San Diego, CA, USA*), following a paired-end 150 cycles protocol, including in-instrument demultiplexing, as described.

For bioinformatic analyses, fragments were mapped to the human cDNA database (GRCh38) using Kallisto, with 100 permutations during pseudo-alignments and using bias correction. Differential expression analysis of RNAseq data was conducted using negative binomial generalized linear models with the DESeq2 R package. Gene abundance differences with a corrected p-value ≤ 0.05 were considered differentially expressed.

2.2.6 Pathway analysis

The pathway analyses were made through Gene Ontology (70, 71) using the transcripts Ensembl IDs from the differential expression analyses. Only the protein-coding transcripts were used for the analysis. Each was selected by transforming the transcripts Ensembl IDs to proteins Ensembl IDs; the sequences that did not have protein Ensembl ID were non-protein-coding transcripts. The list of transcripts was then uploaded on Gene Ontology (GO) (http://geneontology.org/) under the "Biological Process" tab.

3. RESULTS

A multi-omics analysis is a helpful approach for investigating disease pathogenesis (58). For example, transcriptomics is a valuable tool for identifying functional and structural characteristics of genes and differences in gene expression (59, 60). Complimentary, proteomics studies are useful when evaluating protein interaction and post-translation modifications under specific conditions and among different patient groups (59).

Our lab first decided to use this approach to identify possible abnormalities in mitochondria and metabolism in the BEC of PBC patients. The findings were addressed in Filip Wysokinski's thesis and paper (72), where he explained how the metabolic remodelling in mitochondria might account for changes linked with the development of autoimmunity and production of anti-mitochondrial antibodies in relation to PBC pathogenesis.

Although unexplored at the time, these experiments also suggested potential dysregulation in essential splicing proteins in the BEC of PBC patients. Because of the central role of viral infection in disrupting splicing and RNA maturation (73-81) and our lab's interest in studying viral pathogenesis in PBC (33-38), we decided to investigate further. Accordingly, we evaluated our proteomic and transcriptomic datasets for differences in splicing in the available body compartments, including BEC and whole blood, which we studied because PBC is considered a systemic disease. Our results not only confirmed splicing dysregulation but also indicated disturbance in other essential cellular processes in both BEC and blood of PBC patients. These disrupted pathways are explained in detail in the following pages.

3.1 BILIARY EPITHELIAL CELL PROTEOMICS

As mentioned, we decided to use proteomic analysis of biliary cells extracted from the explants of liver transplant recipients to address potential anomalies in the splicing process of PBC patients. For that, we started with a differential protein expression analysis using three PBC and four non-biliary diseases BEC run in duplicate. This analysis showed that 53 proteins were upregulated (p-value ≤ 0.047), and 46 proteins were downregulated (p-value ≤ 0.048) in the BEC of PBC patients when compared to control samples. The pathway analysis of these differentially expressed proteins on STRING® (65) showed a total of 253 single pathways upregulated in the BEC of PBC patients (*vs* control) and 157 single pathways downregulated in the BEC of PBC patients (*vs* control). Because of the sizable number of pathways, we selected the ones with a Fold Enrichment ≥ 1 with the lowest FDR from both analyses to further investigate (Figure 4).

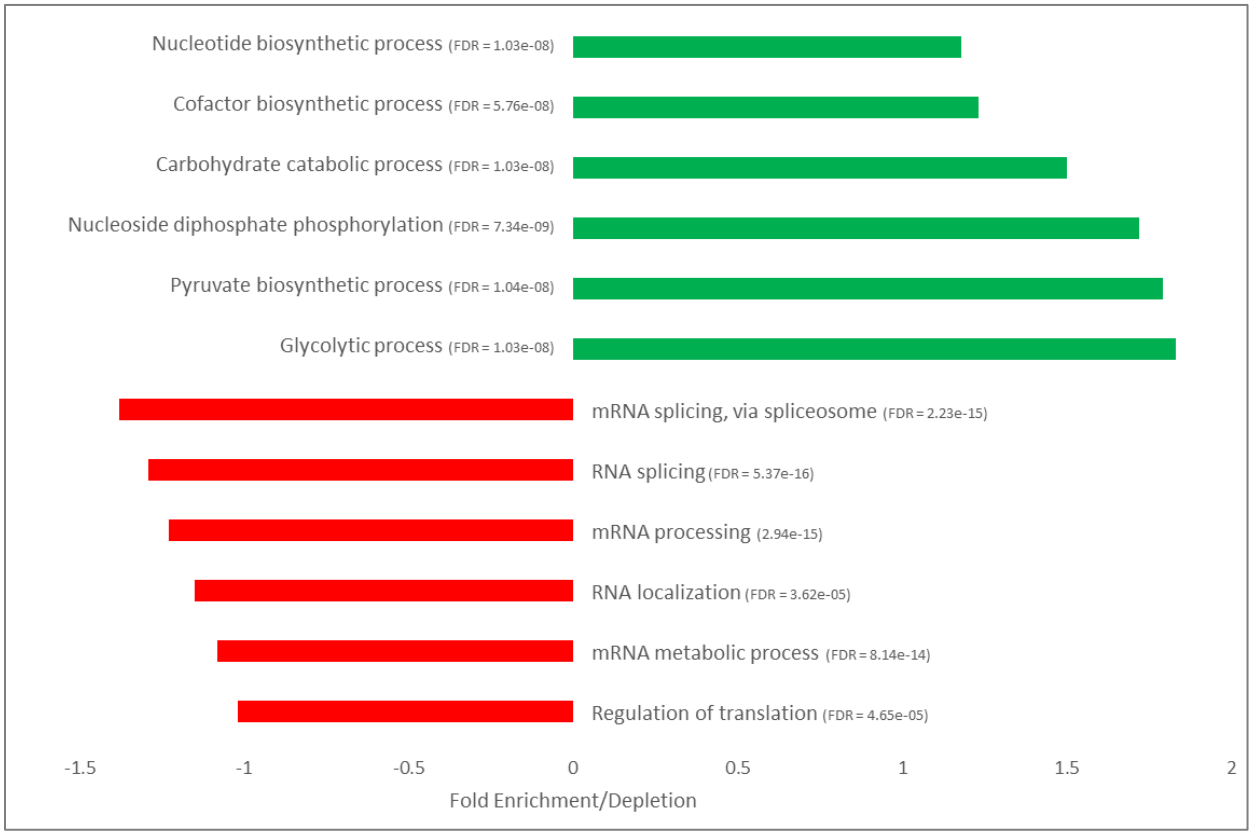

Figure 4: Diminished pathways detected in proteomic study related to splicing, according to STRING® (65). Red represents the pathways downregulated in the BEC of PBC patients (*vs* control), and green represents the pathways upregulated in the BEC of PBC patients (*vs* control).

As indicated by Figure 4, "Biosynthetic processes" pathways ("Nucleotide biosynthetic process", "Cofactor biosynthetic process", and "Pyruvate biosynthetic process"), "Glycolytic process", "Carbohydrate catabolic process" and "Nucleotide diphosphate phosphorylation" are the pathways with the lowest FDR and highest fold enrichment upregulated in the BEC of PBC patients. Further evidence of aerobic glycolysis and abnormal mitochondrial signalling was investigated and confirmed in functional studies using cultured BEC by Filip Wysokinski (72). These findings suggested that PBC BEC were undergoing metabolic remodelling consistent with Warburg metabolism, with hyperpolarized mitochondria and mitochondrial biogenesis. These data

support the central hypothesis that compensatory production and cell surface expression of pyruvate dehydrogenase may lead to loss of tolerance to self and production of AMA (7, 33).

In contrast, the downregulated pathways with the lowest FDR and higher fold depletion in the BEC of PBC patients *vs* control subjects included the following GO terms: "Regulation of translation", "mRNA metabolic process", "RNA localization", "mRNA processing" and splicing related pathways ("mRNA splicing, via spliceosome" and "RNA splicing"). These data suggest a hypothesis of major downregulation of essential cellular processes related to RNA metabolism.

Interestingly, splicing-related pathways were the most significant pathways downregulated in the BEC of PBC patients *vs* control subjects. The splicing machinery is crucial for the process of mRNA maturation leading to protein synthesis. Several splicing pathways appear downregulated in the BEC of PBC patients. Out of the 46 proteins downregulated in the BEC of PBC patients, 18 are present in splicing pathways (39.13% out of the total of downregulated proteins). Therefore, due to the high number of splicing proteins and pathways identified in our analysis, we decided to examine this process in more detail.

The proteomics analysis showed seven splicing pathways downregulated in the BEC of PBC patients *vs* controls with the following GO terms: "RNA splicing, via transesterification reactions", "RNA splicing", "mRNA splicing, via spliceosome", "regulation of RNA splicing", "negative regulation of RNA splicing, via spliceosome", "regulation of RNA splicing, via spliceosome" and "regulation of alternative mRNA splicing, via spliceosome" (Table 1). Of note, these processes overlap because of the involvement of multiple proteins in different but related pathways.

Pathways	GO term	Proteins	FDR
RNA splicing, via transesterification reactions[1]	GO:0000375	ALYREF, DHX9, ELAVL1, hnRNPA2B1, hnRNPA3, hnRNPC, hnRNPD, hnRNPH1, hnRNPH3, hnRNPK, hnRNPL, hnRNPU, KHSRP, NHP2L1, NONO, PCBP2, PTBP1	2.17e-16
RNA splicing[2]	GO:0008380	NONO, hnRNPA2B1, ELAVL1, hnRNPH3, hnRNPL, NHP2L1, hnRNPA3, hnRNPD, PCBP2, PTBP1, ALYREF, TARDBP, hnRNPU, hnRNPK, hnRNPC, DHX9, KHSRP, hnRNPH1	5.37e-16
mRNA splicing, via spliceosome[3]	GO:0000398	ALYREF, DHX9, ELAVL1, hnRNPA2B1, hnRNPA3, hnRNPC, hnRNPD, hnRNPH1, hnRNPH3, hnRNPK, hnRNPL, hnRNPU, KHSRP, NHP2L1, NONO, PCBP2, PTBP1, TARDBP	2.23e-15
Regulation of RNA splicing[4]	GO:0043484	AHNAK, hnRNPA2B1, hnRNPH1, hnRNPK, hnRNPL, hnRNPU, PTBP1	1.32e-06
Negative regulation of mRNA splicing, via spliceosome[5]	GO:0048025	hnRNPA2B1, hnRNPK, hnRNPL, PTBP1	2.03e-05
Regulation of mRNA splicing, via spliceosome[6]	GO:0048024	hnRNPA2B1, hnRNPK, hnRNPL, hnRNPU, PTBP1	5.73e-05
Regulation of alternative mRNA splicing, via spliceosome[7]	GO:0000381	hnRNPL, hnRNPU, PTBP1	0.0023

[1]Splicing of RNA via a series of two transesterification reactions (70, 71).
[2]The process of removing sections of the primary RNA transcript to remove introns and joining of the exons to form the mature structure of the RNA (70, 71).
[3]The joining together of exons from one or more primary transcripts of messenger RNA (mRNA) and the excision of intron sequences via a spliceosomal mechanism so that mRNA consisting only of the joined exons is produced (70, 71, 82, 83).
[4]Any process that modulates the frequency, rate, or extent of RNA splicing (70, 71).
[5]Any process that stops, prevents or reduces the rate or extent of mRNA splicing via a spliceosomal mechanism (70, 71).
[6]Any process that modulates the frequency, rate or extent of mRNA splicing via a spliceosomal mechanism (70, 71).
[7]Any process that modulates the frequency, rate, or extent of alternative splicing of nuclear mRNAs (70, 71).

Table 1: Splicing related pathways downregulated in the BEC of PBC patients (*vs* controls), according to STRING® (65) analysis.

The GO term “RNA splicing” was used as the starting point for our investigation, as it is the broadest and all-encompassing term with the lowest FDR and highest number of differentially expressed proteins highlighted in our analysis (Figure 5).

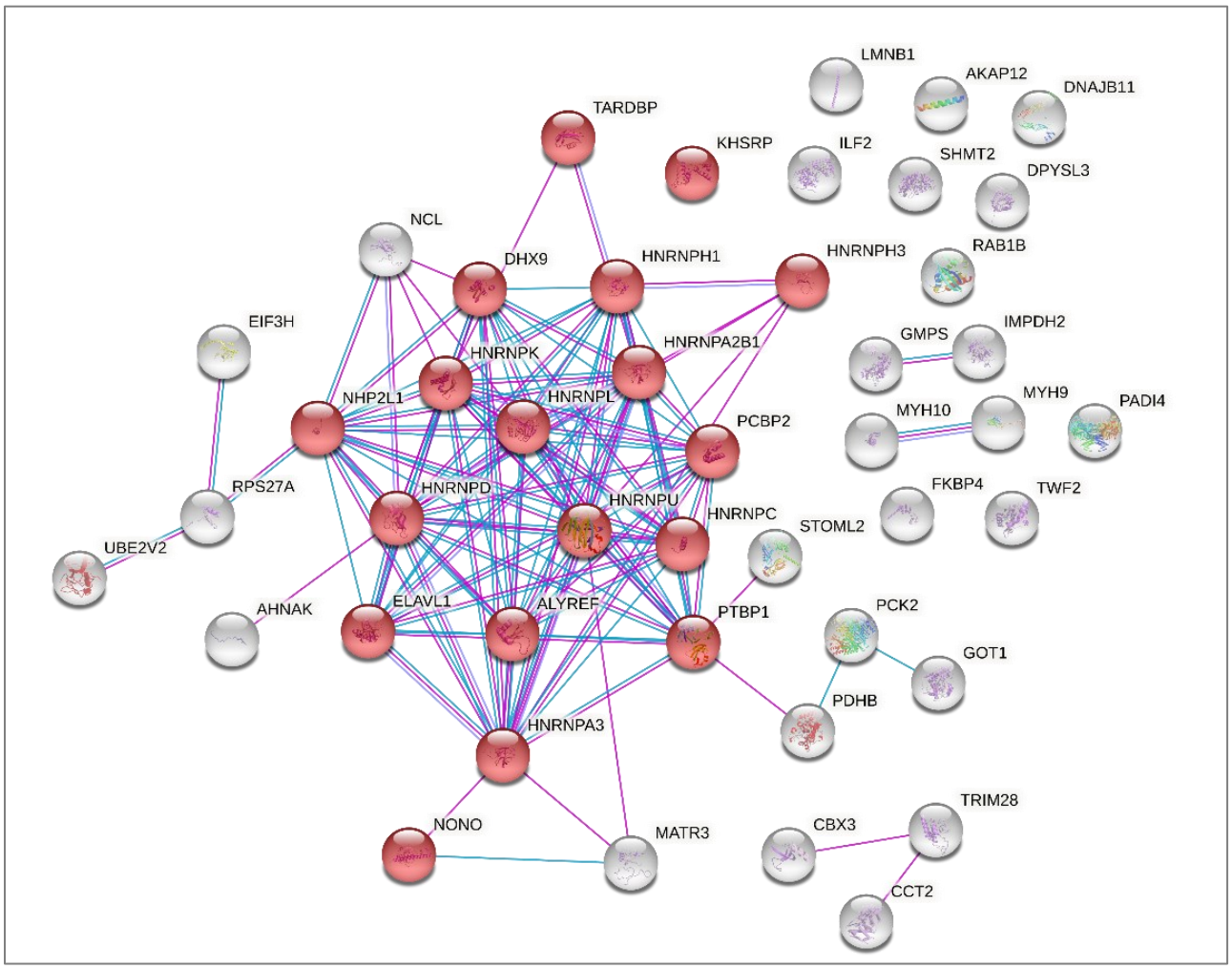

Figure 5: Downregulated proteins in PBC (BEC). The figure is derived from STRING® (65). In red are the proteins that belong to the “RNA splicing” pathway. Blue edges represent interactions found in curated databases, and pink edges represent experimentally determined interactions.

Due to known viral interference with host splicing machinery and our hypothesis of HBRV involvement in PBC pathogenesis, we decided to investigate whether the proteins downregulated in

the BEC of PBC and related to splicing have been linked to viruses in previous studies. The STRING analysis showed that, from the 18 proteins downregulated in PBC patients in the "RNA splicing" pathway, 13 had been described to interact with other viruses and play a role in viral pathogenesis (Table 2).

NONO	hnRNPA2B1	PCBP2	hnRNPK
hnRNPH1	hnRNPL	ELAVL1	PTBP1
hnRNPC	hnRNPD	KHSRP	DHX9
hnRNPU			

Table 2: Proteins downregulated in PBC patients (*vs* control) involved in splicing pathways and previously related to viruses, according to STRING® analysis.

Studies have shown that splicing proteins may modify viral expression in several different ways. For example, some proteins such as hnRNPU and NONO negatively regulate HIV infection (84, 85), hnRNPL negatively regulates the replication of the foot-mouth-disease-virus (86), and KHSRP downregulates enterovirus 71 infection (87, 88). At the same time, others have been reported to enhance viral replication. These include ELAVL1 and hnRNDP, which augment HPV replication (73, 74); hnRNPK, which enhances the Dengue Virus type 2 replication (75); PTBP1, which is required for the picornavirus' gene propagation (76); hnRNPC, that facilitates picornavirus RNA synthesis (77); and PCBP2, which is essential for poliovirus viral translation (77-79).

Relevant to this investigation, both hnRNPA2B1 and hnRNPH1 are modulated *in vitro* by HIV infection with initial downregulation of expression in the first weeks of infection and then subsequently upregulation (80). Furthermore, DHX9 expression is upregulated by the Hepatitis B

virus, enhancing viral replication (89). Taken together with our findings of downregulated splicing protein expression, the collective data point to a hypothesis that RNA splicing may be disrupted in PBC patients, and this process may be linked with HBRV infection. (See discussion and "Table 10" for more detailed information.)

3.2 TRANSCRIPTOMICS

3.2.1 PBC *vs* Non-cholestatic Liver Disease Control (Biliary Epithelial Cell)

To further develop the findings from the proteomics analysis, we took the same unbiased approach to study transcriptomic changes in BEC. The transcriptomics analysis of the BEC from four PBC patients *vs* five non-cholestatic liver controls, using an FDR ≤ 0.05 to indicate differential expression, showed that 172 functional transcripts were overexpressed in PBC patients and 344 functional transcripts were under-expressed in the BEC of PBC *vs* control. We used these transcripts as input for a GO analysis (70, 71), which indicated 10 pathways upregulated and 27 pathways downregulated in the BEC of PBC patients. We observed upregulated pathways consistent with the metabolic remodelling and changes compatible with Warburg metabolism, including "response to hypoxia". To facilitate the visualization of the data, we selected the pathways that better represented our findings regarding RNA splicing and essential cellular processes with the lowest FDR and higher fold enrichment/depletion (Figure 6).

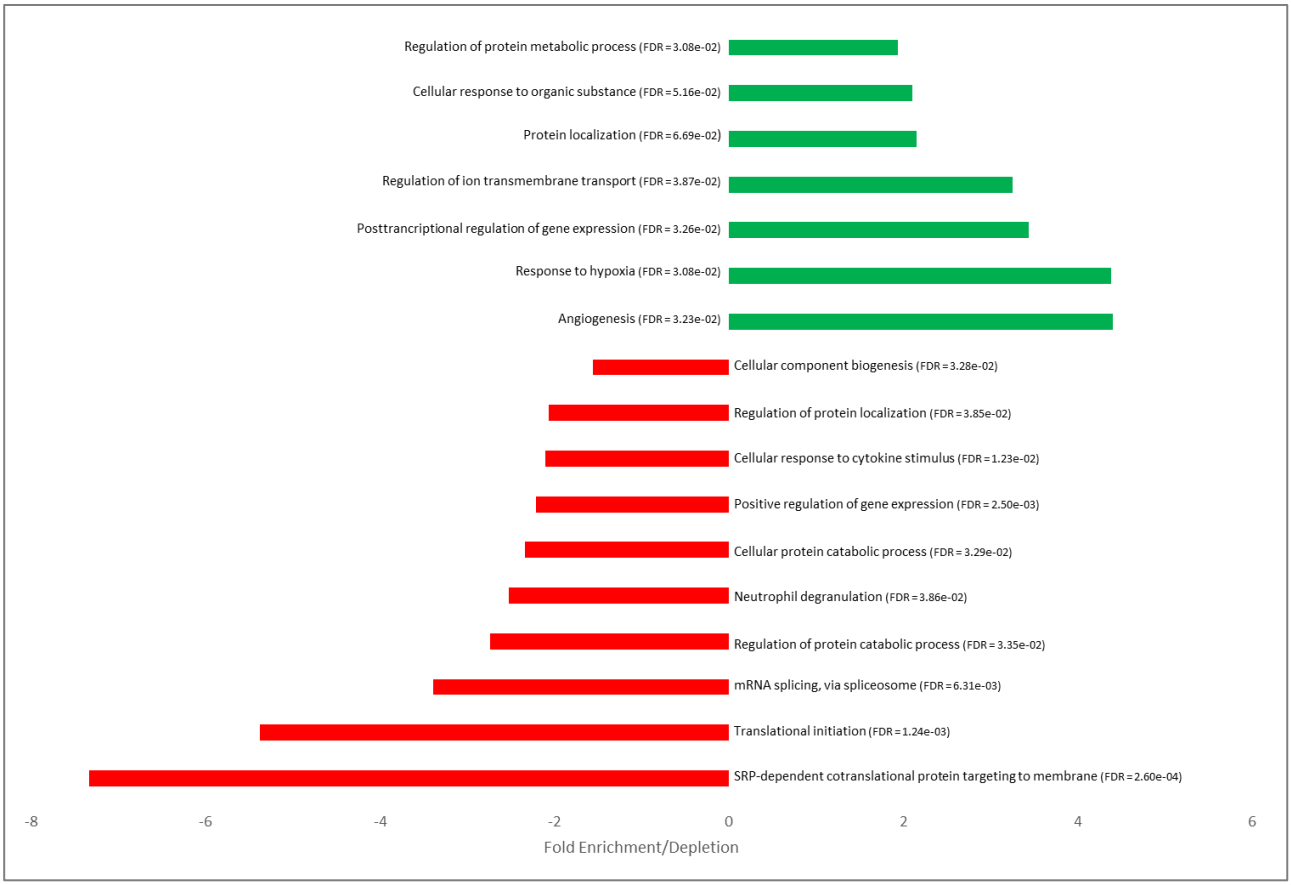

Figure 6: Pathways up and downregulated in the BEC of PBC patients, according to Gene Ontology® (70, 71, 90). Red represents the pathways downregulated in the BEC of PBC patients (*vs* control), and green represents the pathways upregulated in the BEC of PBC patients (*vs* control).

As shown by Figure 6, processes related to mRNA splicing and translation were among the most downregulated pathways at the transcript level within the PBC BEC. It is noteworthy that the transcriptional expression does not directly correlate with protein levels, as post-translational protein modifications impact the rate of protein degradation. Even though the transcripts and proteins abnormally expressed are not necessarily the same, we observed that a few of the dysregulated pathways were highlighted in both analyses. The examples include the metabolic remodelling observed in PBC BEC (72), splicing and other biological processes outlined in Table 3.

Upregulated Pathways	
Transcriptomics	**Proteomics**
“Angiogenesis”	“Angiogenesis”
“Response to hypoxia”	“Response to hypoxia”
“Regulation of ion transmembrane transport”	“Iron ion transmembrane transport”
“Cellular response to organic substance”	“Cellular response to organic substance”
“Signaling”	“Regulation of signalling”
“Cell communication”	“Regulation of cell communication”
Downregulated Pathways	
Transcriptomics	**Proteomics**
“Translational initiation”	“Regulation of translation”, “Positive regulation of translation”, “Negative regulation of translation”
“mRNA splicing, via spliceosome”	“RNA splicing, via transesterification reactions”, “RNA splicing”, “mRNA splicing, via spliceosome”, “Regulation of RNA splicing”, “Negative regulation of mRNA splicing, via spliceosome”, “Regulation of mRNA splicing, via spliceosome”, “Regulation of alternative mRNA splicing, via spliceosome”
“Membrane organization”	“Membrane organization”
“Positive regulation of gene expression”	“Positive regulation of gene expression”, “Regulation of gene expression”, “Negative regulation of gene expression”
“Cellular response to cytokine stimulus”	“Cellular response to cytokine stimulus”
“Positive regulation of cellular protein metabolic process”	“Positive regulation of protein metabolic process”

"Negative regulation of nitrogen compound metabolic process"	"Negative regulation of nitrogen compound metabolic process"
"Negative regulation of cellular metabolic process"	"Negative regulation of cellular metabolic process"

Table 3: Comparison of pathways up and downregulated in transcriptomics and proteomics analysis of the BEC of PBC patients *vs* control.

In concordance with the proteomic data, "mRNA splicing, via spliceosome" is one of the pathways downregulated in the transcriptomics analysis. Indeed, one of the transcripts (PCBP2) downregulated in this pathway was also highlighted in the proteomics study (Table 4).

Pathway	GO Term	Functional transcripts	FDR
mRNA splicing, via spliceosome[1]	GO:0000398	CPSF1, PABPC1, HSPA8, SNRNP70, PRPF40B, PUF60, SRRM2, DHX38, PRPF8, PRPF39, PCBP2, EFTUD2, DDX42, SYMPK, hnRNPR, RBM6	6.31e-03

[1]The joining together of exons from one or more primary transcripts of messenger RNA (mRNA) and the excision of intron sequences via a spliceosomal mechanism so that mRNA consisting only of the joined exons is produced (70, 71, 82, 83).

Table 4: Splicing-related pathway downregulated in the BEC of PBC patients (*vs* control), according to GeneOntology® (70, 71, 90) analysis.

3.2.2 Transcriptomics analyses PBC *vs* Healthy (whole blood RNA)

There is evidence to suggest that PBC is a systemic disease, and the disease process is not just restricted to the liver. Patients commonly develop Sicca syndrome (with dry eyes and mouth) and musculoskeletal symptoms and disease as part of the systemic pathological process. Furthermore, patients can develop incapacitating fatigue, described differently as exhaustion and brain fog in the absence of cirrhosis. Indeed, the persistence of fatigue after liver transplantation without evidence of recurrent disease in the allograft supports the hypothesis of PBC as a systemic disorder (91).

Accordingly, we decided to investigate whole blood at the transcriptomics level to assess whether some of the pathological processes observed in BEC may also be seen in peripheral blood mononuclear cells. This analysis was performed in collaboration with Intercept Pharma, which provided funding and whole blood RNA from serial samples of PBC patients enrolled in a clinical trial. These patients were unresponsive to the standard of care with ursodeoxycholic acid and then randomized to receive either a placebo or the FXR agonist obeticholic acid (OCA). We received samples obtained from patients before and after therapy who had either responded to OCA intervention or failed to sufficiently improve liver function (66). Our initial study assessed the differential expression of transcripts between 87 PBC patients' samples collected before the patients underwent OCA treatment and 13 healthy subjects. Using an FDR ≤ 0.01, we found that a total of 363 unique functional transcripts were upregulated in PBC patients (*vs* healthy), and 385 unique functional transcripts were downregulated in PBC patients (*vs* healthy).

In the Gene Ontology (70, 71) pathway analysis, we input all the functional transcripts (363 transcripts upregulated in PBC and 385 transcripts downregulated in PBC). We found a total of 53 upregulated pathways and 14 downregulated pathways in PBC patients *vs* healthy controls. Because of the large number of pathways, we conducted a more stringent Gene Ontology (70, 71) analysis using as input all the functional transcripts with a log2FoldChange of ≥ 0.5 and ≤ -0.5, accounting for 105 transcripts upregulated in PBC *vs* healthy analysis and 191 transcripts downregulated in PBC *vs* healthy. We found six pathways upregulated in the blood of PBC patients *vs* healthy volunteers and nine pathways downregulated in the blood of PBC patients *vs* healthy volunteers (Figure 7).

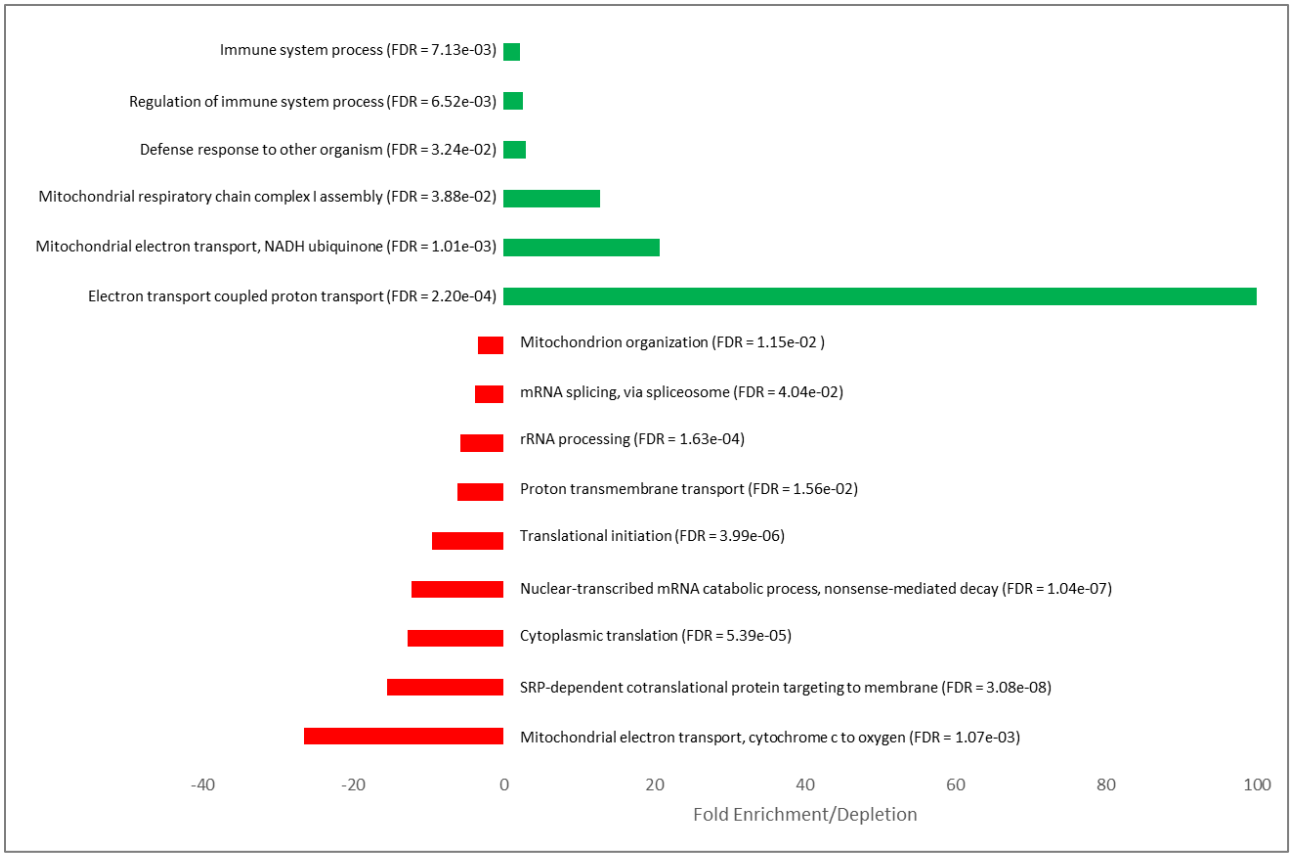

Figure 7: Pathways up and downregulated in the blood of PBC patients, according to Gene Ontology® (70, 71, 90). Red represents the pathways downregulated in the blood of PBC patients (*vs* healthy), and green represents the pathways upregulated in the blood of PBC patients (*vs* healthy).

The transcriptomics analysis of whole blood RNA of PBC patients *vs* healthy controls showed a similar trend to that seen in the proteomics analysis. As indicated by Figure 7, electron transport pathways were paradoxically both upregulated and downregulated in PBC patients. Interestingly, the upregulated transcripts were all mitochondrially encoded genes that belong to the "Mitochondrial electron transport, NADH ubiquinone" and "Electron transport coupled proton transport" (MT-CO1, MT-ND2, MT-ND1, MT-ND4, MT-ND4L, MT-ND5, MT-ND6, MT-CYB). Whereas the downregulated transcripts were all nuclear-encoded genes from the "Mitochondrial electron transport, cytochrome c to oxygen" and "Mitochondrial respiratory chain complex 1 assembly" (NDUFA6, NDUB1, NDUFS5, COX6A1, COX6C, ATP5J, TIMM8B, TOMM7). These results are consistent

with the findings in the proteomic BEC analysis, suggesting that the mitochondrial function is inhibited in PBC patients, and there is evidence of compensatory mitochondrial biogenesis with elevated levels of mitochondrial DNA as reported by Wysokinsky (72). In this study, it was hypothesized that the observed metabolic remodelling with aerobic glycolysis resulted in an inhibition of mitochondrial function, resulted in a compensatory increase in the number of mitochondria (and production of PDC-E2 leading to loss of tolerance to self-proteins). We have suggested a similar hypothesis in the peripheral blood, with the increased expression of mitochondrial genes and decreased nuclear genes encoding mitochondrial proteins currently being investigated using mt-DNA PCR.

We observed similar trends, with the identification of similar pathway changes in the transcriptomics analyses of whole blood as seen with the proteomic and RNAseq analyses in BEC. Specifically, "mRNA splicing, via spliceosome" was one of the pathways downregulated in PBC patients' whole blood RNA (Table 5). This finding is consistent with the idea that PBC is a systemic disease where similar processes, such as metabolic remodelling, were observed in BEC and whole blood. In this example, we observed evidence in both the BEC and whole blood analyses suggesting that the deficiency in splicing is present in the liver and peripheral blood of PBC patients.

Pathway	GO Term	Functional transcripts	FDR
mRNA splicing, via spliceosome[1]	GO:0000398	UBL5, LSM3, LSM7, ZMAT2, PHF5A, SNRPE, MAGOH, GTF2F2, CWC15, SYF2, POLR2K	4.04e-02

[1]The joining together of exons from one or more primary transcripts of messenger RNA (mRNA) and the excision of intron sequences via a spliceosomal mechanism so that mRNA consisting only of the joined exons is produced (70, 71, 82, 83).

Table 5: Splicing-related pathway downregulated in the blood of PBC patients (*vs* healthy), according to Gene Ontology® (70, 71, 90) analysis.

3.2.3 Assessment of splicing in PBC patients in different prognostic categories

From the POISE trial (66), we obtained samples from patients with and without a biochemical response to OCA as judged by serum alkaline phosphatase (ALP) levels, which provides a biochemical determinant of the biliary disease activity referred to as grade of disease. The score also uses bilirubin levels that provide an indication of disease stage, as bilirubin levels rise with loss of bile ducts. The criteria used to determine response to OCA treatment are referred to as the POISE criteria (66): (i) "Responders" if the ALP level was <1.67 times the upper limit of the normal range, or with a 15% decrease in ALP and normalized BLB; or (ii) "Non-responders" to OCA treatment if ALP level was > 1.67 times the upper limit of normal at the end of the 12-months treatment period.

In the first analysis, we compared all the responders (n = 36) with the non-responders (n = 51) before treatment to determine transcripts that may predict response to therapy. In this analysis, only two transcripts were identified as differentially expressed, and both of these are of unknown significance: the ezrin gene (ENST00000392177) and HLA-A-211 (ENST00000550728). The lack of detecting deregulated transcription is likely due to the homogeneity of the samples in PBC patients before treatment, as well as the use of the POISE endpoint criteria. The latter is useful to discriminate responders *vs* non-responders at the end of the study. However, the PBC patients prior to therapy had too similar biochemical profiles to distinguish between the groups and identify prognostic transcripts. Accordingly, we decided to compare transcripts in patients in each group with the lowest and highest ALP levels to increase the likelihood of finding differential transcription patterns. In the new analysis, we selected 15 patients with the highest ALP level in the non-responder group and 15 patients with the lowest ALP levels in the responders group. With this new selection, we found that 1,519 protein-coding unique transcripts were downregulated in non-responders before the commencement of OCA therapy, and another 1,516 protein-coding transcripts were upregulated in non-responders.

The pathway analysis was performed with Gene Ontology (70, 71) using the input of protein-coding transcripts with a log2FoldChange of ≥ 0.5 difference providing 1,052 transcripts downregulated and 1,280 transcripts upregulated in non-responders vs responders. The results showed 37 pathways downregulated in the non-responders group vs responders and no pathways upregulated in non-responders (Figure 8).

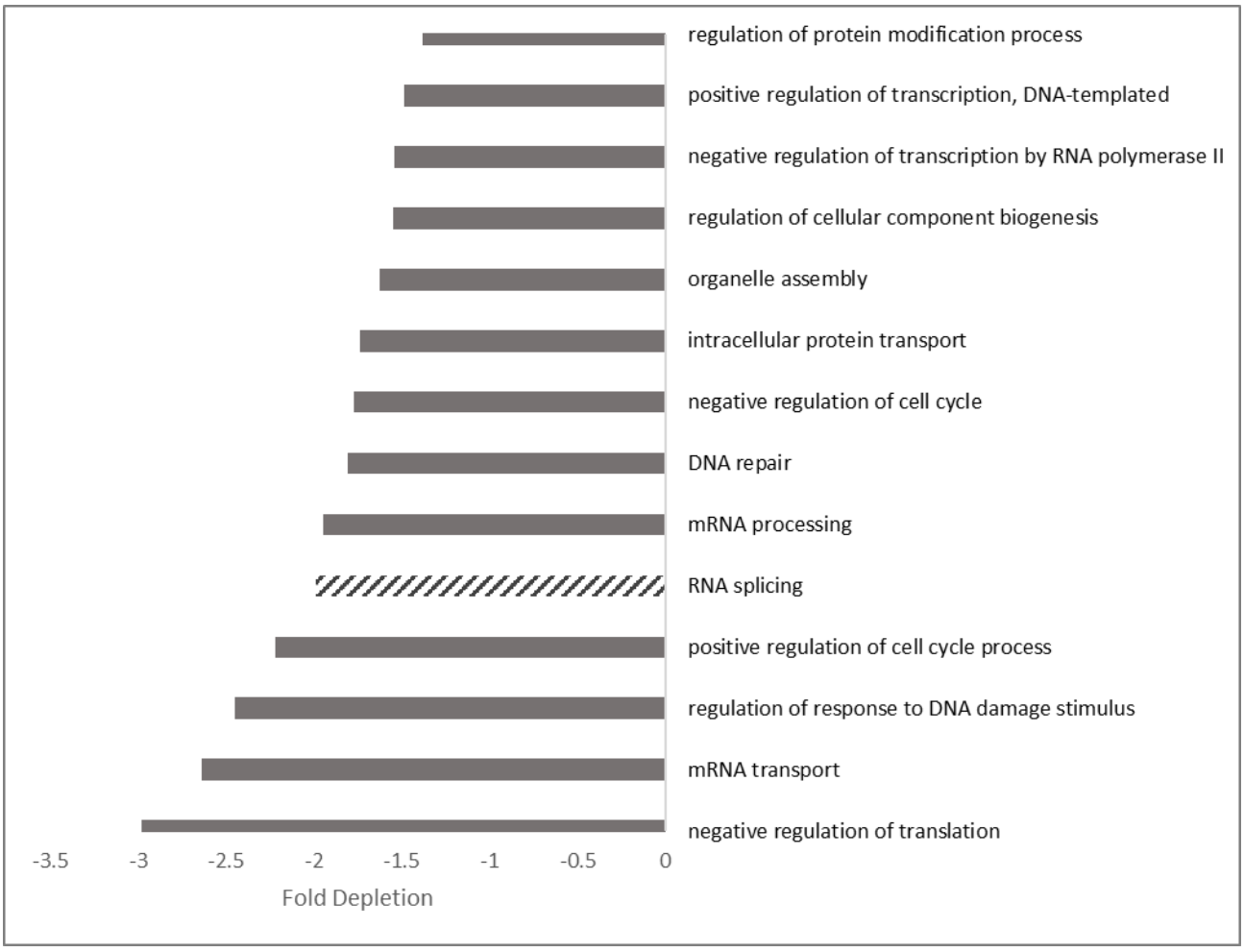

Figure 8: Pathways downregulated in the blood of non-responders *vs* responders, according to the Gene Ontology® (70, 71, 90) analysis.

As indicated by Figure 8, all of the pathways downregulated in the non-responders group are related to cell replication and protein synthesis. This deregulation indicates a disruption in multiple cellular and metabolic processes in patients with a worse prognosis. This finding may be of prognostic relevance, as we see that PBC patients with higher ALP levels and worse treatment responsiveness have downregulation of several essential cellular pathways. Of relevance to the current analysis, "RNA splicing" is one of the downregulated pathways in PBC patients with the highest ALP levels in non-responders to treatment (Table 6). The implication is that splicing may be related to the severity of the disease, as patients with worse disease conditions have significantly lower splicing. Therefore, we directly evaluated whether splicing was altered in the responder vs non-responder groups.

Pathway	GO Term	Functional transcripts	FDR
RNA splicing[1]	GO:0008380	FAM98B, LUC7L3, UPF3B, PAPOLA, AC011511, PPIL1, U2SURP, PCF11, BCLAF1, SREK1IP1, MBNL1, RBM41, QKI, ZNF326, SMNDC1, CPSF2, SCAF11, C2orf49, PPP4R2, RRAGC, hnRNPH3, PTBP3, PNN, PSIP1, RBM25, DDX46, RNPC3, RBMXL1, NSRP1, PRPF40A, PRPF4B, TRA2A, SREK1, SRSF11, MBNL2, THOC2, NUDT21, PRPF38B, AQR, SYNCRIP, hnRNPA3, RSRC1	9.51e-03

[1]The process of removing sections of the primary RNA transcript to remove introns and joining of the exons to form the mature structure of the RNA (70, 71).

Table 6: RNA splicing pathway downregulated in the blood of non-responders (*vs* responders) according to Gene Ontology® (70, 71, 90) analysis.

To determine whether the change in splicing expression was related to prognosis, we elected to analyze the differential exon use in a pairwise comparison between the three groups: healthy subjects, responders, and non-responders (Figure 9). We selected ten patients from each group and calculated the differences in exon processing with an FDR ≤ 0.05 and a fold enrichment ≥ 2. The analysis showed increased differences in exon usage between responders *vs* non-responders, as well

as healthy *vs* non-responders, with limited differences seen in the healthy group *vs* responders. These data indicate disrupted splicing in non-responders because their transcripts contain more exons that were not spliced out, presumably due to inefficient splicing machinery.

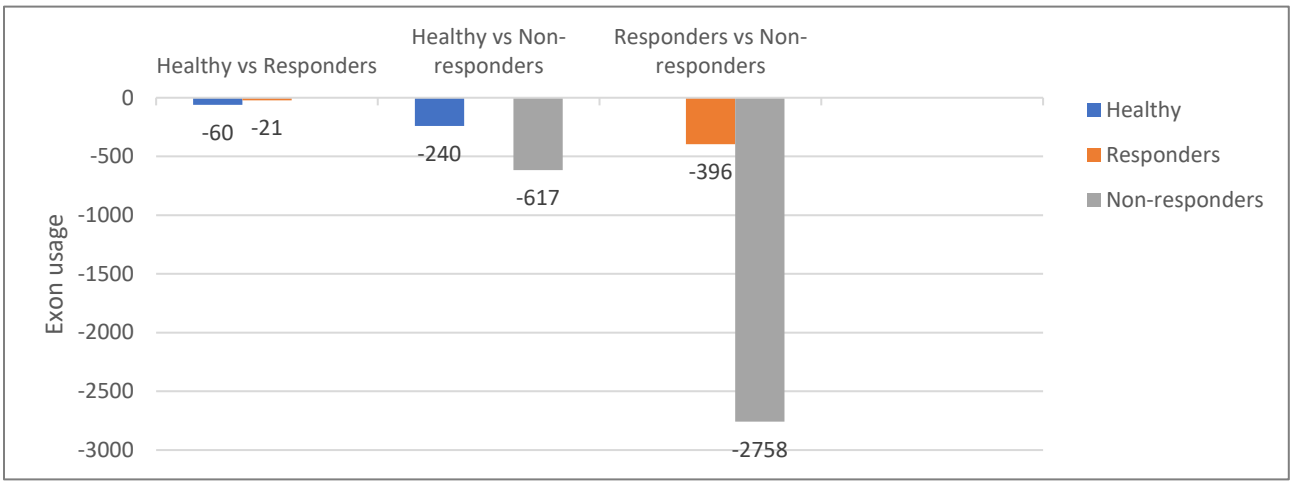

Figure 9: Exon usage among healthy volunteers, non-responders and responders PBC patients.

To evaluate the potential correlation between deregulation in splicing and prognosis, we developed a log-transformed score based on the normalized splicing transcripts downregulated in non-responders. In the first instance, we plotted the individual scores for each differentially expressed transcript on a logistic regression curve with either ALP or total bilirubin levels. We identified individual scores from log-transformed transcripts from 5 genes - PPIL1, PTBP3, RBMXL1, RNPC3 and UPF3B - that demonstrated a significant relationship with both ALP and bilirubin. Then we calculated an aggregate score of all five transcripts and showed that the score negatively correlates with ALP and total bilirubin levels (Figure 10). We also found that the total score was higher in the responders vs non-responders (median 3.62 vs 3.37, $p < 0.025$) with an area under the receiver operator characteristic curve of 0.61 ($p < 0.025$) (Figure 11). Notably the transcript - UPF3B - provided a slightly improved discrimination for prognosis with scores of 0.56 vs. 0.49 in the

responders vs non-responders ($p < 0.007$) and an area under the receiver operator characteristic curve of 0.63 ($p < 0.006$) (Figure 11). These data suggest that patients with a lower splicing score possibly related to diminished splicing had both increased disease activity assessed by higher ALP levels as a measure of disease grade and a more progressive disease as indicated by elevated BLB levels indicative of a higher disease stage (Figure 10 and 11).

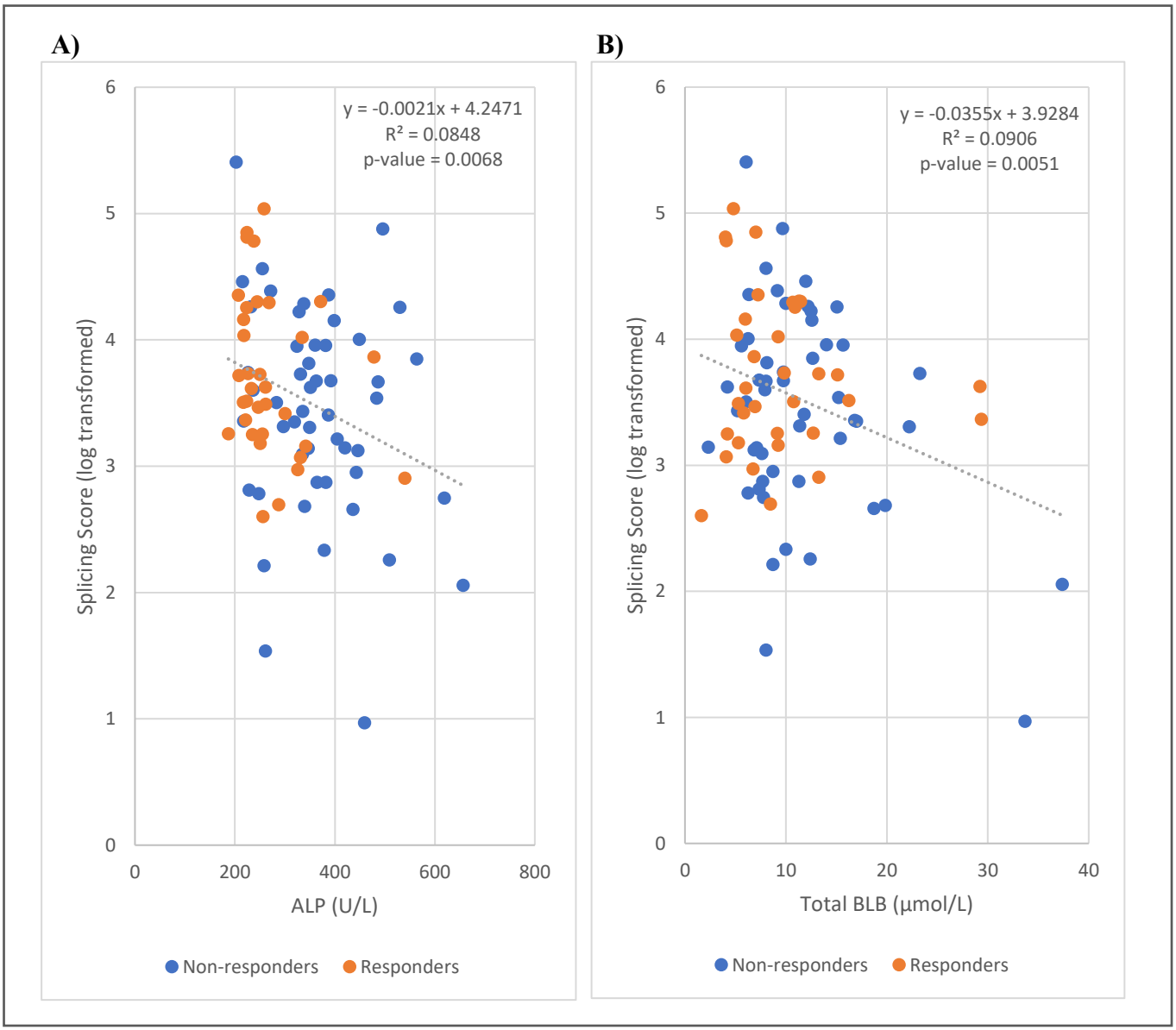

Figure 10: Relationship between splicing and PBC prognosis. (A) compares ALP levels with the splicing score and (B) compared total BLB levels with the splicing score.

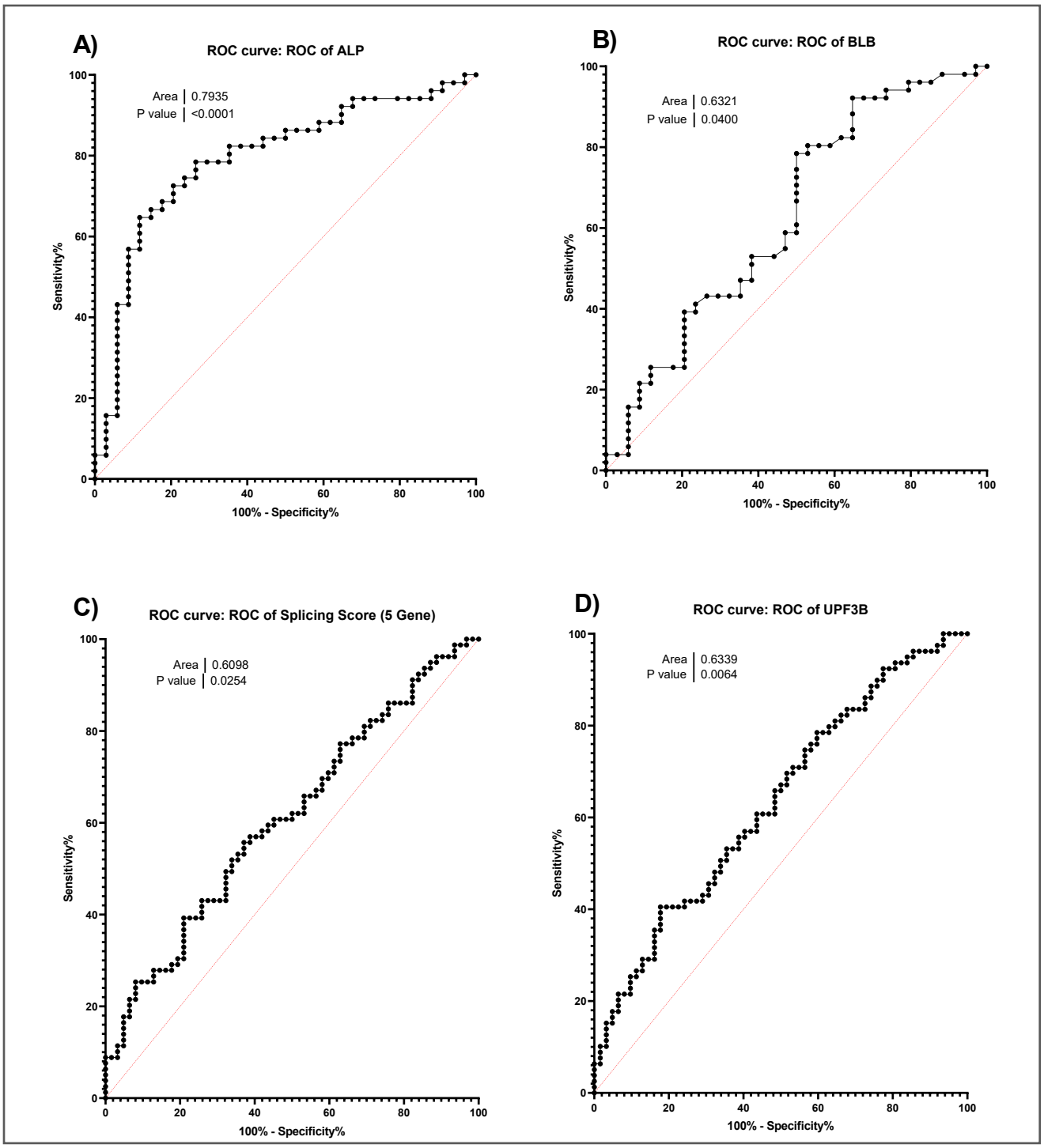

Figure 11: Receiver Operating Characteristic (ROC) curves of the Responders and Non-responders data. (A) shows the ROC curve of the ALP levels of Non-responders against Responders; (B) shows the ROC curve of the total BLB levels of Non-responders against Responders; (C) shows the ROC curve of the splicing score of Non-responders against Responders; (D) shows the ROC curve of the UPF3B gene TPM data fo Non-responders against Responder.

3.3 OTHER PROTEOMICS AND TRANSCRIPTOMICS FINDINGS

On and above the dysregulation in splicing pathways, two other terms were commonly found within the different analyses that included “RNA processing” and “translation”. These pathways may be indicative of global dysregulation of cellular processes in PBC patients. For example, we observed pathways involved in protein-translation are downregulated in the BEC and the blood of PBC patients. The four pathways downregulated in BEC of PBC patients *vs* non-cholestasis controls included “Regulation of translation”, “Positive regulation of translation”, “IRES-dependent viral translational initiation”, and “Negative regulation of translation”. Two pathways downregulated in the blood of PBC patients *vs* healthy subjects were identified as “Translation initiation” and “Cytoplasmatic translation”, and one pathway downregulated in non-responders *vs* responders was “Negative regulation of translation” (Table 7).

Pathways	GO term	Analysis	Proteins/Genes	FDR
Regulation of translation[1]	GO:0006417	Proteomics - Downregulated proteins in PBC BEC (*vs* control)	DHX9, EIF3H, ELAVL1, hnRNPD, hnRNPL, KHSRP, NCL, SHMT2	4.65e-05
Positive regulation of translation[2]	GO:0045727	Proteomics - Downregulated proteins in PBC BEC (*vs* control)	DHX9, ELAVL1, hnRNPD, hnRNPL	0.0022
IRES-dependent viral translational initiation[3]	GO:0075522	Proteomics - Downregulated proteins in PBC BEC (*vs* control)	PCBP2, PTBP1	0.0044
Negative regulation of translation[4]	GO:0017148	Proteomics - Downregulated proteins in PBC BEC (*vs* control)	hnRNPD, KHSRP, NCL	0.0256
Translational initiation[5]	GO:0006413	Transcriptomics – Downregulated transcripts in PBC blood (*vs* healthy)	RPL22, RPL17, RPL23, RPL31, FAU, RPL35A, RPL30, RPS27, RPS17, RPS24, RPS23, RPL5, RPL6	3.99e-06
Cytoplasmic translation[6]	GO:0002181	Transcriptomics - Downregulated transcripts in PBC blood (*vs* healthy)	RPL22, RPL17, RPL31, RPL35A, RPL30, RPS23, RPL6, RPL26L1, TMA7	5.39e-05
Negative regulation of translation[4]	GO:0017148	Transcriptomics – Downregulated transcripts in the blood of non-responders (*vs* responders)	ROCK2, TNRC6A, EIF2AK2, RC3H1, DDX3X, DDX6, DHX36, ROCK1, AGO4, AGO3, FXR1, CNOT6L, CPEB4, TIA1, XRN1, LSM14A, CNOT7, IREB2, SYNCRIP, YTHDF3, ZNF540	3.70e-03
Translational initiation[5]	GO:0006413	Transcriptomics – Downregulated transcripts in PBC BEC (*vs* control)	PABPC1, RPS9, RPS23, RPL23, RPS3, RPS13, RPS7, RPL10, RPS8, RPL13, EIF3C, RPL3	1.24e-03

[1]Any process that modulates the frequency, rate or extent of the chemical reactions and pathways resulting in the formation of proteins by the translation of mRNA or circRNA (70, 71).
[2]Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways resulting in the formation of proteins by the translation of mRNA or circRNA (70, 71).
[3]Process by which viral mRNA translation is initiated, where a domain in the 5' untranslated region (UTR) of the viral mRNA called an internal ribosome entry site (IRES) binds the host 43S preinitiation complex, circumventing regular cap-dependent translation initiation (70, 71, 92).
[4]Any process that stops, prevents or reduces the frequency, rate or extent of the chemical reactions and pathways resulting in the formation of proteins by translating mRNA or circRNA (70, 71).
[5]The process preceding the formation of the peptide bond between the first two amino acids of a protein. This includes forming a complex of the ribosome, mRNA or circRNA, and an initiation complex that contains the first aminoacyl-tRNA (70, 71, 93).
[3]The chemical reactions and pathways resulting in the formation of a protein in the cytoplasm. This is a ribosome-mediated process in which the information in messenger RNA (mRNA) is used to specify the sequence of amino acids in the protein (70, 71).

Table 7: Translation related pathways downregulated in the blood and BEC of PBC patients (*vs* controls), and downregulated in the blood of on non-responders (*vs* responders) according to STRING® (65) and Gene Ontology® (70, 71, 90) analyses

While the cause of these changes also remains to be resolved, it is interesting that viral infection has been linked to the global dysregulation of protein-translation as observed with the unfolded protein response, for example. The host cell can use mechanisms to downregulate translation to limit infection. For example, eukaryotic cells can recognize PAMPs (pathogen-associated molecular patterns) and activate innate immune responses, leading to decreasing translation, mainly translation initiation. Because viruses utilize host cellular processes for their own replication, the host cell downregulates translation to limit the production of viral proteins, and consequently, stop or reduce infection (94). This observation is relevant for our hypothesis that a viral trigger may be central to the perturbation of cellular processes, as protein-translation is downregulated in both blood and BEC of PBC patients.

A further disrupted pathway included RNA processing, an important sequence of events required for the maturation of mRNA molecules (95). We found two RNA processing-related pathways, “mRNA processing” and “Regulation of mRNA processing”, that were downregulated in our proteomic searches in PBC patients *vs* non-cholestasis control BEC. We also found “rRNA processing” downregulated in the transcriptomics analysis of the blood of PBC patients *vs* healthy subjects and “mRNA processing” downregulated in the non-responders *vs* responders (Table 8).

Pathway	GO term	Analysis	Proteins/Genes	FDR
mRNA processing[1]	GO:0006397	Proteomics - Downregulated proteins in PBC BEC (*vs* control)	ALYREF, DHX9, ELAVL1, hnRNPA2B1, hnRNPA3, hnRNPC, hnRNPD, hnRNPH1, hnRNPH3, hnRNPK, hnRNPL, hnRNPU, KHSRP, NHP2L1, NONO, PCBP2, PTBP1, TARDBP	2.94e-15
Regulation of mRNA processing[2]	GO:0050684	Proteomics - Downregulated proteins in PBC BEC (*vs* control)	DHX9, hnRNPA2B1, hnRNPK, hnRNPL, hnRNPU, PTBP1	2.15e-05
rRNA processing[3]	GO:0006364	Transcriptomics – Downregulated transcripts in PBC blood (*vs* healthy)	RPL22, RPL17, RPL23, NGDN, RPL31, PIN4, RPL35A, RPL30, RPS27, RPS17, RPS24, METTL5, RPL5, RPL6	1.63e-04
mRNA processing[1]	GO:0006397	Transcriptomics - Downregulated transcripts in the blood of non-responders (*vs* responders)	GRSF1, LUC7L3, UPF3B, PAPOLA, CMTR2, AC011511, PPIL1, RBBP6, U2SURP, ZC3H11A, PCF11, BCLAF1, SREK1IP1, MBNL1, RBM27, RBM41, MNAT1, QKI, ZNF326, SMNDC1, CPSF2, SCAF11, PPP4R2, PNPT1, hnRNPH3, PTBP3, PNN, PSIP1, PAPOLG, RBM25, DDX46, RNPC3, RBMXL1, NSRP1, PRPF40A, PAN3, PRPF4B, TRA2A, CNOT6L, SREK1, SRSF11, MBNL2, THOC2, NUDT21, PRPF38B, AQR, SYNCRIP, hnRNPA3, RSRC1	3.43e-03

[1]Any process involved in converting a primary mRNA transcript into one or more mature mRNA(s) before translation into polypeptide (70, 71).
[2]Any process that modulates the frequency, rate or extent of mRNA processing (70, 71).
[3]Any process involved in the conversion of a primary ribosomal RNA (rRNA) transcript into one or more mature rRNA molecules (70, 71).

Table 8: RNA processing related pathways downregulated both in the blood and BEC of PBC patients (*vs* controls) and downregulated on non-responders (*vs* responders) according to STRING® (65) and to Gene Ontology® (70, 71, 90) analyses.

Accordingly, protein translation, RNA splicing and RNA processing are all downregulated in most datasets, and these changes have been linked with viral infection. Indeed, this was a consistent finding in all our studies and supports the concept that PBC is a systemic disease with pathological changes in the peripheral blood and BEC of PBC patients. To better visualize our findings regarding the RNA processing and RNA splicing pathways, we condensed the collective data into a *WikiPathways* (96) figure (97) of the mRNA processing pathways that includes relevant RNA splicing proteins and genes. For this, we chose all the functional transcripts downregulated in the PBC *vs* healthy and Responders *vs* Non-responders analyses and all the downregulated proteins from the proteomics analysis (Figure 12).

The figure shows that proteins and transcripts from the BEC and blood of PBC patients are under-expressed in all stages of RNA processing, including RNA splicing. Indeed, the process that contains the most downregulated elements is the formation of pre-mRNPs. These are ribonucleoprotein complexes that process the pre-mRNA molecule associated with the hnRNPs ribonuclear proteins in preparation for splicing (98). The demonstrable downregulation of hnRNPs could cause defective splicing, which was observed in this study.

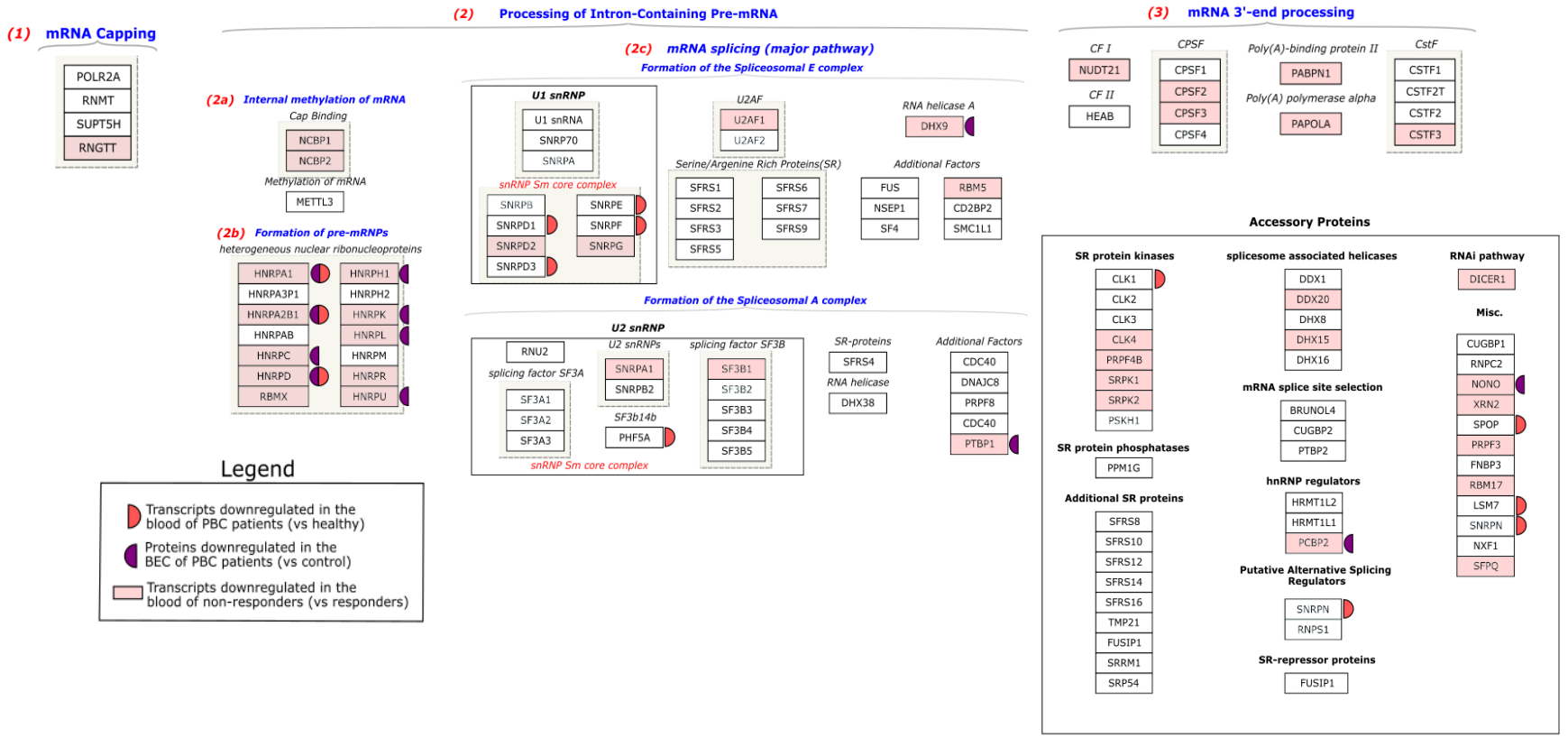

Figure 12: RNA processing pathway (96, 97). The rectangles and semi-circles are colour-coded, representing proteins/genes downregulated in the proteomics analysis of BEC and transcriptomics analysis of peripheral blood of PBC patients.

4. DISCUSSION

After performing several hypothesis-generating proteomic and transcriptomic analyses in both blood and BEC of PBC patients, the key findings of this study are (i) splicing is deregulated in the blood and BEC of PBC patients when compared to healthy volunteers and non-cholestatic liver disease controls, and the process is more advanced in PBC patients with worst prognosis; (ii) differential exon usage analysis showed a significant reduction in splicing in PBC patients with worse prognosis; (iii) previous studies have shown a relation between splicing proteins and viruses.

Splicing involves multiple complex mechanisms. This process can be affected by several external factors, such as xenobiotics and bacterial and viral infections. Many *in vitro* studies have shown that drugs and other compounds can promote splicing modulation (99). Anticancer drugs, such as Spliceostatin A, Meayamycin and Isoginkgetin, are splicing inhibitors (100-103); other common drugs, like salbutamol, valproic acid, and even ethanol, can induce alternative splicing modulation (104-106). Microorganisms are also able to disrupt the splicing process. For example, there is evidence of alteration in the RNA splicing pattern in macrophages post *Mycobacterium tuberculosis* infection (107), and *Salmonella* and *Listeria* infection can induce changes in alternative splicing (108). Ultimately, the specific cause for the splicing deregulation in PBC patients was not analyzed, and therefore, we can only speculate that the human betaretrovirus, shown to be integral to some aspects of PBC pathogenesis, may be playing a role in this process.

It is well known that splicing activity may be altered with a viral infection. A recent study on Sars-Cov-2 showed disruption in splicing and translation caused by the viral proteins, aiming to suppress interferon-gamma responses against the infection (109). Influenza viruses are known to hijack the host splicing machinery to process their RNA segments, leading to alterations in the host spliceosome expression (110). Likewise, oncoviruses, such as Epstein-Barr virus, hepatitis C

and B viruses and Merkel cell polyomavirus, require the RNA machinery from the host cells to stabilize the viral RNA and promote their oncogenic effects (111). These few examples show that splicing is a common cellular process affected by viruses.

The proteins in our proteomics analysis have been linked to both inhibition and stimulation of viral infection (Table 9), viruses can affect splicing, and splicing can impact viral replication (Table 10). "Table 10" references studies that showed how viruses might require splicing proteins for their replication and survival and that some splicing proteins inhibit the viral infection. For the latest, it would be logical that some viruses would target these proteins. This is concordant with our viral hypothesis, meaning that HBRV might be diminishing these proteins for its benefit.

Protein	Virus	Effect on Viral Infection
hnRNPU	HIV	Inhibiting
NONO	HIV	Inhibiting
hnRNPL	Foot-mouth-disease-virus	Inhibiting
ELAVL1	HCV	Stimulating
hnRNDP	HCV	Stimulating
hnRNPK	Dengue Virus (type 2)	Stimulating
hnRNPC	Picornavirus	Stimulating
PTBP1	Picornavirus	Stimulating
PCBP2	Poliovirus	Stimulating
KHSRP	Enterovirus 71	Inhibiting
hnRNPA2B1	HIV-1	Stimulating
hnRNPH1	HIV-1	Stimulating
DHX9	HBV	Stimulating

Table 9: Splicing proteins and their effect on viruses.

Protein	Protein Primary Function	Virus	Effect of the virus on the protein	Impact of the protein on viral infection
hnRNPU	Pre-mRNA alternative splicing modulator (112)	HIV	-	A protein terminal fragment induces an antiviral state in the host cell against HIV, inhibiting the viral gene expression (84)
NONO	Pre-mRNA splicing (113)	HIV	-	Acts as a sensor for the HIV capsid for immune activation of dendritic cell and macrophages against viral infection, helping to decrease the infection rate (85)
hnRNPL	Global regulator of mRNA splicing (114)	Foot-mouth-disease-virus	-	The protein specifically interacts with the viral IRES and negatively regulates the viral replication by inhibiting its RNA replication (86)
ELAVL1	RNA-binding protein. Helps to increase mRNA stability (115)	HCV	-	Regulates HCV replication by interacting with the 3'UTR of the viral RNA, facilitating La-mediated circularization of the HCV genome (74)
hnRNDP	Mediates rapid mRNA decay (116, 117)	HCV	-	Promotes viral protein translation by enhancing HCV IRES-dependent translation (73)
hnRNPK	Multifunctional protein (117). Participates in the regulation of mRNA transcription (118) and translation (119), silencing (120), splicing (121) and stability (122)	Dengue Virus (type 2)	The virus promotes protein cytoplasmatic re-localization (75)	Enhances viral replication (75)
hnRNPC	Splicing regulator with function in the packing of nascent transcripts and in	Picornavirus	-	Helps positive-strand viral RNA synthesis by facilitating

	increasing mRNA translation by binding mRNA's IRES (123)			interaction between the 5' and 3' ends of viral RNAs (77)
PTBP1	Nuclear ribonucleoprotein with splicing regulation functions (117)	Picornavirus	It is mobilized out of the nucleus and then cleaved as the viral infection progresses (76)	Required for picornavirus gene expression and propagation (76)
PCBP2	Binding protein. Binds single-stranded nucleic acids and negatively regulates MA*VS* signalling cellular antiviral responses (78, 124)	Poliovirus	The protein is cleaved by the viral 3CD proteinase, also contributing to viral translation (77-79)	It is necessary for translation initiation, as it binds to the viral IRES and promotes the IRES-mediate translation (77-79)
KHSRP	RNA-binding protein and ITAF. Involved in mRNA localization, transcription, and alternative pre-mRNA (125-127)	Enterovirus 71	-	This protein negatively regulates viral replication by competing with the IRES present in the viral 5' UTR, consequently downregulating IRES-dependant viral translation (87, 88)
hnRNPA2B1 hnRNPH1	Mediators of the pre-mRNA alternative splicing regulation (128, 129)	HIV-1	The virus downregulated the expression of the proteins in the first two weeks of infection, and then when viral production is at its peak, the expression is increased in macrophages (*in vitro*) (80)	The modulation of splicing factors is essential for viral persistence in macrophage reservoir (80)
DHX9	Multifunctional protein. Participates in RNA translation and regulation and DNA replication (130, 131). Also implicated in retroviral RNA expression and nuclear export (132)	HBV	During HBV infection, the virus upregulates the expression of DHX9 (89)	Contributes to viral DNA replication (89)

Table 10: Proteins downregulated in PBC patients (*vs* control) involved in splicing pathways and previously related to viruses, according to STRING® (55) analysis.

The pathogenesis of PBC involves an autoimmune component. We believe that HBRV might be the trigger for this autoimmunity, as we see in other autoimmune diseases, such as systemic lupus erythematosus (SLE), in which studies show that infectious agents can induce the development of autoantibodies (Table 11).

Infectious Agent	Reference
BK virus	(133)
Epstein-Barr virus	(134)
HIV	(135)

Table 11: Some infectious agents that might be linked to the development of autoantibodies in SLE.

We also noticed that patients with higher ALP levels and worse responsiveness to UDCA and OCA treatment have more deregulation in exon usage than patients with a better prognosis, which can be related to a deficiency in the splicing. We can infer it based on the exon usage analysis because if exons are being spliced out when they are not supposed to, there is something wrong with the splicing machinery. Also, the transcriptomics comparison against patients with lower ALP levels and better responsiveness to treatment showed downregulation in splicing related pathways and other essential cellular pathways.

It is important to note that our study presents limitations. First, the proteomics analysis was conducted on BEC from end-stage liver diseases, which could affect the general health of the cell, and then the protein expression. Second, we did not perform proteomics analysis in the POISE study patients, so we can only infer that the changes in expression in the transcripts will translate into changes in the proteins. Also, splicing and other cell process deregulations were not shown in serial samples, meaning that we cannot directly link the worsening in these cellular processes with worsening in the prognosis. Finally, we could not directly link splicing deregulation with HBRV infection, so the viral trigger as the causative for global cell dysregulation can only be suggested.

5. FUTURE DIRECTIONS

To investigate whether the changes outlined herein are linked with HBRV infection, we propose to validate our findings by conducting *in vitro* co-culture studies, as previously conducted by our lab (33). In these studies, for example, we were able to show that both HBRV and the genetically indistinguishable mouse betaretrovirus, MMTV had the capability of triggering a disease-specific phenotype of upregulating the expression of mitochondrial proteins on the cell surface (33). Accordingly, we could proceed with culturing cells with either HBRV or MMTV and evaluate whether the same trends on splicing and global cell deregulation are observed.

In addition, we could prospectively follow PBC patients to determine whether the findings on the RNA level subsequently predict patient outcomes, as suggested by our clinical evaluation of our cellular data. We would also like to conduct a proteomic analysis on the PBC samples obtained from the POISE study (66) to evaluate whether the trends and changes observed in the RNA level are translated to the proteins. Finally, to increase the strength of the results of the omics analysis in the liver, we propose to include healthy liver samples obtained from liver cancer patients that undergo biopsies and conduct proteomics and transcriptomics analysis in these samples, followed by comparison against the data we have on PBC patients.

6. CONCLUSION

Our results suggest dysregulation in splicing, mRNA processing and translation at both the transcriptomics and proteomics levels and BEC and peripheral blood of PBC patients compared to patients with non-cholestatic liver diseases and healthy volunteers, respectively. This deregulation is even more pronounced in PBC patients with progressive disease. These findings may be significant, as they bring new evidence to the light of pathogenic mechanisms implicated in the disease process. With the results we have presented herein, we can infer that the changes in splicing and other cellular processes are either a result of the disease or involved in the pathogenesis. Also, based on other studies relating many of these cellular processes' dysregulation and viral infection, we believe that HBRV may play a causative role for these changes that will require further *in vitro* validation.

REFERENCES

1. Onofrio FQ, Hirschfield GM, Gulamhusein AF. A Practical Review of Primary Biliary Cholangitis for the Gastroenterologist. Gastroenterol Hepatol (N Y). 2019;15(3):145-54.
2. Lleo A, Jepsen P, Morenghi E, Carbone M, Moroni L, Battezzati PM, et al. Evolving Trends in Female to Male Incidence and Male Mortality of Primary Biliary Cholangitis. Sci Rep. 2016;6:25906.
3. Carey EJ, Ali AH, Lindor KD. Primary biliary cirrhosis. The Lancet. 2015;386(10003):1565-75.
4. Boonstra K, Beuers U, Ponsioen CY. Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: A systematic review. Journal of Hepatology. 2012;56(5):1181-8.
5. Yoshida EM, Mason A, Peltekian KM, Shah H, Thiele S, Borrelli R, et al. Epidemiology and liver transplantation burden of primary biliary cholangitis: a retrospective cohort study. CMAJ Open. 2018;6(4):E664-E70.
6. Hirschfield GM, Gershwin ME. The Immunobiology and Pathophysiology of Primary Biliary Cirrhosis. Annual Review of Pathology: Mechanisms of Disease. 2013;8(1):303-30.
7. Mason AL. Is PBC a viral infectious disease? Best Pract Res Clin Gastroenterol. 2018;34-35:27-39.
8. Nakamura M. Clinical Significance of Autoantibodies in Primary Biliary Cirrhosis. Seminars in Liver Disease. 2014;34:334–40.
9. Webb GJ, Hirschfield GM. Primary Biliary Cholangitis. Handbook of Liver Disease: Elsevier; 2018. p. 217-32.
10. Invernizzi P, Floreani A, Carbone M, Marzioni M, Craxi A, Muratori L, et al. Primary Biliary Cholangitis: advances in management and treatment of the disease. Digestive and Liver Disease. 2017;49(8):841-6.
11. Lindor KD, Bowlus CL, Boyer J, Levy C, Mayo M. Primary Biliary Cholangitis: 2018 Practice Guidance from the American Association for the Study of Liver Diseases. Hepatology (Baltimore, Md). 2018.
12. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. J Hepatol. 2017;67(1):145-72.

13. Witt-Sullivan H, Heathcote J, Cauch K, Blendis L, Ghent C, Katz A, et al. The demography of primary biliary cirrhosis in ontario, canada. Hepatology (Baltimore, Md). 1990;12(1):98-105.
14. Chalifoux SL, Konyn PG, Choi G, Saab S. Extrahepatic Manifestations of Primary Biliary Cholangitis. Gut and Liver. 2017;11(6):771-80.
15. Khanna A, Hegade VS, Jones DE. Management of Fatigue in Primary Biliary Cholangitis. Current Hepatology Reports. 2019;18(2):127-33.
16. Heathcote EJ, Cauch-Dudek K, Walker V, Bailey RJ, Blendis LM, Ghent CN, et al. The canadian multicenter double-blind randomized controlled trial of ursodeoxycholic acid in primary biliary cirrhosis. Hepatology (Baltimore, Md). 1994;19(5):1149-56.
17. Talwalkar JA, Souto E, Jorgensen RA, Lindor KD. Natural history of pruritus in primary biliary cirrhosis. Clinical Gastroenterology and Hepatology. 2003;1(4):297-302.
18. Uddenfeldt P, Danielsson Å, Forssell Å, Holm M, Östberg Y. Features of Sjögren's syndrome in patients with primary biliary cirrhosis. Journal of Internal Medicine. 1991;230(5):443-8.
19. Reksten TR, Jonsson MV. Sjögren's Syndrome. Oral and Maxillofacial Surgery Clinics of North America. 2014;26(1):1-12.
20. Crowe JP, Christensen E, Butler J, Wheeler P, Doniach D, Keenan J, et al. Primary Biliary Cirrhosis: The Prevalence of Hypothyroidism and its Relationship to Thyroid Autoantibodies and Sicca Syndrome. Gastroenterology. 1980;78(6):1437-41.
21. Floreani A, Franceschet I, Cazzagon N, Spinazzè A, Buja A, Furlan P, et al. Extrahepatic Autoimmune Conditions Associated with Primary Biliary Cirrhosis. Clinical Reviews in Allergy & Immunology. 2015;48(2-3):192-7.
22. Marasini B. Rheumatic disorders and primary biliary cirrhosis: an appraisal of 170 Italian patients. Annals of the Rheumatic Diseases. 2001;60(11):1046-9.
23. Tanaka A, Leung PSC, Gershwin ME. Pathogen infections and primary biliary cholangitis. Clinical & Experimental Immunology. 2019;195(1):25-34.
24. Tanaka A, Leung PS, Gershwin ME. Environmental basis of primary biliary cholangitis. Experimental Biology and Medicine. 2018;243(2):184-9.

25. Abu-Mouch S, Selmi C, Benson GD, Kenny TP, Invernizzi P, Zuin M, et al. Geographic Clusters of Primary Biliary Cirrhosis. Clinical and Developmental Immunology. 2003;10(2-4):127-31.

26. Selmi C, Mayo MJ, Bach N, Ishibashi H, Invernizzi P, Gish RG, et al. Primary biliary cirrhosis in monozygotic and dizygotic twins: Genetics, epigenetics, and environment. Gastroenterology. 2004;127(2):485-92.

27. Jones DE, Watt FE, Metcalf JV, Bassendine MF, James OFW. Familial primary biliary cirrhosis reassessed: a geographically-based population study. Journal of Hepatology. 1999;30(3):402-7.

28. Gershwin ME, Selmi C, Worman HJ, Gold EB, Watnik M, Utts J, et al. Risk factors and comorbidities in primary biliary cirrhosis: A controlled interview-based study of 1032 patients. Hepatology (Baltimore, Md). 2005;42(5):1194-202.

29. Tanaka A, Leung PSC, Gershwin ME. The genetics of primary biliary cholangitis. Current Opinion in Gastroenterology. 2019;35(2):93-8.

30. Selmi C, Gershwin ME. The role of environmental factors in primary biliary cirrhosis. Trends in Immunology. 2009;30(8):415-20.

31. Ala A, Stanca CM, Bu-Ghanim M, Ahmado I, Branch AD, Schiano TD, et al. Increased prevalence of primary biliary cirrhosis near superfund toxic waste sites. Hepatology (Baltimore, Md). 2006;43(3):525-31.

32. Koutsoumpas AL, Kriese S, Rigopoulou EI. Popular and unpopular infectious agents linked to primary biliary cirrhosis. Autoimmunity Highlights. 2012;3(3):95-104.

33. Xu L, Shen Z, Guo L, Fodera B, Keogh A, Joplin R, et al. Does a betaretrovirus infection trigger primary biliary cirrhosis? Proceedings of the National Academy of Sciences. 2003;100(14):8454-9.

34. Mason AL, Xu L, Guo L, Munoz S, Jaspan JB, Bryer-Ash M, et al. Detection of retroviral antibodies in primary biliary cirrhosis and other idiopathic biliary disorders. Lancet. 1998;351(9116):1620-4.

35. Wang W, Indik S, Wasilenko ST, Faschinger A, Carpenter EJ, Tian Z, et al. Frequent proviral integration of the human betaretrovirus in biliary epithelium of patients with autoimmune and idiopathic liver disease. Aliment Pharmacol Ther. 2015;41(4):393-405.

36. Xu L, Sakalian M, Shen Z, Loss G, Neuberger J, Mason A. Cloning the human betaretrovirus proviral genome from patients with primary biliary cirrhosis. Hepatology (Baltimore, Md). 2004;39(1):151-6.

37. Zhang G, Bashiri K, Kneteman M, Cave K, Hong Y, Mackey JR, et al. Seroprevalence of Human Betaretrovirus Surface Protein Antibodies in Patients with Breast Cancer and Liver Disease. Journal of Oncology. 2020;2020:1-9.

38. Lytvyak E, Hosamani I, Montano-Loza AJ, Saxinger L, Mason AL. Randomized clinical trial: Combination antiretroviral therapy with tenofovir-emtricitabine and lopinavir-ritonavir in patients with primary biliary cholangitis. Canadian Liver Journal. 2019;2(1):31-44.

39. Griffiths AJF, Wessler SR, Carroll SB, Doebley JF. Introduction to genetic analysis. Eleventh edition. ed. xxiii, 868 pages p.

40. Ashraf U, Benoit-Pilven C, Lacroix V, Navratil V, Naffakh N. Advances in Analyzing Virus-Induced Alterations of Host Cell Splicing. Trends Microbiol. 2019;27(3):268-81.

41. Newman A. RNA splicing. Current Biology. 1998;8(25):R903-R5.

42. Rio DC. Splicing of pre-mRNA: mechanism, regulation and role in development. Current Opinion in Genetics & Development. 1993;3(4):574-84.

43. Björk P, Wieslander L. Integration of mRNP formation and export. Cellular and Molecular Life Sciences. 2017;74(16):2875-97.

44. Wang Y, Liu J, Huang B, Xu Y-M, Li J, Huang L-F, et al. Mechanism of alternative splicing and its regulation. Biomedical Reports. 2015;3(2):152-8.

45. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010;463(7280):457-63.

46. Harwig A, Landick R, Berkhout B. The Battle of RNA Synthesis: Virus versus Host. Viruses. 2017;9(10):309.

47. Herbert K, Nag A. A Tale of Two RNAs during Viral Infection: How Viruses Antagonize mRNAs and Small Non-Coding RNAs in The Host Cell. Viruses. 2016;8(6):154.

48. Rivas H, Schmaling S, Gaglia M. Shutoff of Host Gene Expression in Influenza A Virus and Herpesviruses: Similar Mechanisms and Common Themes. Viruses. 2016;8(4):102.

49. De Maio FA, Risso G, Iglesias NG, Shah P, Pozzi B, Gebhard LG, et al. The Dengue Virus NS5 Protein Intrudes in the Cellular Spliceosome and Modulates Splicing. PLOS Pathogens. 2016;12(8):e1005841.

50. Hu B, Li X, Huo Y, Yu Y, Zhang Q, Chen G, et al. Cellular responses to HSV-1 infection are linked to specific types of alterations in the host transcriptome. Scientific Reports. 2016;6(1):28075.

51. Paronetto MP, Passacantilli I, Sette C. Alternative splicing and cell survival: from tissue homeostasis to disease. Cell Death & Differentiation. 2016;23(12):1919-29.

52. Zhu C, Xiao F, Hong J, Wang K, Liu X, Cai D, et al. EFTUD2 Is a Novel Innate Immune Regulator Restricting Hepatitis C Virus Infection through the RIG-I/MDA5 Pathway. Journal of Virology. 2015;89(13):6608-18.

53. Shkreta L, Chabot B. The RNA Splicing Response to DNA Damage. Biomolecules. 2015;5(4):2935-77.

54. Bagga S, Bouchard MJ. Cell Cycle Regulation During Viral Infection. Methods in Molecular Biology: Springer New York; 2014. p. 165-227.

55. Fan Y, Sanyal S, Bruzzone R. Breaking Bad: How Viruses Subvert the Cell Cycle. Frontiers in Cellular and Infection Microbiology. 2018;8.

56. Song W, Tang D, Chen D, Zheng F, Huang S, Xu Y, et al. Advances in applying of multi-omics approaches in the research of systemic lupus erythematosus. International Reviews of Immunology. 2020;39(4):163-73.

57. Wheelock CE, Goss VM, Balgoma D, Nicholas B, Brandsma J, Skipp PJ, et al. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. European Respiratory Journal. 2013;42(3):802-25.

58. Sun YV, Hu Y-J. Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. Advances in Genetics: Elsevier; 2016. p. 147-90.

59. Yadav D, Tanveer A, Malviya N, Yadav S. Overview and Principles of Bioengineering. Omics Technologies and Bio-Engineering: Elsevier; 2018. p. 3-23.

60. Eraslan G, Avsec Ž, Gagneur J, Theis FJ. Deep learning: new computational modelling techniques for genomics. Nature Reviews Genetics. 2019;20(7):389-403.

61. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biology. 2017;18(1).

62. Yu L-R, Stewart NA, Veenstra TD. Proteomics. Essentials of Genomic and Personalized Medicine: Elsevier; 2010. p. 89-96.

63. Clish CB. Metabolomics: an emerging but powerful tool for precision medicine. Molecular Case Studies. 2015;1(1):a000588.
64. Joplin R, Kachilele S. Human Intrahepatic Biliary Epithelial Cell Lineages: Studies In Vitro. Methods in Molecular Biology: Humana Press; 2009. p. 193-206.
65. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research. 2019;47(D1):D607-D13.
66. Nevens F, Andreone P, Mazzella G, Strasser SI, Bowlus C, Invernizzi P, et al. A Placebo-Controlled Trial of Obeticholic Acid in Primary Biliary Cholangitis. New England Journal of Medicine. 2016;375(7):631-43.
67. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biology. 2013;14(4):R36.
68. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Research. 2012;22(10):2008-17.
69. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.
70. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nature Genetics. 2000;25(1):25-9.
71. Carbon S, Douglass E, Good BM, Unni DR, Harris NL, Mungall CJ, et al. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Research. 2021;49(D1):D325-D34.
72. Wysokinski F. Characterization of the Mitochondrial Phenotype Associated with Primary Biliary Cholangitis. Education and Research Archives: University of Alberta; 2017.
73. Paek KY, Kim CS, Park SM, Kim JH, Jang SK. RNA-Binding Protein hnRNP D Modulates Internal Ribosome Entry Site-Dependent Translation of Hepatitis C Virus RNA. Journal of Virology. 2008;82(24):12082-93.
74. Shwetha S, Kumar A, Mullick R, Vasudevan D, Mukherjee N, Das S. HuR Displaces Polypyrimidine Tract Binding Protein To Facilitate La Binding to the 3′ Untranslated Region and Enhances Hepatitis C Virus Replication. Journal of Virology. 2015;89(22):11356-71.

75. Brunetti JE, Scolaro LA, Castilla V. The heterogeneous nuclear ribonucleoprotein K (hnRNP K) is a host factor required for dengue virus and Junín virus multiplication. Virus Research. 2015;203:84-91.
76. Florez PM, Sessions OM, Wagner EJ, Gromeier M, Garcia-Blanco MA. The Polypyrimidine Tract Binding Protein Is Required for Efficient Picornavirus Gene Expression and Propagation. Journal of Virology. 2005;79(10):6172-9.
77. Holmes AC, Semler BL. Picornaviruses and RNA Metabolism: Local and Global Effects of Infection. Journal of Virology. 2019;93(21).
78. Walter BL, Parsley TB, Ehrenfeld E, Semler BL. Distinct Poly(rC) Binding Protein KH Domain Determinants for Poliovirus Translation Initiation and Viral RNA Replication. Journal of Virology. 2002;76(23):12008-22.
79. Chase AJ, Daijogo S, Semler BL. Inhibition of Poliovirus-Induced Cleavage of Cellular Protein PCBP2 Reduces the Levels of Viral RNA Replication. Journal of Virology. 2014;88(6):3192-201.
80. Dowling D, Nasr-Esfahani S, Tan CH, O'Brien K, Howard JL, Jans DA, et al. HIV-1 infection induces changes in expression of cellular splicing factors that regulate alternative viral splicing and virus production in macrophages. Retrovirology. 2008;5(1):18.
81. Kung YA, Hung CT, Chien KY, Shih SR. Control of the negative IRES trans-acting factor KHSRP by ubiquitination. Nucleic Acids Res. 2017;45(1):271-87.
82. Smith AD. Oxford dictionary of biochemistry and molecular biology. Oxford: Oxford University Press; 2004.
83. Atkins JF, Cech TR, Gesteland RF. The RNA world : the nature of modern RNA suggests a prebiotic RNA. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; 1999.
84. Valente ST, Goff SP. Inhibition of HIV-1 Gene Expression by a Fragment of hnRNP U. Molecular Cell. 2006;23(4):597-605.
85. Lahaye X, Gentili M, Silvin A, Conrad C, Picard L, Jouve M, et al. NONO Detects the Nuclear HIV Capsid to Promote cGAS-Mediated Innate Immune Activation. Cell. 2018;175(2):488-501.e22.
86. Sun C, Liu M, Chang J, Yang D, Zhao B, Wang H, et al. Heterogeneous Nuclear Ribonucleoprotein L Negatively Regulates Foot-and-Mouth Disease Virus Replication through

Inhibition of Viral RNA Synthesis by Interacting with the Internal Ribosome Entry Site in the 5′ Untranslated Region. Journal of Virology. 2020;94(10).

87. Lin J-Y, Li M-L, Shih S-R. Far upstream element binding protein 2 interacts with enterovirus 71 internal ribosomal entry site and negatively regulates viral translation. Nucleic Acids Research. 2009;37(1):47-59.

88. Briata P, Bordo D, Puppo M, Gorlero F, Rossi M, Perrone-Bizzozero N, et al. Diverse roles of the nucleic acid-binding protein KHSRP in cell differentiation and disease. Wiley Interdiscip Rev RNA. 2016;7(2):227-40.

89. Shen B, Chen Y, Hu J, Qiao M, Ren J, Hu J, et al. Hepatitis B virus X protein modulates upregulation of DHX9 to promote viral DNA replication. Cellular Microbiology. 2020;22(3).

90. Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Research. 2010;38(suppl_1):D204-D10.

91. Griffiths L JD. Pathogenesis of primary biliary cirrhosis and its fatigue. Digestive Diseases. 2014;32(5):615-25.

92. Balvay L, Rifo RS, Ricci EP, Decimo D, Ohlmann T. Structural and functional diversity of viral IRESes. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2009;1789(9-10):542-57.

93. Lewin B. Genes. Oxford: Oxford University Press; 2000.

94. Jaafar ZA, Kieft JS. Viral RNA structure-based strategies to manipulate translation. Nature Reviews Microbiology. 2019;17(2):110-23.

95. Potter K, Cremona N, Wise JA. Messenger RNA Processing in Eukaryotes. Encyclopedia of Biological Chemistry: Elsevier; 2013. p. 59-64.

96. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter N, Denise, Hanspers K, et al. WikiPathways: connecting communities. Nucleic Acids Research. 2021;49(D1):D613-D21.

97. Nathan Salomonis KH, Alexander Pico, Martijn van Iersel, Andrew Mason, Daniela Digles, Friederike Ehrhart, Martina Kutmon, Zahra Roudbari. mRNA Processing (*Homo sapiens*) [Available from: https://www.wikipathways.org/index.php/Pathway:WP411_r105857.

98. Daneholt B. Pre-mRNP particles: From gene to nuclear pore. Current Biology. 1999;9(11):R412-R5.

99. Zaharieva E, Chipman JK, Soller M. Alternative splicing interference by xenobiotics. Toxicology. 2012;296(1-3):1-12.
100. Kaida D, Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., Tani, T., Horinouchi, S., Yoshida, M. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nature Chemical Biology. 2007;3:576-83.
101. Roybal GA, Jurica MS. Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. Nucleic Acids Research. 2010;38(19):6664-72.
102. Albert BJ, Mcpherson PA, O'Brien K, Czaicki NL, Destefino V, Osman S, et al. Meayamycin inhibits pre–messenger RNA splicing and exhibits picomolar activity against multidrug-resistant cells. Molecular Cancer Therapeutics. 2009;8(8):2308-18.
103. O'Brien K, Matlin AJ, Lowell AM, Moore MJ. The Biflavonoid Isoginkgetin Is a General Inhibitor of Pre-mRNA Splicing*. Journal of Biological Chemistry. 2008;283(48):33147-54.
104. Angelozzi C BF, Tiziano FD, et al. Salbutamol increases SMN mRNA and protein levels in spinal muscular atrophy cells. Journal of Medical Genetics. 2008;45:29-31.
105. Brichta L. Valproic acid increases the SMN2 protein level: a well-known drug as a potential therapy for spinal muscular atrophy. Human molecular genetics. 2003;12(19):2481-9.
106. Oomizu S, Boyadjieva N, Sarkar DK. Ethanol and Estradiol Modulate Alternative Splicing of Dopamine D2 Receptor Messenger RNA and Abolish the Inhibitory Action of Bromocriptine on Prolactin Release From the Pituitary Gland. Alcoholism: Clinical & Experimental Research. 2003;27(6):975-80.
107. Agrawal AK, Ranjan R, Chandra S, Rout TK, Misra A, Reddy TJ. Some proteins of M. tuberculosis that localise to the nucleus of THP-1-derived macrophages. Tuberculosis. 2016;101:75-8.
108. Pai AA, Baharian G, Pagé Sabourin A, Brinkworth JF, Nédélec Y, Foley JW, et al. Widespread Shortening of 3' Untranslated Regions and Increased Exon Inclusion Are Evolutionarily Conserved Features of Innate Immune Responses to Infection. PLOS Genetics. 2016;12(9):e1006338.
109. Banerjee AK, Blanco MR, Bruce EA, Honson DD, Chen LM, Chow A, et al. SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. Cell. 2020;183(5):1325-39.e21.

110. Dubois J, Terrier O, Rosa-Calatrava M. Influenza Viruses and mRNA Splicing: Doing More with Less. mBio. 2014;5(3):e00070-14-e.

111. Ajiro M, Zheng Z-M. Oncogenes and RNA splicing of human tumor viruses. Emerging Microbes & Infections. 2014;3(1):1-16.

112. Ye J, Beetz N, O'Keeffe S, Tapia JC, Macpherson L, Chen WV, et al. hnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. Proceedings of the National Academy of Sciences. 2015;112(23):E3020-E9.

113. Feng P, Li L, Deng T, Liu Y, Ling N, Qiu S, et al. NONO and tumorigenesis: More than splicing. Journal of Cellular and Molecular Medicine. 2020.

114. Hung L-H, Heiner, M., Hui, J., Schreiner, S., Benes, V., Bindereif, A. Diverse roles of hnRNP L in mammalian mRNA processing: A combined microarray and RNAi analysis. RNA. 2008;14:284–96.

115. Zhang W, Vreeland AC, Noy N. The RNA-binding protein HuR regulates protein nuclear import. Journal of Cell Science. 2016:jcs.192096.

116. Fialcowitz EJ, Brewer BY, Keenan BP, Wilson GM. A Hairpin-like Structure within an AU-rich mRNA-destabilizing Element Regulates trans-Factor Binding Selectivity and mRNA Decay Kinetics. Journal of Biological Chemistry. 2005;280(23):22406-17.

117. Geuens T, Bouhy D, Timmerman V. The hnRNP family: insights into their role in health and disease. Human Genetics. 2016;135(8):851-67.

118. Stains P, Joseph, Lecanda F, Towler A, Dwight, Civitelli R. Heterogeneous nuclear ribonucleoprotein K represses transcription from a cytosine/thymidine-rich element in the osteocalcin promoter. Biochemical Journal. 2005;385(2):613-23.

119. Habelhah H, Shah K, Huang L, Ostareck-Lederer A, Burlingame AL, Shokat KM, et al. ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. Nature Cell Biology. 2001;3(3):325-30.

120. Fan X, Xiong H, Wei J, Gao X, Feng Y, Liu X, et al. Cytoplasmic hnRNPK interacts with GSK3β and is essential for the osteoclast differentiation. Scientific Reports. 2016;5(1):17732.

121. Cao W, Razanau, A., Feng, D., Lobo, V.G., Xie, J. Control of alternative splicing by forskolin through hnRNP K during neuronal differentiation. Nucleic Acids Research. 2012;40:8059–71.

122. Fukuda T, Naiki T, Saito M, Irie K. hnRNP K interacts with RNA binding motif protein 42 and functions in the maintenance of cellular ATP level during stress conditions. Genes to Cells. 2009;14(2):113-28.
123. Han P, Siew, Tang H, Yue, Smith R. Functional diversity of the hnRNPs: past, present and perspectives. Biochemical Journal. 2010;430(3):379-92.
124. You F, Sun H, Zhou X, Sun W, Liang S, Zhai Z, et al. PCBP2 mediates degradation of the adaptor MAVS via the HECT ubiquitin ligase AIP4. Nature Immunology. 2009;10(12):1300-8.
125. Min H, Turck CW, Nikolic JM, Black DL. A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. Genes & Development. 1997;11(8):1023-36.
126. Gherzi R, Lee K-Y, Briata P, Wegmüller D, Moroni C, Karin M, et al. A KH Domain RNA Binding Protein, KSRP, Promotes ARE-Directed mRNA Turnover by Recruiting the Degradation Machinery. Molecular Cell. 2004;14(5):571-83.
127. Chen L-L, Kung Y-A, Weng K-F, Lin J-Y, Horng J-T, Shih S-R. Enterovirus 71 Infection Cleaves a Negative Regulator for Viral Internal Ribosomal Entry Site-Driven Translation. Journal of Virology. 2013;87(7):3828-38.
128. Paul S, Dansithong W, Kim D, Rossi J, Webster NJG, Comai L, et al. Interaction of musleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. The EMBO Journal. 2006;25(18):4271-83.
129. Markovtsov V, Nikolic JM, Goldman JA, Turck CW, Chou M-Y, Black DL. Cooperative Assembly of an hnRNP Complex Induced by a Tissue-Specific Homolog of Polypyrimidine Tract Binding Protein. Molecular and Cellular Biology. 2000;20(20):7463-79.
130. Tetsuka T, Uranishi H, Sanda T, Asamitsu K, Yang J-P, Wong-Staal F, et al. RNA helicase A interacts with nuclear factor κB p65 and functions as a transcriptional coactivator. European Journal of Biochemistry. 2004;271(18):3741-51.
131. Hartman TR, Qian S, Bolinger C, Fernandez S, Schoenberg DR, Boris-Lawrie K. RNA helicase A is necessary for translation of selected messenger RNAs. Nature Structural & Molecular Biology. 2006;13(6):509-16.
132. Singh G, Heng X, Boris-Lawrie K. Cellular RNA Helicases Support Early and Late Events in Retroviral Replication. Retrovirus-Cell Interactions: Elsevier; 2018. p. 253-71.

133. Flaegstad T, Fredriksen K, Dahl B, Traavik T, Rekvig OP. Inoculation with BK virus may break immunological tolerance to histone and DNA antigens. Proceedings of the National Academy of Sciences. 1988;85(21):8171-5.
134. Sundar K, Jacques S, Gottlieb P, Villars R, Benito M-E, Taylor DK, et al. Expression of the Epstein-Barr virus nuclear antigen-1 (EBNA-1) in the mouse can elicit the production of anti-dsDNA and anti-Sm antibodies. Journal of Autoimmunity. 2004;23(2):127-40.
135. González CM L-LF, Monteagudo I, et al. Antiribonucleoprotein antibodies in children with human immunodeficiency virus infection: comparative study with childhood-onset systemic lupus erythematosus. Pediatric AIDS and HIV Infection. 1996;7(6)::401-8.