# Identifying RNA-binding landscapes of PXR using enhanced UV-crosslinking and immunoprecipitation followed by sequencing

by

**Xiaofei Wang**

BS, Wuhan University, 2019

Submitted to the Graduate Faculty of

School of Pharmacy

of the requirements for the degree of

Master of Science

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

SCHOOL OF PHARMACY

This thesis/dissertation was presented

by

**Xiaofei Wang**

It was defended on

April 2, 2021

and approved by

Da Yang, Associate Professor, Department of Pharmaceutical Science

Xie Wen, Professor, Department of Pharmaceutical Science

Xiaochao Ma, Associate Professor, Department of Pharmaceutical Science

Min Zhang, Research Assistant Professor, Pharmaceutical Sciences

Thesis Advisor/Dissertation Director: Da Yang, Associate Professor, Department of Pharmaceutical Science

**Identifying RNA-binding landscapes of PXR using enhanced UV-crosslinking and immunoprecipitation followed by sequencing**

Xiaofei Wang, BS

University of Pittsburgh, 2021

Pregnane X Receptor (PXR) is a nuclear receptor whose primary function is to sense the presence of foreign toxic substances and subsequently up-regulate the expression of proteins involved in the clearance of these substances transcriptionally. Intriguingly, emerging evidence suggests that PXR may also bind with RNAs and directly regulate them. For example, PXR was reported to decrease the stability of TLR4 mRNA after treated with its ligand. However, the underlying mechanism of the PXR's regulation of RNAs remains elusive. Here, we aim to provide insights into the post-transcriptionally regulatory function of PXR. Specifically, we used enhanced UV-crosslinking and immunoprecipitation (eCLIP) to identify the RNAs that can bind with PXR in both colorectal cancer cell lines and mouse livers with humanized PXR. This analysis revealed that PXR can bind with thousands of mRNAs in both human colorectal cancer cell lines and mouse livers. Moreover, analysis of PXR's binding location demonstrated that PXR eCLIP peaks are highly enriched in the 3'UTR region of mRNAs. Treatment of rifampicin, a ligand of PXR, can further increase PXR binding preference in mRNA 3'UTR. Genes with PXR binding to their 3'UTR were enriched in several pathways like cell migration and apoptosis, metabolism of glucose and triglyceride, and vascular functions, which are all known to be related to PXR functions. Since 3'UTR is responsible for RNA stability, it is possible that PXR can post-transcriptionally regulate these signaling pathways. In this regard, we detected the correlation between the expression of PXR and its eCLIP targets using the TCGA database. PXR expression level are positively

corrected with some of its eCLIP-seq targets like VEGFA and Igf1, suggesting PXR may regulate VEGFA and Igf1 expression post-transcriptionally. In summary, our study suggests PXR can bind to the 3'UTR of mRNAs and regulate their expression level post-transcriptionally.

## Table of Contents

## List of Figures

# 1.0 Introduction

## 1.1 RNA-binding protein

Disordered gene expression is the major feature of various diseases including metabolic disorders and cancers. [1] However, the regulation of gene expression is not confined to the differential transcription level from DNA. The rate of mRNA degradation can determine how long the mRNA can serve as a template for translation, also plays an important role in determining the gene expression levels.[2] Pre-mRNA splicing is a common regulatory mechanism for regulating mRNA stability. [3] Besides, the 3’ untranslated region (3’UTR) is pivotal in mRNA stability. Specifically, the single-stranded regions in 3'UTR are susceptible to cleavage by ribonucleases. The cleavage on the 3’UTR region will separate the polyA tail from the coding DNA sequence (CDS) and 5' untranslated region (5'UTR), which will cause the mRNA to be degraded rapidly. [4] Emerging evidence has shown that these processes, which determine the half-life of a specific mRNA, can be regulated by several types of regulators, including RNA-binding protein (RBP).

RBPs are a diverse class of proteins that bind with RNAs and modulate the stability or function of the bound RNAs. Some RBPs can participate in the maturation and fate of their target RNA substrates and regulate numerous aspects of gene expression, including pre-mRNA splicing, cleavage and polyadenylation, RNA stability, RNA localization, RNA editing, and translation. [5] Some RBPs can bind to the 3’UTR and destabilize mRNAs by recruiting exosomes to the mRNAs or stabilize mRNAs by competing with other proteins that normally have a destabilizing effect on mRNAs. [6] Human antigen R (HuR), a well-known RNA binding protein, can bind to the cis-

acting AU-rich elements at 3'UTR and stabilize a variety of mRNAs including mRNAs of estrogen receptor (ER), Erb-B2 receptor tyrosine kinase 2 (ERBB2) and aromatase enzyme (CYP19A1). [7] [8] [9] And many RBPs are involved in more than one of these processes, such as NOVA2 regulating both alternative splicing and poly(A) site. These roles are critical for normal human physiology, since RBP functions are associated with genetic and somatic disorders, such as neurodegeneration, autoimmunity, and cancer. The regulatory roles of RBPs are also affected by the subcellular localization of RBPs and their RNA targets, since post-transcriptional regulation occurs in both membrane- and phase-separated sub-cellular compartments. [5] For instance, Trichostatin A (TSA) and 5-Aza 2' deoxycytidine (AZA), two well-characterized pharmacologic inhibitors of histone deacetylation and DNA methylation, can decrease the cytoplasmic abundance of HuR, so that less HuR will be able to bind with the 3'UTR of ERα mRNAs to stabilize them, and the level of ERα mRNA will decrease. [7]

### 1.2 Enhanced UV-crosslinking and immunoprecipitation followed by sequencing

As RBPs play essential roles in cellular physiology by interacting with target RNA molecules, a pivitol step in the functional and mechanistic characterization of the target RBP is to identify the RNAs it binds. [10] To discover the mechanisms through which RBPs affect RNA, technologies such as RNA immunoprecipitation (RIP) and UV-crosslinking and immunoprecipitation (CLIP) that comprehensively identify the RNA substrates each RBP interacts with are widely used. [11] However, these technologies have limitations such as technically demanding, low yield, low resolution, and high experimental failure rates. A novel technology called enhanced UV-crosslinking and immunoprecipitation followed by sequencing (eCLIP-seq)

has been developed in recent years. Briefly, in this approach, the cells are crosslinked by UV irradiation, which can only introduce covalent bonds between the RBP and the RNAs that are directly contact. Then the RNAs are fragmented by RNase. The target RBP and the RNA fragments it binds are pulled down by immunoprecipitation. Then adaptors with barcodes to demultiplex the samples after sequencing are added to 3' of the binding-RNA fragments. The complex of RBP and RNAs is resolved on an SDS-PAGE gel and them transferred to a nitrocellulose membrane. The membrane areas containing the target RNA-protein complexes are cut into small pieces, and proteinase K digestion is used to remove the RBP from the complexes. The extracted RNA fragments are reverse transcripted and subject to 3' DNA adaptor ligation followed by high throughput sequencing. Mapping and analyzing the sequencing reads reveal where the target RBP binds in the transcriptome. [10] eCLIP improves these adaptor ligations to >70% efficiency compared to previous CLIP and individual-nucleotide resolution crosslinking and immunoprecipitation (iCLIP), and significantly increasing the number of non-PCR duplicates reads that can be obtained after high-throughput sequencing. [11] Thus, it provides a robust and standardized framework for large-scale generation of transcriptome-wide binding maps for RNA binding protein. [11]

### 1.3 Pregnane X Receptor

PXR is a well-known transcription factor that belongs to the nuclear receptor superfamily encoded by Nuclear Receptor Subfamily 1 Group I Member 2 (NR1I2 ) gene. Its primary function is to sense the presence of foreign toxic substances and subsequently upregulate the expression of

genes that are involved in the clearance of these foreign substances [12, 13]. It can be activated by a range of compounds like rifampicin and induce the expression of its target genes encoding cytochrome P450 enzymes, conjugation enzymes, and transporters. [14, 15]. PXR activated by its ligands can increase the level of cytochrome P450 enzymes 3A4 (CYP3A4) [16], the most abundant phase I enzyme in the liver and colon. Activated PXR can also strongly upregulated multidrug resistance 1 (MDR1) in the colon, which is a protective physiological barrier of cells and responsible for exporting hydrophobic compounds in several organs. [17]

Beyond canonical xenobiotic-sensing function, several other functions of PXR have been described. The role of PXR in regulating glycogen, energy homeostasis, and insulin resistance has been revealed. [18] PXR is involved in atherosclerosis development and vascular functions. [19, 20] And PXR has been shown to play an essential role in maintaining the intestinal wall integrity through affecting inflammation through regulating NF-nB, TNF-a, and TLR4. [21, 22]

PXR is also pivotal in cell proliferation, migration, and apoptosis. As for cell proliferation, PXR serves a dual role. PXR can upregulate cell cycle suppressor gene p21 to inhibit the G2-M phase progression of the cell cycle or decrease p27 and p130 to enhance the G1-S phase transition of the cell cycle. [23, 24] Besides, PXR can induce cell migration through the GADD45b-p38, HNF4a-IGFBP1, and FGF19 regulatory axes.[25, 26] And PXR affects apoptosis through upregulating anti-apoptotic genes like BAG3 and downregulating pro-apoptotic genes such as p53. [27]

So far, most studies of PXR focus on the transcriptional regulatory function. However, PXR has also been reported to post-transcriptionally regulate RNAs directly. PXR was shown to negatively regulate TLR4 signaling which contributes to the pathogenesis of necrotizing enterocolitis. This regulation of TLR4 signaling is not dependent on TLR4 gene transcription. Instead, PXR decreased the half-life of TLR4 mRNA by 47% after rifampicin treatment compared to untreated cells. [22] This novel discovery encouraged us to conjecture that PXR is also an RBP and can directly regulate RNA metabolism.

### 1.4 Overview of thesis

In this study, we performed a comprehensive investigation into the RNA bind landscape of PXR by using eCLIP-seq. Specifically, we used eCLIP-seq to detect RNAs bind to PXR in both PXR-humanized mouse liver and colorectal cancer cell lines with and without rifampicin treatment. We found that PXR can bind to thousands of mRNAs. Moreover, PXR's binding peaks are highly enriched in the 3'UTR of each mRNA. We have further identified the pathways that are enriched with genes that have PXR binding signals in their 3'UTR. Finally, we performed the correlation analysis between the expression levels of PXR and its eCLIP targets using the TCGA database. In conclusion, our study suggests PXR may regulate the RNA stability of genes in signaling pathways such as cell migration, motility and apoptosis, platelet degranulation, blood vessel remodeling, glucose metabolism, and triglyceride.

## 2.0 Experimental Section

### 2.1 Methods

#### 2.1.1 Cell culture and mice

Human colorectal cancer cell line LS174T and PXR knockout LS174T were kindly provided by Prof. Sridhar Mani (Albert Einstein College of Medicine) and cultured with Eagle's Minimum Essential Medium (EMEM) supplemented with 10% FBS and 1% penicillin-streptomycin.[28]

Mice with humanized PXR were kindly provided by Prof. Xiaochao Ma (University of Pittsburgh) which were pretreated with 10 mg/kg rifampicin or DMSO once a day for four days through intraperitoneal injection.

#### 2.1.2 SDA-PAGE and Western Blot

Cells were lysed in RIPA lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 1x protease inhibitor cocktail). Total protein concentrations were measured with a BCA protein assay kit (ThermoFisher, 23225). 20 µg of protein with 2x SDS sample buffer were loaded and then transferred to PVDF membranes (Bio-Rad, 162-0177). Membranes were blocked in 5% non-fat milk for one hour and then incubated in primary antibodies diluted 1:1000 at 4 ℃. PXR antibody was purchased from Santa Cruz (sc-48340). On the second day, membranes were incubated in horseradish peroxidase-conjugated secondary

antibodies for one hour (Anti-mouse IgG: ThermoFisher 31430). Bands were detected by ECL (electrogenerated chemiluminescence) solution (ThermoFisher, 32106) on films

### 2.1.3 Immunoprecipitation

The cell lysate was prepared as described above. The cleared lysates were incubated overnight at 4 ℃ with the corresponding antibodies. On the second day, prewashed beads were added to the lysate and incubated at 4 ℃for one hour. Then beads were washed three times in RIPA buffer. 1x SDS buffer was used to prepare the protein samples for later western blot tests.

### 2.1.4 eCLIP-seq library preparation

eCLIP-seq was performed as previously reported. [11] Briefly, proteins and the binding RNAs were covalently linked by UV irradiation, followed by lysis in 0.5 mL of iCLIP lysis buffer, digestion with RNase I for a limited time, immunoprecipitation of RNA-protein complexes with specific primary antibodies of interest (5 μg per sample) using magnetic protein A beads. After dephosphorylation with FastAP (ThermoFisher) and T4 PNK (NEB), a barcoded RNA adapter was ligated to the 3′ end of RNA fragments on beads (T4 RNA Ligase, NEB). Samples were then run on prepared SDS-page gels and transferred to nitrocellulose membranes, and a region from the protein size to 75 kDa above was cut and treated with proteinase K (NEB) to extract RNA. RNAs were reverse transcribed with AffinityScript (Agilent), and treated with ExoSAP-IT (Affymetrix) to remove the excess oligonucleotides. A DNA adapter (containing a random-mer of 10 ($N_{10}$) random bases at the 5′ end) was then ligated to the 3’ end of DNA fragments (T4 RNA Ligase, NEB). After cleanup (Dynabeads MyOne Silane, ThermoFisher), a small amount of the DNA

fragments from each sample was subjected to qPCR and PCR amplification (Q5, NEB). The size of DNA fragments was selected through agarose gel electrophoresis. Samples were then sequenced on the Illumina NextSeq as 2x paired-end 150bp.

### 2.1.5 eCLIP-seq data processing

After standard NextSeq demultiplexing, eCLIP libraries with distinct barcodes were demultiplexed using custom scripts. Then we use CLIP Tool Kit (CTK) software package [29] to analyze the data. Briefly, 3' RNA adaptor sequences were trimmed, and reads less than 20 bp were discarded. Then exact PXR duplicates were removed so that if multiple reads had the same sequence, only one would be kept. The random barcodes were stripped. Mapping was then performed against the full human genome (hg38) or the full mouse genome (mm10) with BWA. Then the output sam files were transformed to bed files and only unique mappings (with MAPQ >=1) and a minimum mapping size of 18 nt would be kept. Due to abundant repetitive RNAs like rRNAs, we removed reads from rRNA and other repetitive RNA to avoid false-positive results. Then the PCR duplicates were removed again since some reads have slight differences due to sequencing errors occurring in the random barcode. Then we checked the summary file for the percentage of tags mapped to different regions. Bedgraphs for visualization in the genome browser were generated and loaded into UCSC genome browser for visualization of peaks at each genomic position. Peak calling was performed to get bed files with information of significant peaks. The number of reads within each peak was count to evaluate the correlation between replicates by Pearson correlation statistics and calculate fold enrichment. The significant difference of peaks among groups was detected by Yates' Chi-Square test.

### 2.1.6 Motif analysis

*De novo* motif analysis for eCLIP-seq peaks were performed with HOMER's findMotifs program (-len 6 -rna), where the sequences of eCLIP-seq peaks were compared against a set of background 'clusters'. Three random regions which were same-sized as the real eCLIP-seq peak were selected for each real peak corresponding to the same type of genic region.

## 3.0 Result

### 3.1 Expression pattern of PXR in tumor and normal tissue

PXR is a nuclear receptor abundant in the colon and liver. [30] The TCGA data shows the significant expression of PXR in colorectal cancer cells, normal liver tissue, and bile duct (Figure 1A and 1B). We confirmed that PXR is highly expressed in LS174T cell line and LS180 cell line, which are colorectal cancer cell lines, and in hepG2 cell line, which is a liver cancer cell line, at both RNA and protein level (Figure 1C and 1D). Since PXR is significantly upregulated in colorectal cancer compared to its normal counterpart, and PXR has a higher expression in the LS174T cell line than in the LS180 cell line, we select LS174T cancel cell line to elucidate the role of PXR in post-transcriptional regulation. Besides, because PXR is highly expressed and plays a pivotal role in the liver, we also select mouse liver with humanized PXR to identify the RNA-binding targets of PXR. To further determine the feasibility of performing eCLIP-seq in LS174T cell line and mouse liver with humanized PXR, we used immunoprecipitation (IP) to show that PXR can be successfully pulled down by antibody, which is a prerequisite for the success of this experiment (Figure 1E and 1F ).

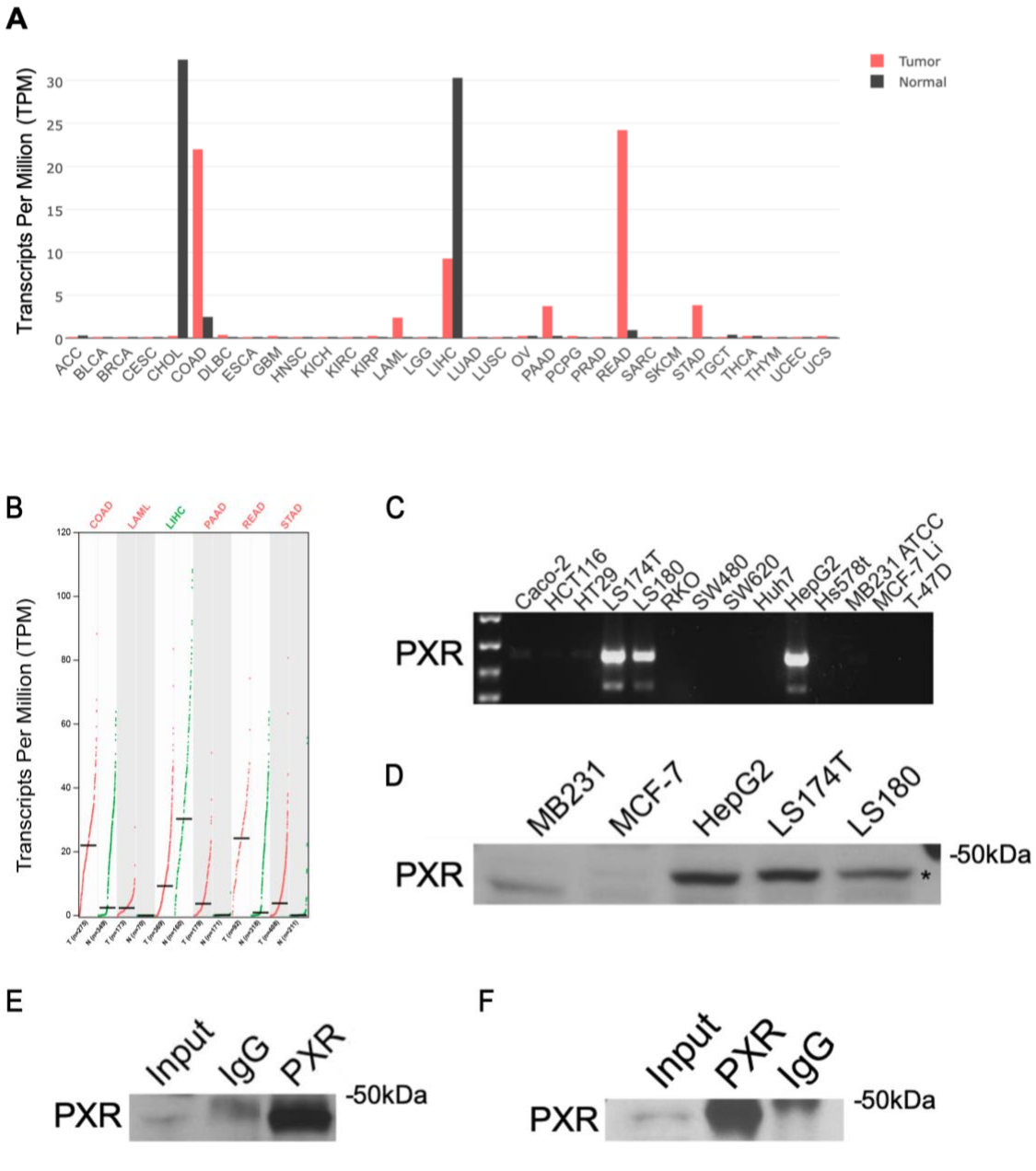



**Figure 1. Expression pattern of PXR in tumor and normal tissue**

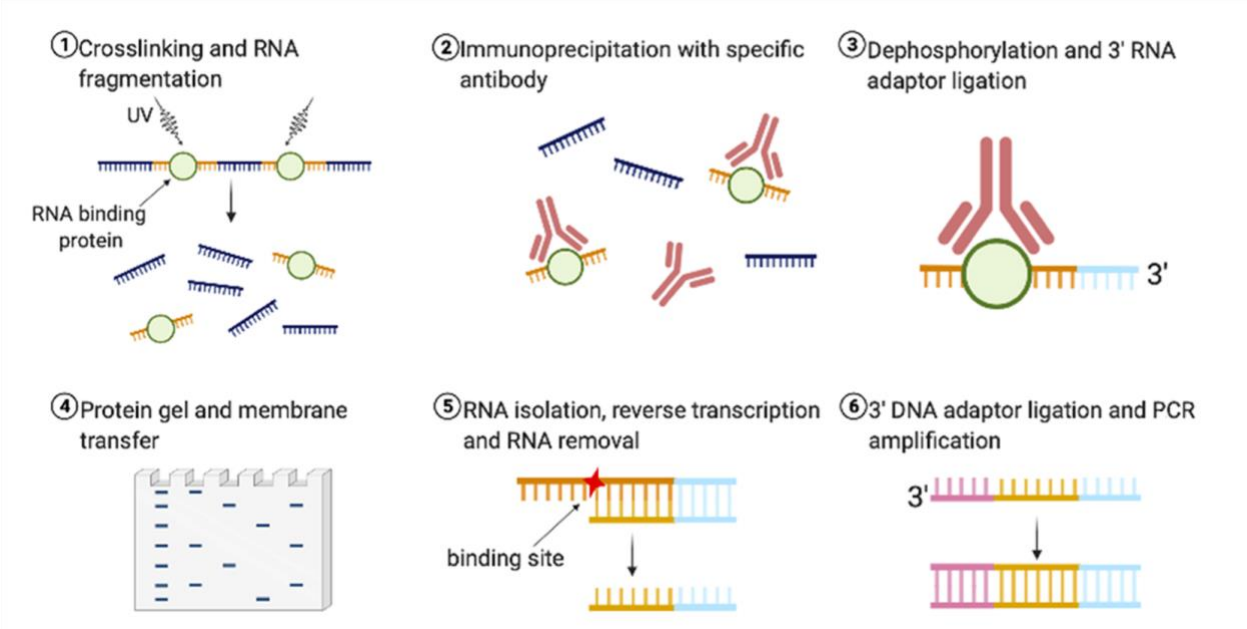
A. The gene expression profile of PXR across all tumor samples and paired normal tissues. The y axis indicates The median expression of certain tumors and paired normal tissue. B. The gene expression profile across tumor samples and paired normal tissues which highly express PXR. Each dots represent the expression of samples C. PCR and agarose gel electrophoresis show RNA levels of PXR in different cell lines D. Western blot shows protein levels of

PXR in different cell lines. E. Immunoprecipitation of PXR in LS174T cell line. F. Immunoprecipitation of PXR in mouse liver tissues.

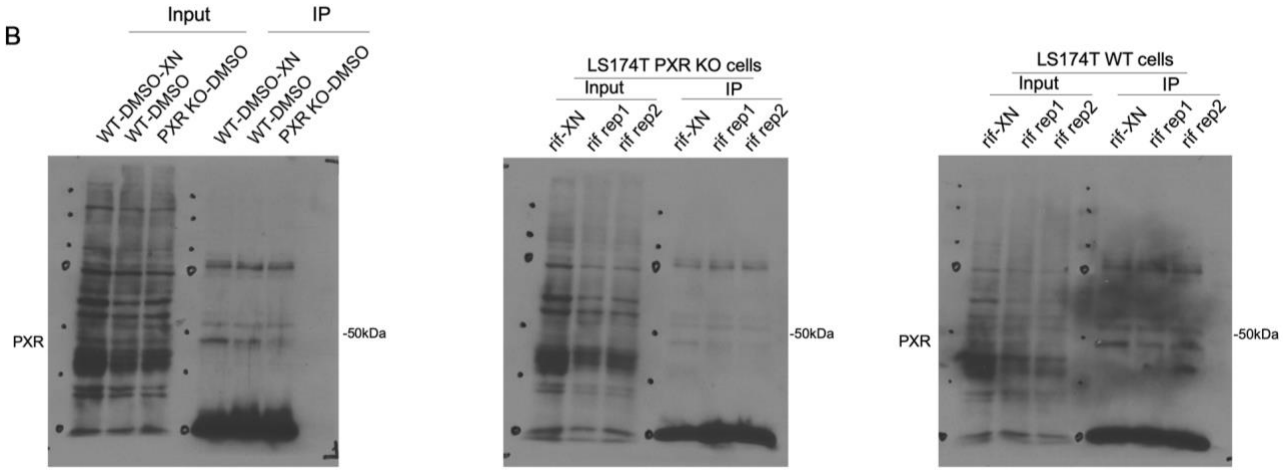
### 3.2 eCLIP-seq process of PXR in colorectal cancer cell line and mouse liver

We next utilized eCLIP to identify its transcriptome-wide RNA targets. [11] LS174T cells and mouse liver cells were UV-crosslinked, lysed, and treated with RNase for a limited period. Then a few lysates were stored as input samples. IP was performed in remaining lysate to precipitate PXR-RNA complexes using antibodies that specifically recognize PXR. Next, RNA fragments binding with PXR were ligated to 3' RNA adaptors which include barcodes for demultiplexing after high-throughput sequencing. A fraction of the samples were saved for the western bolt to test the IP efficiency (Figure 2B and 2C). All remaining samples were used to extract RNA. Next, the RNA fragments were reverse transcribed and 3'DNA adaptors were added which include short random sequences to remove PCR duplicates during later analysis. A small amount of the DNA samples was used for qPCR to detect the quantity of DNA fragments and be a reference to the cycle number of PCR amplification. As we expected, input samples had a higher amount of DNA fragments and lower CT value, while the samples which were not UV-crosslinked had the highest CT value since RNA fragments had separated from PXR during the experiment (Figure 2D and 2E). Finally, we performed PCR amplification and checked the size distribution to make sure the length of the fragments was between 175 and 350 bp (Figure 2F and 2G).

A


1 Crosslinking and RNA fragmentation
UV
RNA binding protein
2 Immunoprecipitation with specific antibody
3 Dephosphorylation and 3' RNA adaptor ligation
3'
4 Protein gel and membrane transfer
5 RNA isolation, reverse transcription and RNA removal
binding site
6 3' DNA adaptor ligation and PCR amplification
3'

B


Input
IP
WT-DMSO-XN
WT-DMSO
PXR KO-DMSO
WT-DMSO-XN
WT-DMSO
PXR KO-DMSO
PXR
-50kDa
LS174T PXR KO cells
Input
IP
rif-XN
rif rep1
rif rep2
rif-XN
rif rep1
rif rep2
-50kDa
LS174T WT cells
Input
IP
rif-XN
rif rep1
rif rep2
rif-XN
rif rep1
rif rep2
PXR
-50kDa

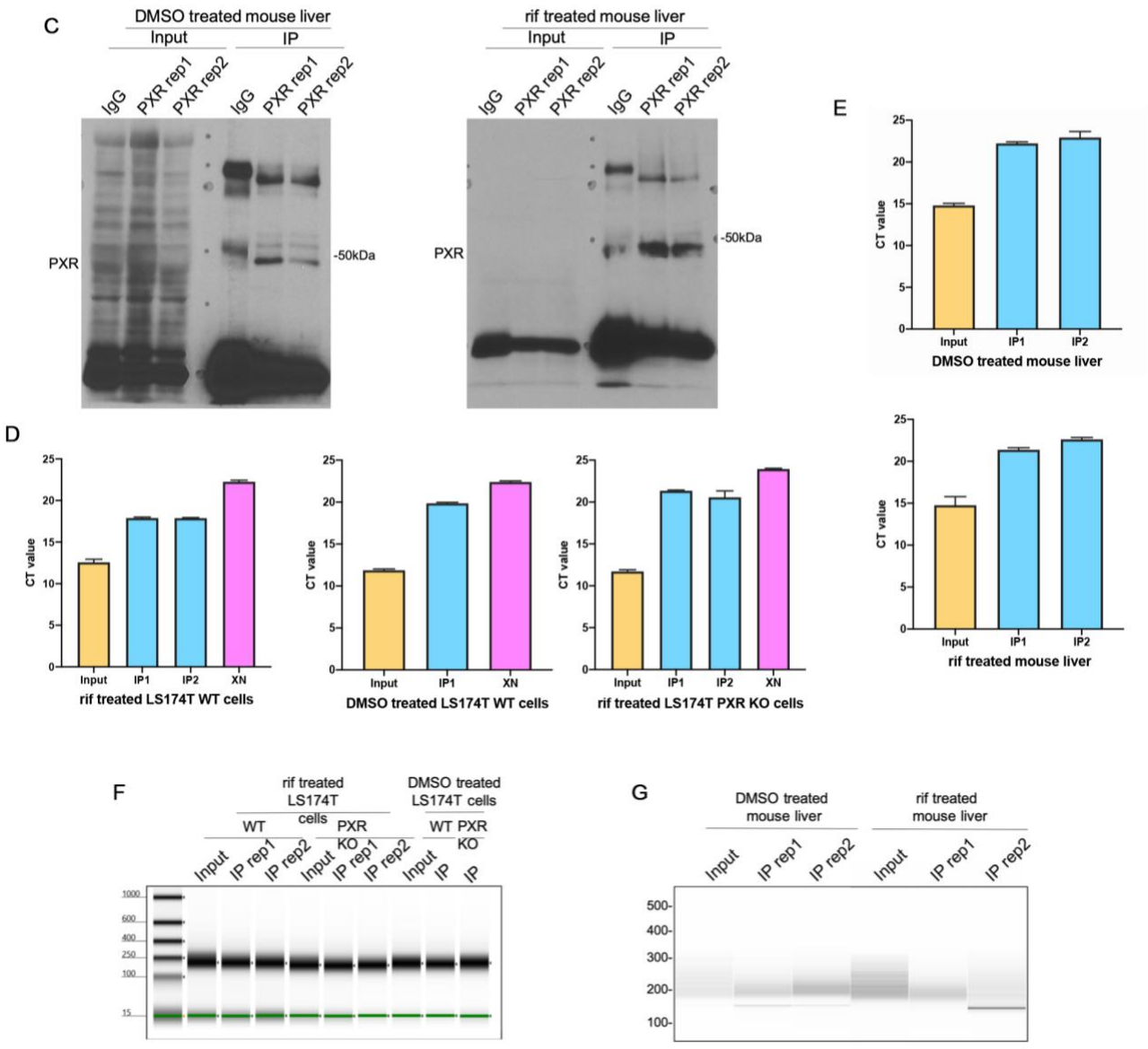



**Figure 2. eCLIP-seq process of PXR in colorectal cancer cell line and mouse liver**

A. Schematic of eCLIP-seq protocol. (Created with BioRender.com) In brief, LS174T cells and mouse liver cells were subjected to UV-mediated crosslinking, lysis, and treatment of RNase, followed by immunoprecipitation of PXR and its binding RNAs. RNA fragments were subjected to 3'RNA adaptor ligation,reverse-transcription, and 3'DNA adaptor ligation to generate eCLIP libraries for high-throughput Illumina sequencing. B. Immunoprecipitation of PXR for eCLIP-seq in LS174T cell line. XN is the sample without UV-crosslinking C. Immunoprecipitation of PXR for eCLIP-seq in mouse liver. D. qPCR of DNA fragments obtained from eCLIP-seq in LS174T cell line. F. qPCR of

DNA fragments obtained from eCLIP-seq in mouse liver. E. Size distribution of eCLIP-seq libraries of LS174T cell line. F. Size distribution of eCLIP-seq libraries of mouse liver.

### 3.3 RNA-binding profile of PXR in colorectal cancer cell line and mouse liver

In total, nine eCLIP libraries of PXR in the LS174T cell line were sequenced, including samples of WT LS174T cells treated with DMSO or rifampicin and those of PXR knockout LS174T cells for control. And six eCLIP libraries in mouse liver with humanized PXR were sequenced, including mouse livers also treated with DMSO or rifampicin. All libraries passed our quality control and CLIP Tool Kit (CTK) was utilized to analyzed the sequencing data. For rifampicin treated LS174T WT and PXR knockout cells, we generated replicate libraries of IP groups, which were highly correlated across the same samples ($R^2$ ~ 0.7, Figure 3A). In mouse livers, we also had replicates for both DMSO treated and rifampicin treated mouse liver IP groups, but due to the individual difference among different mice, the correlation ($R^2$~0.4) was not as high as that of the cell line. While overall the peaks are still positively corrected especially for the peaks with a larger number of reads mapped to (Figure 3B).

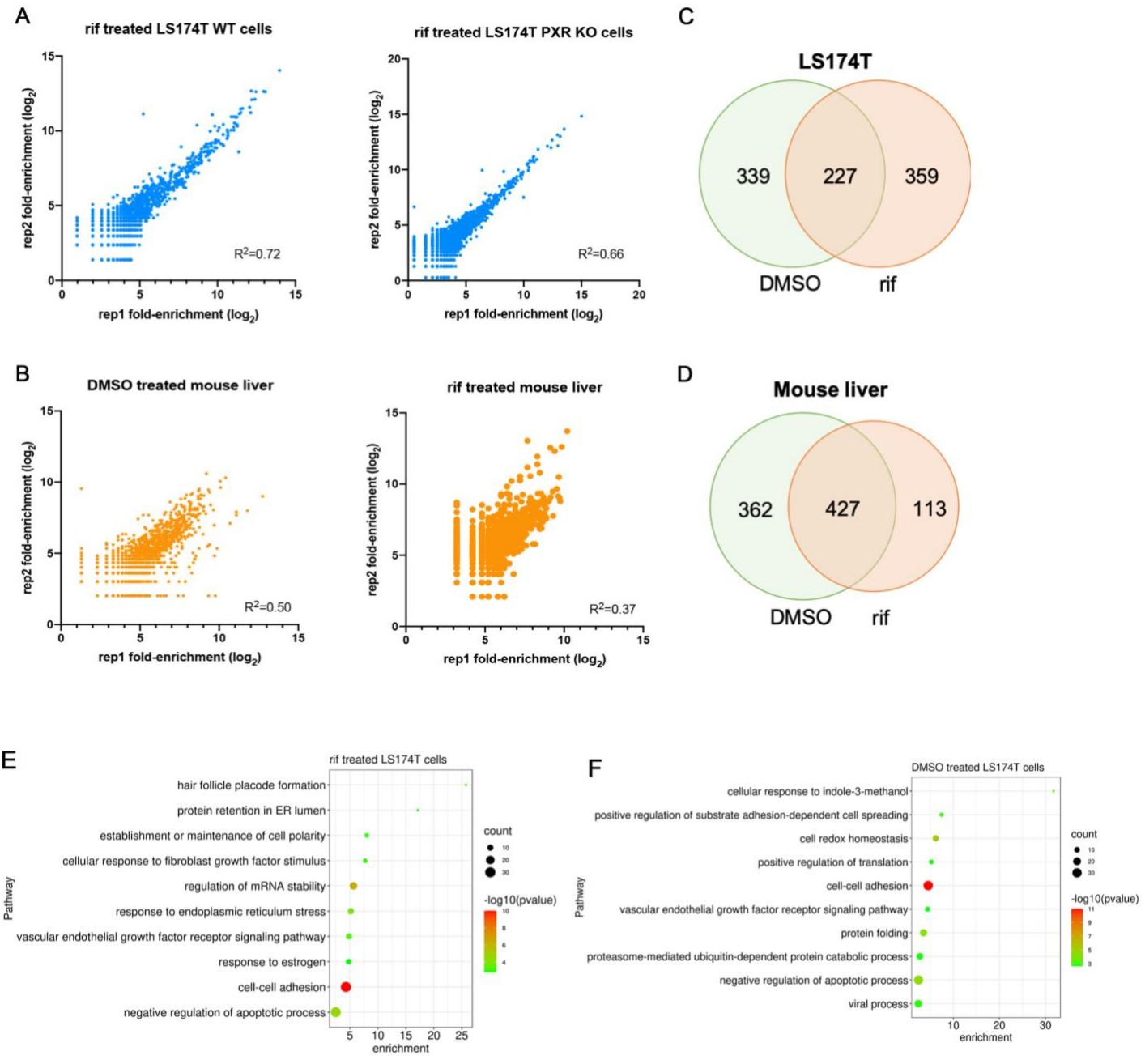
In the LS174T cell line, after compared with input samples and samples of PXR knockout LS174T cells, we got 1578 peaks encoding 789 genes from DMSO treated group and 1067 peaks encoding 540 genes from rifampicin treated group. Among them, 227 genes were overlapped between DMSO treated and the rifampicin treated group (Figure 3C). In mouse liver cells, we compared IP groups with input groups and got 1223 peaks encoding 566 genes from DMSO treated

group and 1276 peaks encoding 586 genes from the rifampicin treated group. 427 genes were overlapped between DMSO treated and rifampicin treated group (Figure 3D).

To identify the potential pathways post-transcriptionally regulated by PXR, we performed GO (Gene Ontology) analysis of each group separately. We found out RNAs binding to PXR in LS174T cells without rifampicin treatment were enriched in several pathways including translation, protein folding, ubiquitin-dependent protein catabolic process, cell-cell adhesion, cell spreading, apoptosis and vascular endothelial growth factor receptor signaling pathway (Figure 3F). After rifampicin treatment, cell-cell adhesion was still the most significant pathway, and apoptosis and vascular endothelial growth factor receptor signaling pathway also showed up. And we also observed a significant enrichment in several pathways other than those in DMSO treated group, like regulation of mRNA stability, response to endoplasmic reticulum stress, cellular response to fibroblast growth factor stimulus, and response to estrogen (Figure 3E). Due to differences in gene expression and tissue specificity, the GO analysis results were different in mouse liver groups from LS174T cell groups. The binding RNAs are enriched in acute-phase response, translation, cholesterol metabolic process, and triglyceride metabolic process related signaling pathways (Figure 3G and 3H). Without rifampicin treatment, there was also enrichment in fibrinolysis, response to peptide hormone and cytokine, and steroid metabolic process pathways (Figure 3). And after the treatment, we then observed blood coagulation, plasminogen activation, apoptosis, and response to glucagon stimulus and drug(Figure 3H).

To further understand the interaction between PXR and RNAs, we utilized de novo motif analysis and the result shows that the possible RNA binding motif could be'UGAUG' (Figure 3I).

There was a higher significance in the LS174T cell group ($p = 10^{-35}$ for DMSO treated group and $10^{-56}$ for rifampicin treated group). In mouse liver, the p values are not that significant ($10^{-8}$ for both groups), while the motif seemed consistent as in LS174T cells.



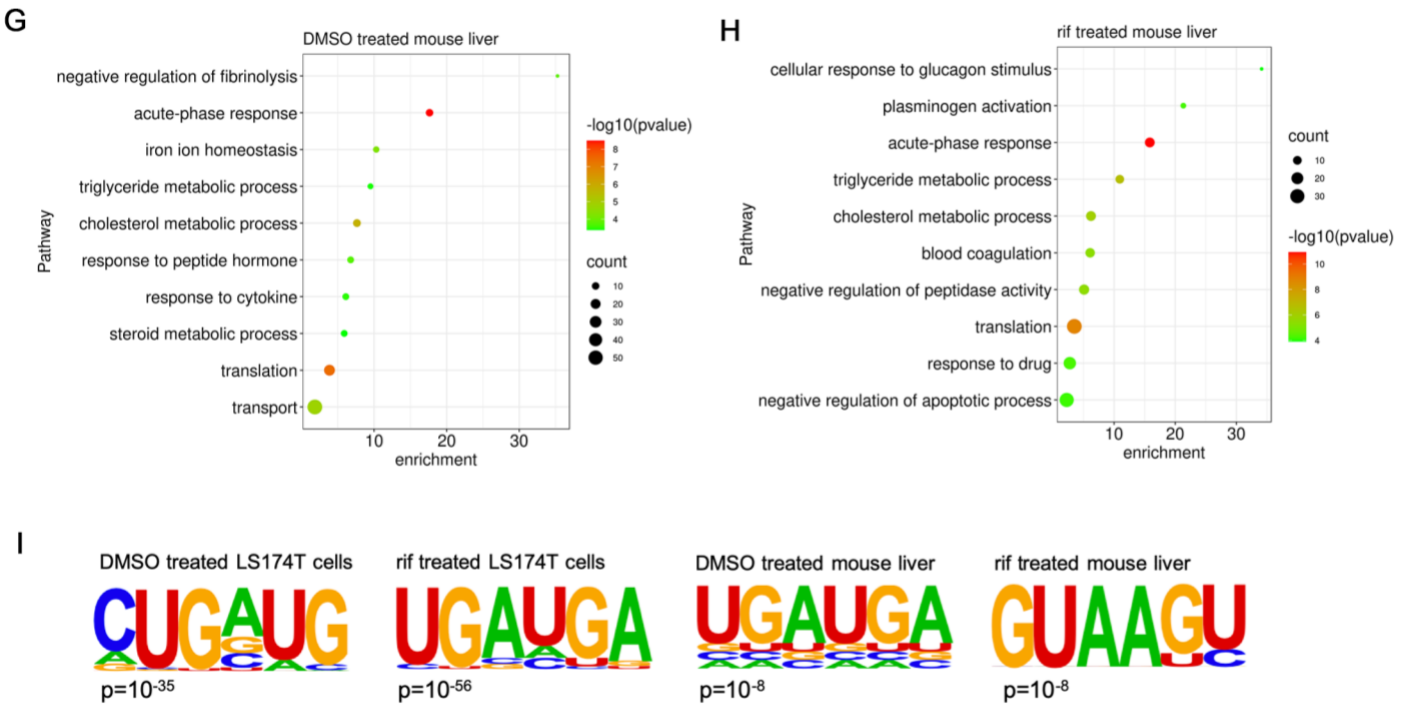



**Figure 3. RNA-binding profile of PXR in colorectal cancer cell line and mouse liver**

A. Correlation between region based normalized reads number in eCLIP-seq of LS174T cells. B. Correlation between region based normalized reads number in eCLIP-seq of mouse liver. C to D. Number of genes obtained from different groups of eCLIP-seq. E to H. GO term analysis of eCLIP-seq targets from different groups. I. Motif analysis of peaks from different eCLIP-seq groups.

### 3.4 RNA-binding preference of PXR in colorectal cancer cell line and mouse liver

Next, we determined RNA binding region specificities by using two distinct methods. During the conventional CTK pipeline, we got the reads in enriched peaks and used the provided tool to quantify the numbers of reads binding to a different region on RNA including the coding sequence (CDS), 5'untranslated region (5'UTR), 3'untranslated region (3'UTR), intron, upstream and downstream 10 kbp of genes and deep intergenic region (Figure 4A and 4B). The results

showed that in LS174T cells, PXR displayed a strong 3’UTR preference (Figure 4A). While in mouse liver, the percentage of reads mapping to 3’UTR is just slightly higher in IP groups than in input groups (Figure 4B). In eCLIP-seq results, there are normally some false positive peaks for snoRNAs, which locate in the intron region of mRNAs or lncRNAs. Since we have PXR knockout LS174T cells to remove those false positive peaks, we rely more on the results getting from the LS174T cell line that PXR binding peaks are enriched in the 3'UTR of RNAs.i

Besides, we generated our own pipeline to map reads to CDS, 5'UTR, and 3'UTR and calculate the fold enrichment in IP groups normalized to their input samples. In both LS174T cells and mouse livers, we found an increase in 3'UTR binding of PXR after rifampicin treatment (Figure 4C and 4D).

3’UTR plays important role in mRNA stability. RNA binding proteins can bind to 3’UTR and stabilize the mRNA by competing with exosomes or destabilize mRNAs by recruiting them. [6] According to a published paper, PXR can actually destabilize the TLR4 mRNA after rifampicin treatment. Therefore, we thought that PXR can regulate mRNA stability by binding to their 3’UTR. Then to delve deeper into the specific mRNA targets of PXR, we obtained the genes with PXR binding on their 3’UTR and performed pathway analysis again. In LS174T cells, those genes were enriched in pathways related to cell migration and motility, platelet degranulation and actin filament bundle assembly (Figure 4E). And in mouse livers, the enriched pathways were insulin-like growth factor receptor signaling pathway, glucose metabolism, triglyceride metabolism, blood vessel remodeling and apoptosis (Figure 4F).

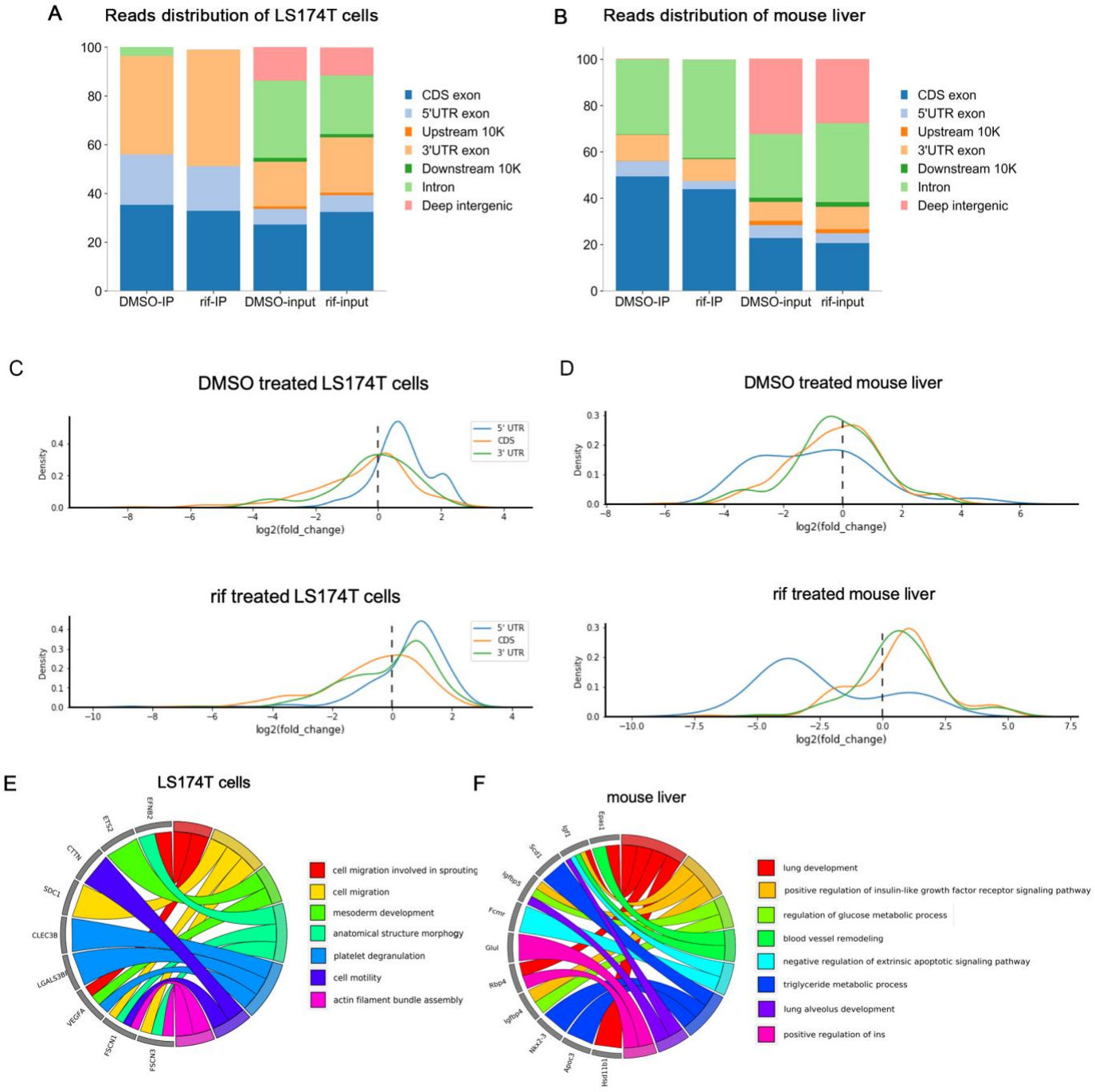



**Figure 4. RNA-binding preference of PXR in colorectal cancer cell line and mouse liver**

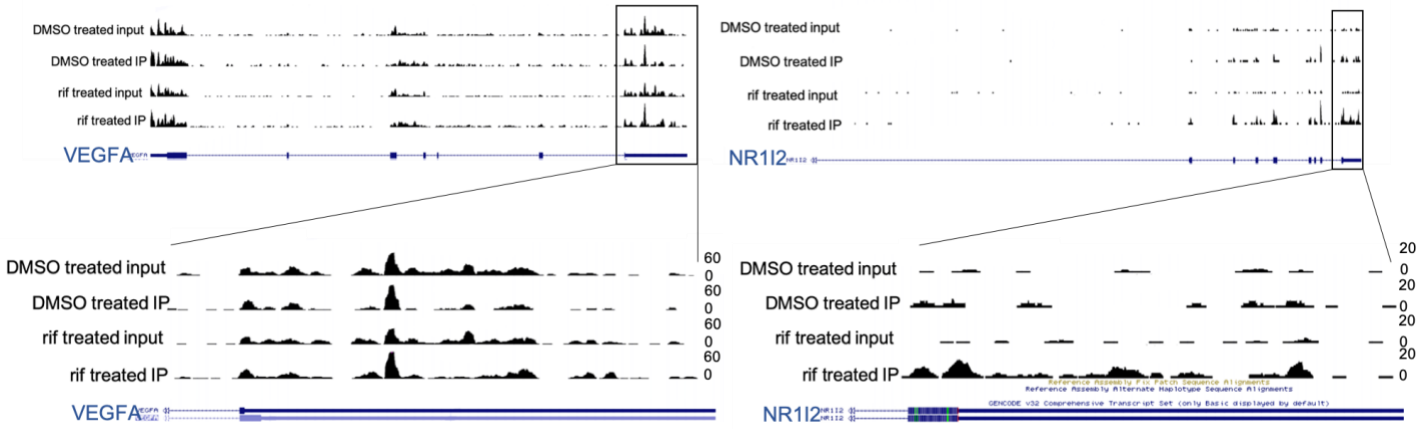
A to B. The distribution of reads in significantly enriched peaks from different eCLIP-seq groups. C to D. Binding preference of reads from different eCLIP-seq groups. E to F. GO term analysis of genes with 3'UTR binding of PXR from different eCLIP-seq groups.

### 3.5 RNA-binding targets of PXR in colorectal cancer cell line and mouse liver

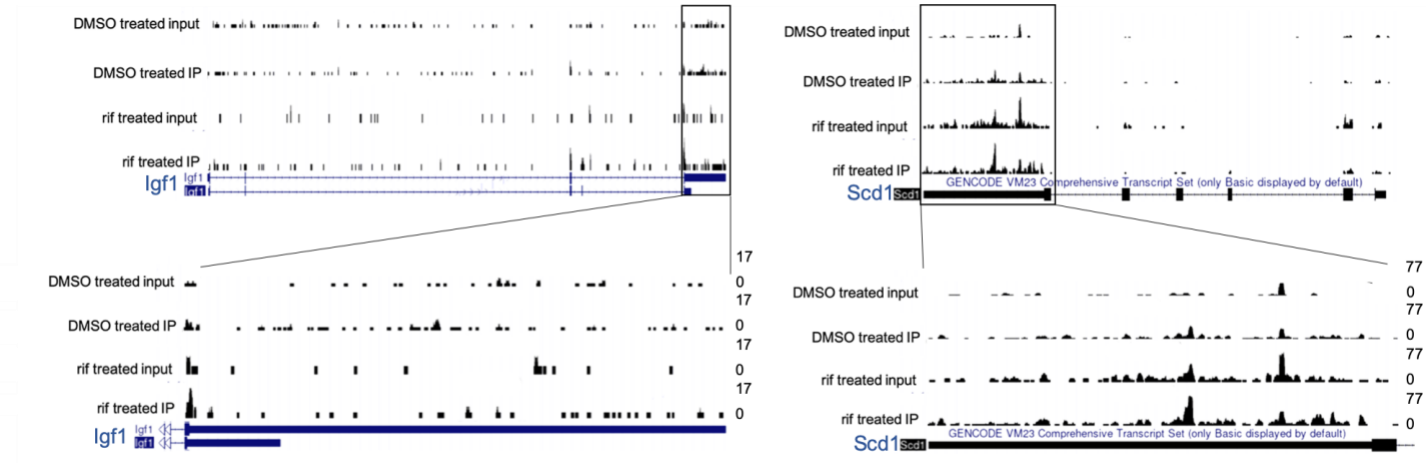
Based on our finding that PXR binds to the 3'UTR and regulate the RNA stability, we checked more details about the binding pattern of PXR. VEGFA belongs to several enriched pathways like cell migration, mesoderm development, and platelet degranulation (Figure 4E). We checked the binding peaks of PXR on VEGFA and found that there were enriched peaks within 3'UTR in both DMSO treated and rifampicin treated IP group. And the signal was even higher in the rifampicin-treated group compared to DMSO treated group (Figure 5A). And VEGFA has been reported to be regulated by rifampicin, another PXR ligand, through a PXR-dependent pathway in colon cancer cells. [31] Interestingly, PXR can bind to its own 3'UTR in LS174T cells (Figure 5B). Activated PXR was shown to destabilize PXR mRNA by upregulating miR-18a-5p, which can induce the degradation of PXR mRNA. [32] Binding to its own 3'UTR might partially contribute to this regulation. And in mouse liver groups, we checked the peaks of Igf1 and scd1. Igf1 is related to lung development, insulin-like growth factor receptor signaling pathway, glucose metabolism and apoptosis. And scd1 belongs to the triglyceride metabolism signaling pathway. As we expected, the binding peaks are enriched in IP groups and higher in rifampicin-treated groups (Figure 5B).

To validate the regulation of PXR on its RNA targets, we detected the correlation of PXR expression level and its eCLIP-seq targets. TCGA database showed that in colon adenocarcinoma and the normal counterpart, the level of VEGFA was positively corrected with that of PXR (Figure 5C). While in liver hepatocellular carcinoma, there was also a positive correction between IGF1 and PXR (Figure 5D). These results at least suggested that PXR is related to the expression of these eCLIP-seq targets.

A




B



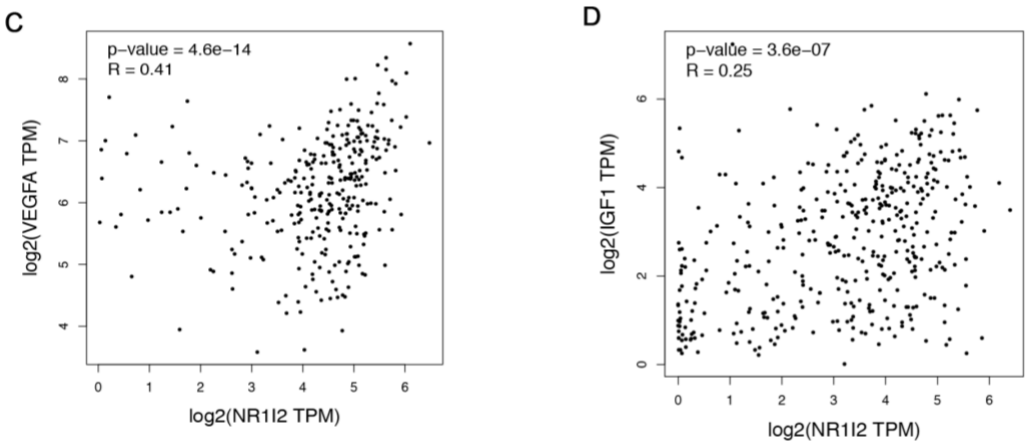



**Figure 5. RNA-binding targets of PXR in colorectal cancer cell line and mouse liver**

A to B. Read density of PXR in full-length as well as 3'UTR of VEGFA (A left), PXR (A right), Igf1 (B left), Scd1 (B right) mRNAs. C to D. Correlation of PXR expression level and the level of its eCLIP-seq targets (C. VEGFA D. IGF1) from TCGA database.

## 4.0 Discussion and future plan

In our study, we demonstrated that PXR can bind to mRNAs in both the colorectal cancer cell line and the mouse liver. These mRNAs were enriched in a variety of signaling pathways including translation, protein folding, protein catabolic process, cell-cell adhesion, cell spreading and apoptosis, the response to vascular endothelial growth factor, endoplasmic reticulum stress, and estrogen. The potential bind motif of PXR was 'UGAUG'. Some well-known RBPs, like HuR, binds to the conventional 'AUUUA' motif on AU-rich elements of the 3'UTR to regulate the mRNA stability. [33] While in fact many other RBPs have distinct motifs. The binding motif of IFIT2 and BOLL, two RBPs bind to 3'UTR and stabilize mRNA, is 'AGUGUA'. [5]

PXR shows a 3'UTR binding preference mainly in the colorectal cancer cell line. In the mouse liver, 3'UTR binding is also enriched after rifampicin treatment like in the colorectal cancer cell line. Since we didn't have PXR knockout mouse liver as control, some false positive peaks of snoRNAs would also be annotated to the intron region of mRNAs and lncRNAs, which led to a higher enrichment of reads in the intron region and lower in 3'UTR. In order to get cleaner and more convincing data, we would either add control groups of mouse liver or remove those snoRNA peaks by some computational tools.

As for the 3'UTR binding targets of PXR, we showed that these targets were enriched in cell migration and motility, platelet degranulation, and actin filament bundle assembly related pathways in colorectal cancer cells (Figure 4E) and in apoptosis, glucose metabolism, triglyceride metabolism, blood vessel remodeling and insulin-like growth factor receptor related signaling

pathway in mouse liver (Figure 4F). Previous studies have demonstrated that some of these pathways can be regulated by PXR. Those pathways include cell migration and apoptosis, glucose metabolism and triglyceride metabolism, vascular functions and the function of the platelet. Our analysis of the TCGA database indicates that the expression level of VEGFA and Igf1, whose 3'UTRs can bind with PXR, are positively correlated with the PXR level. It is promising that PXR can regulate the stability of VEGFA mRNA and Igf1 mRNA. While this hypothesis needs to be tested by the future study, our data suggest that the post-transcriptionally regulatory function of PXR may contribute to its regulation of vascular function and glucose metabolism. [18, 19, 23-25]

In the future, we plan to perform RNA-seq to find out if the level of PXR eCLIP-seq targets is post-transcriptionally regulated by PXR. And we will use luciferase reporter assay and Actinomycin D chase assay to determine if this regulation is through an effect on the RNA stability. Some RBPs regulate RNA stability by recruiting other RBPs. For example, RBMS1 can stabilize mRNA by recruiting HuR. [34] Since the de novo RNA binding motif of PXR is not the traditional sequence of AU-rich element (i.e., 'AUUUA'), it is possible that PXR can form a complex with some AU-rich element-binding RBPs that is well-studied to regulate mRNA stability. We may perform co-immunoprecipitation (Co-IP ) followed by mass spectrometry and RNA pull-down-coupled mass spectrometry to find out if PXR regulates RNA stability by recruiting other RBPs. Moreover, we will identify the RNA binding domain of PXR protein using computational tools like catRAPID. catRAPID is an RNA binding prediction tool that combines secondary structure, hydrogen bonding, and van der Waals' interaction between proteins and RNAs to predict their interaction profile [35]. If we successfully identify the RNA binding domain of PXR, we plan to generate the mutant PXR, which can separate PXR's binding to DNA and RNA. In this way, we

hope to determine the PXR functions and phenotypes mediated only by PXR's post-transcriptional regulation and elucidate the underlying mechanism of PXR in regulating RNA stability.

## Bibliography

[1] G.J. Goodall, V.O. Wickramasinghe, RNA in cancer, Nat Rev Cancer (2020).
[2] J.L. Hargrove, F.H. Schmidt, The role of mRNA and protein stability in gene expression, FASEB J 3(12) (1989) 2360-70.
[3] R. Lemaire, J. Prasad, T. Kashima, J. Gustafson, J.L. Manley, R. Lafyatis, Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins, Genes Dev 16(5) (2002) 594-607.
[4] A. Bevilacqua, M.C. Ceriani, S. Capaccioli, A. Nicolin, Post-transcriptional regulation of gene expression by degradation of messenger RNAs, J Cell Physiol 195(3) (2003) 356-72.
[5] E.L. Van Nostrand, P. Freese, G.A. Pratt, X. Wang, X. Wei, R. Xiao, S.M. Blue, J.Y. Chen, N.A.L. Cody, D. Dominguez, S. Olson, B. Sundararaman, L. Zhan, C. Bazile, L.P.B. Bouvrette, J. Bergalet, M.O. Duff, K.E. Garcia, C. Gelboin-Burkhart, M. Hochman, N.J. Lambert, H. Li, M.P. McGurk, T.B. Nguyen, T. Palden, I. Rabano, S. Sathe, R. Stanton, A. Su, R. Wang, B.A. Yee, B. Zhou, A.L. Louie, S. Aigner, X.D. Fu, E. Lecuyer, C.B. Burge, B.R. Graveley, G.W. Yeo, A large-scale binding and functional map of human RNA-binding proteins, Nature 583(7818) (2020) 711-719.
[6] H. Otsuka, A. Fukao, Y. Funakami, K.E. Duncan, T. Fujiwara, Emerging Evidence of Translational Control by AU-Rich Element-Binding Proteins, Front Genet 10 (2019) 332.
[7] P. Pryzbylkowski, O. Obajimi, J.C. Keen, Trichostatin A and 5 Aza-2' deoxycytidine decrease estrogen receptor mRNA stability in ER positive MCF7 cells through modulation of HuR, Breast Cancer Res Treat 111(1) (2008) 15-25.
[8] H. Wang, R. Li, Y. Hu, The alternative noncoding exons 1 of aromatase (Cyp19) gene modulate gene expression in a posttranscriptional manner, Endocrinology 150(7) (2009) 3301-7.
[9] S. Tan, K. Ding, Q.Y. Chong, J. Zhao, Y. Liu, Y. Shao, Y. Zhang, Q. Yu, Z. Xiong, W. Zhang, M. Zhang, G. Li, X. Li, X. Kong, A. Ahmad, Z. Wu, Q. Wu, X. Zhao, P.E. Lobie, T. Zhu, Post-transcriptional regulation of ERBB2 by miR26a/b and HuR confers resistance to tamoxifen in estrogen receptor-positive breast cancer cells, J Biol Chem 292(33) (2017) 13551-13564.
[10] X. Chen, S.A. Castro, Q. Liu, W. Hu, S. Zhang, Practical considerations on performing and analyzing CLIP-seq experiments to identify transcriptomic-wide RNA-protein interactions, Methods 155 (2019) 49-57.
[11] E.L. Van Nostrand, G.A. Pratt, A.A. Shishkin, C. Gelboin-Burkhart, M.Y. Fang, B. Sundararaman, S.M. Blue, T.B. Nguyen, C. Surka, K. Elkins, R. Stanton, F. Rigo, M. Guttman, G.W. Yeo, Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP), Nat Methods 13(6) (2016) 508-14.
[12] U. Christians, V. Schmitz, M. Haschke, Functional interactions between P-glycoprotein and CYP3A in drug metabolism, Expert Opin Drug Metab Toxicol 1(4) (2005) 641-54.
[13] X. Cai, G.M. Young, W. Xie, The xenobiotic receptors PXR and CAR in liver physiology, an update, Biochim Biophys Acta Mol Basis Dis 1867(6) (2021) 166101.

[14] J. Sonoda, W. Xie, J.M. Rosenfeld, J.L. Barwick, P.S. Guzelian, R.M. Evans, Regulation of a xenobiotic sulfonation cascade by nuclear pregnane X receptor (PXR), Proc Natl Acad Sci U S A 99(21) (2002) 13801-6.
[15] X. Chai, S. Zeng, W. Xie, Nuclear receptors PXR and CAR: implications for drug metabolism regulation, pharmacogenomics and beyond, Expert Opin Drug Metab Toxicol 9(3) (2013) 253-66.
[16] Y. Wei, C. Tang, V. Sant, S. Li, S.M. Poloyac, W. Xie, A Molecular Aspect in the Regulation of Drug Metabolism: Does PXR-Induced Enzyme Expression Always Lead to Functional Changes in Drug Metabolism?, Curr Pharmacol Rep 2(4) (2016) 187-192.
[17] A. Kodan, R. Futamata, Y. Kimura, N. Kioka, T. Nakatsu, H. Kato, K. Ueda, ABCB1/MDR1/P-gp employs an ATP-dependent twist-and-squeeze mechanism to export hydrophobic drugs, FEBS Lett 595(6) (2021) 707-716.
[18] K. Spruiell, D.Z. Jones, J.M. Cullen, E.M. Awumey, F.J. Gonzalez, M.A. Gyamfi, Role of human pregnane X receptor in high fat diet-induced obesity in pre-menopausal female mice, Biochem Pharmacol 89(3) (2014) 399-412.
[19] P.O. Oladimeji, T. Chen, PXR: More Than Just a Master Xenobiotic Receptor, Mol Pharmacol 93(2) (2018) 119-127.
[20] M. Karpale, A.J. Karajamaki, O. Kummu, H. Gylling, T. Hyotylainen, M. Oresic, A. Tolonen, H. Hautajarvi, M.J. Savolainen, M. Ala-Korpela, J. Hukkanen, J. Hakkola, Activation of nuclear receptor PXR induces atherogenic lipids and PCSK9 through SREBP2-mediated mechanism, Br J Pharmacol (2021).
[21] C. Zhou, M.M. Tabb, E.L. Nelson, F. Grun, S. Verma, A. Sadatrafiei, M. Lin, S. Mallick, B.M. Forman, K.E. Thummel, B. Blumberg, Mutual repression between steroid and xenobiotic receptor and NF-kappaB signaling pathways links xenobiotic metabolism and inflammation, J Clin Invest 116(8) (2006) 2280-2289.
[22] K. Huang, S. Mukherjee, V. DesMarais, J.M. Albanese, E. Rafti, A. Draghi Ii, L.A. Maher, K.M. Khanna, S. Mani, A.P. Matson, Targeting the PXR-TLR4 signaling pathway to reduce intestinal inflammation in an experimental model of necrotizing enterocolitis, Pediatr Res 83(5) (2018) 1031-1040.
[23] H. Huynh, Overexpression of tumour suppressor retinoblastoma 2 protein (pRb2/p130) in hepatocellular carcinoma, Carcinogenesis 25(8) (2004) 1485-94.
[24] D. Sun, H. Ren, M. Oertel, R.S. Sellers, D.A. Shafritz, L. Zhu, Inactivation of p27Kip1 promotes chemical mouse liver tumorigenesis in the resistant strain C57BL/6J, Mol Carcinog 47(1) (2008) 47-55.
[25] H. Wang, M. Venkatesh, H. Li, R. Goetz, S. Mukherjee, A. Biswas, L. Zhu, A. Kaubisch, L. Wang, J. Pullman, K. Whitney, M. Kuro-o, A.I. Roig, J.W. Shay, M. Mohammadi, S. Mani, Pregnane X receptor activation induces FGF19-dependent tumor aggressiveness in humans and mice, J Clin Invest 121(8) (2011) 3220-32.
[26] S. Kodama, Y. Yamazaki, M. Negishi, Pregnane X Receptor Represses HNF4alpha Gene to Induce Insulin-Like Growth Factor-Binding Protein IGFBP1 that Alters Morphology of and Migrates HepG2 Cells, Mol Pharmacol 88(4) (2015) 746-57.
[27] J. Zhou, M. Liu, Y. Zhai, W. Xie, The antiapoptotic role of pregnane X receptor in human colon cancer cells, Mol Endocrinol 22(4) (2008) 868-80.
[28] Z. Wang, B. Yang, M. Zhang, W. Guo, Z. Wu, Y. Wang, L. Jia, S. Li, N. Cancer Genome Atlas Research, W. Xie, D. Yang, lncRNA Epigenetic Landscape Analysis Identifies EPIC1 as an Oncogenic lncRNA that Interacts with MYC and Promotes Cell-Cycle Progression in Cancer, Cancer Cell 33(4) (2018) 706-720 e9.

[29] A. Shah, Y. Qian, S.M. Weyn-Vanhentenryck, C. Zhang, CLIP Tool Kit (CTK): a flexible and robust pipeline to analyze CLIP sequencing data, Bioinformatics 33(4) (2017) 566-567.
[30] I. Koutsounas, S. Theocharis, E. Patsouris, C. Giaginis, Pregnane X receptor (PXR) at the crossroads of human metabolism and disease, Curr Drug Metab 14(3) (2013) 341-50.
[31] G. Esposito, S. Gigli, L. Seguella, N. Nobile, A. D'Alessandro, M. Pesce, E. Capoccia, L. Steardo, C. Cirillo, R. Cuomo, G. Sarnelli, Rifaximin, a non-absorbable antibiotic, inhibits the release of pro-angiogenic mediators in colon cancer cells through a pregnane X receptor-dependent pathway, Int J Oncol 49(2) (2016) 639-45.
[32] T. Smutny, J. Dusek, L. Hyrsova, J. Nekvindova, A. Horvatova, S. Micuda, S. Gerbal-Chaloin, P. Pavek, The 3'-untranslated region contributes to the pregnane X receptor (PXR) expression down-regulation by PXR ligands and up-regulation by glucocorticoids, Acta Pharm Sin B 10(1) (2020) 136-152.
[33] N. Ripin, J. Boudet, M.M. Duszczyk, A. Hinniger, M. Faller, M. Krepl, A. Gadi, R.J. Schneider, J. Sponer, N.C. Meisner-Kober, F.H. Allain, Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM, Proc Natl Acad Sci U S A 116(8) (2019) 2935-2944.
[34] J. Yu, A. Navickas, H. Asgharian, B. Culbertson, L. Fish, K. Garcia, J.P. Olegario, M. Dermit, M. Dodel, B. Hanisch, Y. Luo, E.M. Weinberg, R. Dienstmann, R.S. Warren, F.K. Mardakheh, H. Goodarzi, RBMS1 Suppresses Colon Cancer Metastasis through Targeted Stabilization of Its mRNA Regulon, Cancer Discov 10(9) (2020) 1410-1423.
[35] C.M. Livi, P. Klus, R. Delli Ponti, G.G. Tartaglia, catRAPID signature: identification of ribonucleoproteins and RNA-binding regions, Bioinformatics 32(5) (2016) 773-5.