DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Transcriptional regulation of carboxylesterase 1 (*CES1*) in human liver: role of the nuclear receptor *NR1H3* (LXRα) and its splice isoforms

[1]Joseph M. Collins, [2]Rong Lu, [3]Xinwen Wang, [4]Hao-Jie Zhu and [1]Danxin Wang

[1]Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics, College of Pharmacy, University of Florida, Gainesville, Florida 32610, USA; [2]The Quantitative Sciences Unit, Division of Biomedical Informatics Research, Department of Medicine, Stanford University, Stanford, California; [3]Department of Pharmaceutical Sciences, Northeast Ohio Medical University, Rootstown, OH, 44272; [4]Department of Clinical Pharmacy, University of Michigan, Ann Arbor, MI, 48109-1065.

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Regulation of *CES1* expression by *NR1H3*

Corresponding author:

Danxin Wang, MD, PhD

Department of Pharmacotherapy and Translational Research, College of Pharmacy, University of Florida, PO Box 100486, 1345 Center Drive MSB PG-05B, Gainesville, FL 32610

Tel: 352-273-7673; Fax: 352-273-6121

Email: Danxin.wang@cop.ufl.edu

Number of text pages: 17

Number of tables: 0

Number of figures: 6 (main) + 5 (supplemental)

Number of references: 39

Number of words in abstract: 229

Number of words in introduction: 595

Number of words in discussion: 857

Abbreviations:

ACE: angiotensin-converting enzyme

AHR: aryl hydrocarbon receptor

CAR: constitutive androstane receptor

CES1: Carboxylesterase 1

ChIP-Seq: chromatin immunoprecipitation followed by sequencing

CRISPR: clustered regularly interspaced short palindromic repeats

gRNA: guide RNA

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

GTEx: Genotype-Tissue Expression

HNF4α: hepatocyte nuclear factor 4α

iPSCs: induced pluripotent stem cells

KD: knock-down

LXR: Liver X Receptor

NR1H3: Nuclear Receptor Subfamily 1 Group H Member 3

NRF2: nuclear factor erythroid related factor 2

PXR: pregnane X receptor

siRNA: small interfering RNA

SSI: Sobol's Sensitivity Indices

TA: transcriptional activation

TFs: Transcription Factors

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract:

Carboxylesterase 1 (*CES1*) is the predominant carboxylesterase in the human liver, involved in metabolism of both xenobiotics and endogenous substrates. Genetic or epigenetic factors that alter *CES1* activity or expression are associated with changes in drug response, lipid, and glucose homeostasis. However, the transcriptional regulation of *CES1* in the human liver remains uncertain. By applying both the random forest and Sobol's Sensitivity Indices (SSI) to analyze existing liver RNA expression microarray data (GSE9588), we identified *NR1H3* (LXRα) as a key factor regulating constitutive *CES1* expression. This model prediction was validated using siRNA knockdown and CRISPR-mediated transcriptional activation of *NR1H3* in Huh7 and HepG2 cells. We found that *NR1H3*'s activation of *CES1* is splice isoform-specific, namely that increased expression of the *NR1H3-211* isoform increased *CES1* expression while *NR1H3-201* did not. Also, in human liver samples, expression of *NR1H3*-211 and *CES1* are correlated, while *NR1H3-201* and *CES1* are not. This trend also occurs during differentiation of induced pluripotent stem cells (iPSCs) to hepatocytes, where only expression of the *NR1H3-211* isoform parallels expression of *CES1*. Moreover, we found that treatment with the *NR1H3* agonist T0901317 in HepG2 cells had no effect on *CES1* expression. Overall, our results demonstrate a key role of *NR1H3* in maintaining the constitutive expression of *CES1* in the human liver. Furthermore, our results support that the effect of *NR1H3* is splice isoform-specific and appears to be ligand independent.

Significance statement:

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Despite the central role of CES1 in metabolism of numerous medications, little is known about its transcriptional regulation. Here we identify NR1H3 as a key regulator of constitutive CES1 expression, and therefore is a potential target for future studies to understand inter-person variabilities in CES1 activity and drug metabolism.

Introduction:

Carboxylesterase 1 (CES1) is the predominant carboxylesterase in the human liver and intestine. CES1 catalyzes the ester cleavage of a large number of structurally diverse ester- or amide-containing substrates and is involved in the metabolism of both xenobiotics and endogenous compounds. CES activity is also a major determinant for the bioconversion of prodrugs to the active parent drugs (Imai and Hosokawa, 2010). Common drugs metabolized by CES1 include the anti-platelet prodrug clopidogrel (Lins et al., 1999), angiotensin-converting enzyme (ACE) inhibitors (imidapril, enalapril, trandolapril, ramipril) (Song and White, 2002; Thomsen et al., 2014), chemotherapeutic agents (irinotecan) (Humerickhouse et al., 2000), ADHD medications (methylphenidate) (Sun et al., 2004) and others. CES1 is also known to metabolize endogenous esters including cholesteryl esters, triacylglycerols, and other endogenous lipids that have vital physiological functions in lipid homeostasis (Lian et al., 2018). For example, reduced DNA methylation of the *CES1* gene is associated with childhood obesity (Li et al., 2018), and *CES1* knockout mice are more susceptible to high cholesterol diet-induced liver injury (Li et al., 2017).

There exists large inter-person variability in *CES1* expression and activity, which affects drug response. Nonsynonymous loss of function genetic polymorphisms in *CES1* have been associated with prodrug (e.g. dabigatran, etexilate, and oseltamivir)

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

activation, pharmacokinetics, and efficacy (Shi et al., 2016a; Shi et al., 2016b; Mu et al., 2020), and some variants have been proposed to serve as biomarkers for predicting clopidogrel efficacy (Lewis et al., 2013). However, the allele frequencies of these coding region variants are low, and therefore cannot explain the large variability in CES1 activity between individuals. Several potential *CES1* regulatory polymorphisms have also been identified (Geshi et al., 2005; Bruxel et al., 2013; Johnson et al., 2013), including structural variants arising from genomic translocation of the 5' region from the poorly expressed pseudogene *CES1P* to *CES1* (Sanford et al., 2016). However, the functional consequences of these regulatory variants are uncertain.

Little is known about transcriptional regulation of *CES1* in the human liver. Chemical induction experiments in mice showed that the transcription factors (TFs) aryl hydrocarbon receptor (AHR), constitutive androstane receptor (CAR, *NR1I3*), pregnane X receptor (PXR, *NR1I2*), and the nuclear factor erythroid related factor 2 (NRF2, *NFE2L2*) were involved in expression of the *CES* genes (Zhang et al., 2012). In patients with alcoholic steatohepatitis, the mRNA of both the hepatocyte nuclear factor 4α (*HNF4A*) and *CES1* were markedly reduced (Xu et al., 2016), implying a potential regulatory role of HNF4A on *CES1* expression during alcoholic steatohepatitis. In HepG2 cells, PXR is involved in insulin- (Yang et al., 2019) and fluoxetine-mediated (Shang et al., 2016) *CES1* transcriptional regulation and a variety of stimuli that alter signaling pathways have been shown to change *CES1* expression, including: the steroid hormone (17β-estradiol) (Wu et al., 2018), antioxidants (Chen et al., 2012), and disease states (e.g. type 2 diabetics) (Chen et al., 2015). However, the primary TFs controlling constitutive *CES1* expression remains largely unknown.

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

The purpose of this study was to identify TFs regulating constitutive *CES1* expression in the human liver. We applied both random forest and Sobol's Sensitivity Indices (SSI) (Lu et al., 2017) on existing microarray liver gene expression data (GSE9588) (Yang et al., 2010), as described previously for CYP3A4 (Wang et al., 2019). Of the 44 liver enriched TFs (Yang et al., 2010) analyzed, we identified *NR1H3* (LXRα) and several others as the top TFs associated with *CES1* expression. siRNA-mediated knock down (KD), CRISPR-mediated transcriptional activation (TA), and quantitative liver gene expression validated the regulatory role of *NR1H3* in constitutive *CES1* expression and demonstrate that this role of *NR1H3* is splice isoform specific.

Material and Methods:

Human liver samples. Human liver samples were obtained from the Cooperative Human Tissue Network (CHTN, Bethesda, MD). Demographics of liver samples are mean age-60 ± 13 years, female-52%, all samples were from Caucasian American donors (n=140). The University of Florida internal review board approved the human tissue study.

Random forest and SSI analysis of TF interactions with CES1. The mRNA dataset used is published microarray data (GSE9588) from 427 liver samples (Yang et al., 2010). We selected 44 liver-enriched TFs (Yang et al., 2010), represented by 78 probes in microarray data (some TFs were measured by multiple probes) (Supplemental Table 1). We estimated the mean decreases in Gini by fitting a random forest classifier of CES1 and estimated the main effect Sobol's indices by using the empirical variance of the best-fitting polynomial expression (Lu et al., 2017; Wang et al., 2019). The most influential TF was identified by the largest mean decreases in Gini, the

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

largest Sobol's indices, and the shortest distance between CES1 and TF in network analysis, which represents the strength of the interaction between CES1 and TF (the shorter the distance, the stronger the interaction).

Cell culture and iPSC differentiation. Huh7 and HepG2 cells were cultured at 37°C in a humidified incubator at 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and penicillin/streptomycin (100 U/100 µg). Human iPSCs (ASE-9203) were purchased from Applied StemCell (Milpitas, CA, USA) and were cultured at 37°C in a humidified incubator at 5% CO_2 in DEF-CS medium (Takara Bio, Mountain View, CA, USA). iPSC to hepatocyte differentiation was performed using the Cellartis iPSC to hepatocyte differentiation system (Takara Bio, Mountain View, CA, USA) according to manufacturer's instructions. The system progresses through directed differentiation of iPSCs into definitive endoderm (DE) (completed on day 14), which are then differentiated into hepatocytes (completed on day 32).

Gene knockdown using siRNA. Silencer siRNA targeting *NR1H3* (#138007), *NR1I3* (#5535), *HNF4A* (#290203) and *NR1I2* (#6638) and the negative control #1 were purchased from Thermo Fisher Scientific (Waltham, MA, USA). siRNA was introduced into cells using the lipofectamine RNAiMAX reagent (Thermo Fisher Scientific, Waltham, MA, USA). After incubation for 72 hours, the cells were harvested for total RNA preparation, reverse transcription, and quantitative real-time polymerase chain reaction (qRTPCR).

CRISPR-mediated NR1H3 transcriptional activation. By fusing VP64 (the universal transcriptional activator) with an inactive mutant Cas9 protein, dCas9, the dCas9-VP64

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

fusion protein can specifically activate transcription when directed by a guide RNA (gRNA) to a target gene promoter (Konermann et al., 2015). We used the lentiviral-based vectors lentiviral-VP64-dCas9 (#61429; Addgene) and LentisgRNA vector (#61427; Addgene), for VP64-dCas9 fusion protein and gRNA delivery. We used previously reported gRNA sequences to target *NR1H3* (Konermann et al., 2015). We targeted three separate promoters of *NR1H3* corresponding to three different isoforms: *NR1H3*-235 (NM_001251934), *NR1H3-211* (NM_001130101), and *NR1H3-201* (NM_001130102) (see Supplemental Figure 1 for liver *NR1H3* splice isoforms). Each promoter was targeted with three gRNAs, and a gRNA without a human genome target served as a negative control (Supplemental Table 2 contains the gRNA sequences). Lentiviral particles containing the expression vectors for VP64-dCas9 and a mix of the three gRNAs targeting a specific gene promoter or the negative control gRNA were incubated with Huh7 or HepG2 cells in the presence of 8 µg/ml SureEntry transduction reagent (Qiagan, Valencia, CA) for 24 hours. Cells were harvested 72 hours after transduction.

RNA preparation and gene expression analysis. Total RNA was prepared using RNA mini prep kits from Zymo Research (Irvine, CA). RNA was reverse transcribed into cDNA using the RTIV reverse transcriptase (Life Technologies, Carlsbad, CA). Gene expression levels were measured using qRTPCR with gene-specific primers (Supplemental Table 2) and the SYBR Green PCR master mix (Life Technologies, Carlsbad, CA), using GAPDH as an internal control as described (Collins and Wang, 2021). The Quantabio Q real-time PCR instrument (VWR, PA, USA) was used to measure the signal. The relative expression of each gene was calculated using the

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

following formula: expression level of tested gene = antilog2(mean Ct value of GAPDH – mean Ct value of tested gene)*10^6. After Log10 transformation, the expression level of *NR1H3* and *CES1* in liver samples followed a normal distribution.

CES1 protein quantification in human liver tissues. Relative CES1 protein expression in 46 individual human liver tissues was determined using a western blot assay that we described in a previous publication (Sanford et al., 2016).

Results

SSI and random forest analysis identified NR1H3 as a main regulator of CES1 expression. By applying SSI analysis to published microarray mRNA expression data (GSE9588) from 427 liver samples (Yang et al., 2010), we identified PGRMC1, NR1I3, NHF4A, NR1H3 and ARNT1 as the top five TFs with largest Sobol's indices for *CES1* expression among the 44 liver-enriched TFs tested (Supplemental Table 3). Similarly, these five TFs also showed the largest mean decreases in Gini using random forest classification (Supplemental Figure 1). We also employed SSI network analysis to help determine which TF may be directly regulating *CES1* expression in the liver. Compared to the other four TFs, NR1H3 has the shortest distance to CES1 (Figure 1), and therefore became the primary focus for experimental validation.

siRNA knockdown (KD) of NR1H3 and other TFs. To validate the *in-silico* predictions, we used siRNA KD in HepG2 and Huh7 cells to determine the effect of decreased NR1H3 on *CES1* expression. *NR1H3* siRNA reduced *NR1H3* expression 83% in Huh7 and 43% in HepG2 cells and significantly decreased *CES1* mRNA in both cell lines. *CES1* was decreased to a greater extent in HepG2 cells compared to Huh7 cells (Figure 2), which may be due to different expression of *CES1* in these two cells lines, as

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Huh7 has much lower *CES1* levels (313-fold) compared to HepG2 (Huh7: 8.6 ± 1.1 and HepG2: 2691 ± 712, arbitrary units resulting from comparison to an internal control GAPDH). We also tested four other TFs based on their SSI values (Supplemental Table 1) and network distance to CES1 (Figure 1). These included (as compared to NR1H3): NR1I3 (larger SSI value and similar distance), PGRMC1 and HNF4A (larger SSI value and more distal), and NR1I2 (smaller SSI value and closer distance). siRNA KD of all four of these TFs did not affect CES1 expression in Huh7 cells (Supplemental Figure 2), indicating that they do not directly control the constitutive expression of CES1. Instead, their associations from the models may have resulted from indirect regulation or inducible expression of CES1, consistent with previous studies showing involvement of NR1I3, NR1I2 and HNF4A in chemical- and lipid-mediated CES1 induction (Zhang et al., 2012; Shang et al., 2016; Xu et al., 2016; Yang et al., 2019).

CRISPR-mediated transcriptional activation (TA) of NR1H3 and the effects on CES1 expression. To further validate the impact of *NR1H3* on *CES1* expression, we used CRISPR-mediated TA to increase the expression of *NR1H3* in the same two cell lines, as reported (Wang et al., 2019). Based on Genotype-Tissue Expression (GTEx) data (Carithers et al., 2015), transcription of *NR1H3* can be initiated from at least three different promoters, producing three main splice isoforms, *NR1H3-211*, *NR1H3-201* and *NR1H3*-235, all of which are expressed in human livers (Supplemental Figure 3). We designed gRNA targeting all three promoters, using three gRNAs per promoter (see Supplemental Table 2 for gRNA sequences). Transcription of *NR1H3-201* and *NR1H3-211* was significantly increased (1.4-2-fold) by CRISPR-mediated TA in both Huh7 and HepG2 cells, while *NR1H3*-235 was not (Figure 3). *NR1H3*-235 is expressed at a low

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

level in the liver (Supplemental Figure 3), implying that additional regulatory mechanisms are controlling its expression. Increased expression of *NR1H3-211* enhanced CES1 mRNA levels in both Huh7 and HepG2 cells, in agreement with the KD results. In contrast, enhanced expression of *NR1H3-201* failed to increase CES1 in either cell line (Figure 3). Overall, these results agree with the SSI prediction indicating NR1H3 as a key TF controlling constitutive *CES1* expression. The results also indicate that the regulation of *CES1* by NR1H3 is splice isoform-specific.

Changes in NR1H3 splice isoforms and CES1 expression during iPSC to hepatocyte differentiation. We next leveraged an iPSC to hepatocyte differentiation model to determine changes in the expression of *CES1*, *NR1H3*, and its splice isoforms during development. We measured developmental markers during the different cell stages to confirm proper differentiation: *POU5F1* (Oct4) for the iPSC stage (day 0), CER1 for the definitive endoderm (DE) stage (day 14), and CYP3A4 for hepatocyte-like cells (day 32) (Ghosheh et al., 2017) . The markers followed an expression pattern in agreement with previous report (Ghosheh et al., 2017): Oct4 expression was high in iPSCs and declined throughout differentiation, CER1 peaked at day14, and CYP3A4 progressed from being undetectable in the iPS and DE cells to a marked increase in expression in the hepatocytes at day 32 (Supplemental Figure 4). These results indicated that our cell differentiation was successful.

Expression of CES1 increased with differentiation of the iPSCs: its relative expression was low in iPSCs, increased roughly two-fold at day 14 in the DE cells, and drastically increased (15-fold, compared to iPSCs) at day 32 in the hepatocyte cells (Figure 4a). *NR1H3* followed a very similar pattern to *CES1*, while a different type of

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

LXR, *NR1H2* (LXRβ), did not (Figure 4a). Furthermore, only the expression of splice isoform *NR1H3-211* peaked at day 32, while the other isoforms, *NR1H3-201* and *NR1H3-235* peaked at day 14, and then declined by day 32 (Figure 4b). Thus, at day 32 in the hepatocyte cells, *NR1H3-211* is the primary isoform (Figure 4c) and thereby coincides with the highest level of *CES1* expression (Figure 4a). These co-expression results support that *NR1H3* (particularly *NR1H3-211*) may have a regulatory role in controlling *CES1* expression during the transition from iPSCs to hepatocytes, and thereby contribute to *CES1* expression in the human liver.

Correlation between expression of CES1 and NR1H3 splice isoforms in cell lines and liver samples.

We compared the expression levels of *CES1* to the overall expression of *NR1H3* and its three splice isoforms in Huh7 cells, HepG2 cells, and liver samples. *CES1* expression varies in the different cell lines; compared to its lowest expression in Huh7 cells, *CES1* is 313-fold higher in HepG2 cells and 4308-fold higher in the liver (average of 140 liver samples). In contrast, when considering total NR1H3 mRNA levels, *NR1H3* expression is relatively similar across all three cell types and is only 1.2-fold higher in HepG2 cells and 3.8-fold higher in liver tissues (compared to Huh7 cells). However, analysis of the individual *NR1H3* splice isoforms shows large differences between the liver and the two cell lines. Over 90% of the total *NR1H3* transcripts in liver tissues are the *NR1H3-211* isoform, while *NR1H3-211* only comprises ~50% of the total *NR1H3* transcript pool in the Huh7 and HepG2 cells (Supplemental Figure 5). Furthermore, in 140 liver samples, we observed a strong positive correlation between levels of CES1 and NR1H3-211 ($P<0.0001$) (Figure 5a), while there was no correlation between

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

expression of CES1 and NR1H3-201 (P=0.367) (Figure 5b). We also measured CES1 protein levels in 46 samples using a western blot approach. Similar to the mRNA expression results, CES1 protein levels are positively correlated with NR1H3-211 (P<0.0001) (Figure 5c) but not with NR1H3-201 (r=-0.115, P=0,451) (Figure 5d). These results indicate that NR1H3-211 is the predominate splice isoform regulating expression of *CES1.*

NR1H3 agonists do not activate CES1 expression. NR1H3 encodes LXRα, a transcriptional regulator that has previously been shown to be strongly activated by T0901317 (Hoang et al., 2012). We therefore tested whether this agonist would also cause a concomitant increased in *CES1* expression. For positive controls, we also tested expression of two genes ABCG1 and FAS that are known to be induced by T0901317 (Hoang et al., 2012). While T0901317 treatment (0.3 or 1 μM, 24h) drastically increased the expression of ABCG1 (~100-fold) and FAS (~4-fold) (Figure 6b), it did not alter *CES1* expression (Figure 6a). Similar results were observed in Huh7 cells (data not shown).

Discussion:

We have identified NR1H3 as a key regulator for constitutive *CES1* expression in the human liver using SSI and random forest analyses. siRNA mediated KD or CRISPR-mediated gene TA of *NR1H3* in HepG2 and Huh7 cells also caused a corresponding change in *CES1* expression. To our knowledge, this is the first study showing the regulation of *CES1* by NR1H3. Moreover, our results demonstrate that the NR1H3-211 splicing isoform is the key NR1H3 splice isoform controlling constitutive *CES1* expression.

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

NR1H3 (LXRα) is a ligand-activated TF of the nuclear receptor superfamily, playing important roles in lipid and carbohydrate metabolism (Baranowski, 2008). The role of ligand-activated NR1H3 in gene expression regulation is well studied; for example, *NR1H3* agonists are known to increase the expression of many genes related to lipid and glucose homeostasis, and display potent antiatherogenic and antidiabetic effects (Baranowski, 2008). *NR1H3* agonists also induce the expression of several phase I and phase II drug metabolizing enzymes, for example, CYP1A1 (Shibahara et al., 2011), CYP3A4, CYP2B6 (Duniec-Dmuchowski et al., 2007) and UGT1s (Hansmann et al., 2020), in cells and mouse models. However, a role of unliganded *NR1H3* has yet to be reported. Our results, for the first time, demonstrate the critical role of *NR1H3* on maintaining basal *CES1* expression and showed correlation between the expression of *NR1H3* and *CES1* in human liver. No *NR1H3* agonists were added during our siRNA or CRISPR-mediated transcription activation experiments, where we saw corresponding changes in expression of both NR1H3 and CES1 (Figures 2 & 3), supporting a role of unliganded *NR1H3* in controlling *CES1* expression in hepatic cells. Furthermore, activation of *NR1H3* by agonist T0901317 did not induce the expression of *CES1* in HepG2 and Huh7 cells, suggesting potential different signaling pathways of ligand-free and ligand-bound *NR1H3*. This result is consistent with our recent findings regarding a different nuclear receptor, the estrogen receptor α (ESR1). Although canonically ESR1 is considered a ligand-activated nuclear receptor, we demonstrated the different roles of ligand-free and ligand-bound ESR1 in regulating the expression of cytochrome P450s (Wang et al., 2019) and our ChIP-Seq experiments showed distinct binding motifs and binding sites for these two forms of ESR1 (Collins et al., 2021).

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

These results suggest that having different chromatin binding and signaling pathways in the presence or absence of ligands may be a general phenomenon of the nuclear receptors. However, we cannot rule out that endogenous NR1H3 ligands may have different effects on NR1H3-mediated regulation than synthetic ligands, and thus, the contribution of endogenous NR1H3 ligands on regulation of CES1 remains unclear and will require further investigation.

NR1H3 has numerous splice isoforms exist, with 35 transcripts listed in the Ensembl database (Howe et al., 2021). According to the GTEx portal (Carithers et al., 2015), thirteen of these transcripts are expressed in the liver, with five of them (*NR1H3-211, -201, -217, -235* and *-221*) being predicted as protein coding (Supplemental Figure 3). These five transcripts are initiated from three different promoters, and we chose a major isoform from each promoter for this study (*NR1H3-211*, *NR1H3-201* and *NR1H3-*235, Supplemental Figure 3). *NR1H3-221* is initiated from the same promoter as *NR1H3-201* but is not (or nearly not) expressed in the liver, while *NR1H3-217* shares the same promoter with *NR1H3-211* and has low liver expression (Supplemental Figure 3). *NR1H3-211* and *NR1H3-217* encode the same protein but differ at the 5'UTR due to retention of an intron in *NR1H3-217*. The qPCR primers used in this study cannot differentiate NR1H3-211 from -217, and thus, may represent the sum of these two isoforms. The expression level of *NR1H3-*235 is low in hepatic cells and in the liver (<1% of total) and therefore may not play major regulatory role. Conversely, the expression levels of the other two isoforms, *NR1H3-211* and *NR1H3-201*, are dynamic depending on the cell type (Figures 3 & 4). Although the expression of both *NR1H3-211* and *NR1H3-201* were activated by CRISPR-mediated TA in HepG2 and Huh7 cells,

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

only *NR1H3-211* enhanced *CES1* expression, indicating different regulatory roles of these two splice isoforms. In further support of this, expression of *CES1* only paralleled the *NR1H3-211* isoform during iPSC to hepatocyte differentiation, and only *NR1H3-211* is correlated with *CES1* expression in human liver samples. These results indicate that the regulation of *NR1H3* on *CES1* transcription is mediated by *NR1H3-211* but not the *NR1H3-201* isoform. These two isoforms differ at their 5'UTR and have different translation start sites (Supplemental Figure 3). Compared to *NR1H3-211*, the *NR1H3-201* protein is shorter and lacks 45 amino acids at the N-terminal. A previous study showed that the N-terminal truncated *NR1H3-201* isoform has lower basal and agonist-induced transcriptional activity than the full-length isoform, indicating that the N-terminal 50 amino acids are critical for full NR1H3 transcriptional function (Chen et al., 2005). Indeed, further studies are needed to elucidate the mechanisms underlying NR1H3 regulation of basal CES1 expression in human liver.

In conclusion, our results demonstrate the regulation of *CES1* by the nuclear receptor *NR1H3* in a ligand-independent and splice isoform-specific manner. Therefore, genetic, or epigenetic factors affecting the expression of *NR1H3* will have the potential to alter *CES1* expression, opening new research directions for understanding variable expression of *CES1* in the human liver.

Authorship Contributions:

Participated in Research design: D. Wang and Collins

Conducted experiments: D. Wang, Collins, X. Wang, and Zhu

Performed data analysis: Lu, Collins and D. Wang

Wrote or contributed to the writing of the manuscript: D. Wang, Collins, Lu, X. Wang and Zhu

References:

Baranowski M (2008) Biological role of liver X receptors. *J Physiol Pharmacol* **59 Suppl 7:**31-55.

Bruxel EM, Salatino-Oliveira A, Genro JP, Zeni CP, Polanczyk GV, Chazan R, Rohde LA, and Hutz MH (2013) Association of a carboxylesterase 1 polymorphism with appetite reduction in children and adolescents with attention-deficit/hyperactivity disorder treated with methylphenidate. *Pharmacogenomics J* **13:**476-480.

Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, Compton CC, DeLuca DS, Peter-Demchok J, Gelfand ET, Guan P, Korzeniewski GE, Lockhart NC, Rabiner CA, Rao AK, Robinson KL, Roche NV, Sawyer SJ, Segrè AV, Shive CE, Smith AM, Sobin LH, Undale AH, Valentino KM, Vaught J, Young TR, Moore HM, and Consortium G (2015) A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank* **13:**311-319.

Chen M, Beaven S, and Tontonoz P (2005) Identification and characterization of two alternatively spliced transcript variants of human liver X receptor alpha. *J Lipid Res* **46:**2570-2579.

Chen R, Wang Y, Ning R, Hu J, Liu W, Xiong J, Wu L, Liu J, Hu G, and Yang J (2015) Decreased carboxylesterases expression and hydrolytic activity in type 2 diabetic mice through Akt/mTOR/HIF-1α/Stra13 pathway. *Xenobiotica* **45:**782-793.

Chen YT, Shi D, Yang D, and Yan B (2012) Antioxidant sulforaphane and sensitizer trinitrobenzene sulfonate induce carboxylesterase-1 through a novel element

transactivated by nuclear factor-E2 related factor-2. *Biochem Pharmacol* **84:**864-871.

Collins JM, Huo Z, and Wang D (2021) ESR1 ChIP-Seq Identifies Distinct Ligand-Free ESR1 Genomic Binding Sites in Human Hepatocytes and Liver Tissue. *Int J Mol Sci* **22**.

Collins JM and Wang D (2021) Co-expression of drug metabolizing cytochrome P450 enzymes and estrogen receptor alpha (ESR1) in human liver: racial differences and the regulatory role of ESR1. *Drug Metab Pers Ther*.

Duniec-Dmuchowski Z, Ellis E, Strom SC, and Kocarek TA (2007) Regulation of CYP3A4 and CYP2B6 expression by liver X receptor agonists. *Biochem Pharmacol* **74:**1535-1540.

Geshi E, Kimura T, Yoshimura M, Suzuki H, Koba S, Sakai T, Saito T, Koga A, Muramatsu M, and Katagiri T (2005) A single nucleotide polymorphism in the carboxylesterase gene is associated with the responsiveness to imidapril medication and the promoter activity. *Hypertens Res* **28:**719-725.

Ghosheh N, Küppers-Munther B, Asplund A, Edsbagge J, Ulfenborg B, Andersson TB, Björquist P, Andersson CX, Carén H, Simonsson S, Sartipy P, and Synnergren J (2017) Comparative transcriptomics of hepatic differentiation of human pluripotent stem cells and adult human liver tissue. *Physiol Genomics* **49:**430-446.

Hansmann E, Mennillo E, Yoda E, Verreault M, Barbier O, Chen S, and Tukey RH (2020) Differential Role of Liver X Receptor (LXR). *Drug Metab Dispos* **48:**255-263.

Hoang MH, Jia Y, Jun HJ, Lee JH, Lee BY, and Lee SJ (2012) Fucosterol is a selective liver X receptor modulator that regulates the expression of key genes in cholesterol homeostasis in macrophages, hepatocytes, and intestinal cells. *J Agric Food Chem* **60:**11567-11575.

Howe KL, Achuthan P, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, El Houdaigui B, Fatima R, Gall A, Garcia Giron C, Grego T, Guijarro-Clarke C, Haggerty L, Hemrom A, Hourlier T, Izuogu OG, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Gonzalez Martinez J, Marugán JC, Maurel T, McMahon AC, Mohanan S, Moore B, Muffato M, Oheh DN, Paraschas D, Parker A, Parton A, Prosovetskaia I, Sakthivel MP, Salam AIA, Schmitt BM, Schuilenburg H, Sheppard D, Steed E, Szpak M, Szuba M, Taylor K, Thormann A, Threadgold G, Walts B, Winterbottom A, Chakiachvili M, Chaubal A, De Silva N, Flint B, Frankish A, Hunt SE, IIsley GR, Langridge N, Loveland JE, Martin FJ, Mudge JM, Morales J, Perry E, Ruffier M, Tate J, Thybert D, Trevanion SJ, Cunningham F, Yates AD, Zerbino DR, and Flicek P (2021) Ensembl 2021. *Nucleic Acids Res* **49:**D884-D891.

Humerickhouse R, Lohrbach K, Li L, Bosron WF, and Dolan ME (2000) Characterization of CPT-11 hydrolysis by human liver carboxylesterase isoforms hCE-1 and hCE-2. *Cancer Res* **60:**1189-1192.

Imai T and Hosokawa M (2010) Prodrug approach using carboxylesterases activity: catalytic properties

and gene regulation of carboxylesterase in mammalian tissue. *Journal of Pesticide Science* **35:**229-239.

Johnson KA, Barry E, Lambert D, Fitzgerald M, McNicholas F, Kirley A, Gill M, Bellgrove MA, and Hawi Z (2013) Methylphenidate side effect profile is influenced by genetic variation in the attention-deficit/hyperactivity disorder-associated CES1 gene. *J Child Adolesc Psychopharmacol* **23:**655-664.

Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, and Zhang F (2015) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517:**583-588.

Lewis JP, Horenstein RB, Ryan K, O'Connell JR, Gibson Q, Mitchell BD, Tanner K, Chai S, Bliden KP, Tantry US, Peer CJ, Figg WD, Spencer SD, Pacanowski MA, Gurbel PA, and Shuldiner AR (2013) The functional G143E variant of carboxylesterase 1 is associated with increased clopidogrel active metabolite levels and greater clopidogrel response. *Pharmacogenet Genomics* **23:**1-8.

Li J, Wang Y, Matye DJ, Chavan H, Krishnamurthy P, Li F, and Li T (2017) Sortilin 1 Modulates Hepatic Cholesterol Lipotoxicity in Mice via Functional Interaction with Liver Carboxylesterase 1. *J Biol Chem* **292:**146-160.

Li Y, Zhou Y, Zhu L, Liu G, Wang X, Wang J, You L, Ji C, Guo X, Zhao Y, and Cui X (2018) Genome-wide analysis reveals that altered methylation in specific CpG loci is associated with childhood obesity. *J Cell Biochem* **119:**7490-7497.

Lian J, Nelson R, and Lehner R (2018) Carboxylesterases in lipid metabolism: from mouse to human. *Protein Cell* **9:**178-195.

Lins R, Broekhuysen J, Necciari J, and Deroubaix X (1999) Pharmacokinetic profile of 14C-labeled clopidogrel. *Semin Thromb Hemost* **25 Suppl 2:**29-33.

Lu R, Wang D, Wang M, and Rempala GA (2017) Estimation of Sobol's sensitivity indices under generalized linear models. *Communications in Statistics - Theory and Methods* **47:**5163-5195.

Mu G, Xiang Q, Zhang Z, Liu C, Zhang H, Liu Z, Pang X, Jiang J, Xie Q, Zhou S, Wang Z, Hu K, Jiang S, Qin X, and Cui Y (2020) PNPT1 and PCGF3 variants associated with angiotensin converting enzyme inhibitor-induced cough: a nested case-control genome wide study. *Pharmacogenomics* **21:**601-614.

Sanford JC, Wang X, Shi J, Barrie ES, Wang D, Zhu HJ, and Sadee W (2016) Regulatory effects of genomic translocations at the human carboxylesterase-1 (CES1) gene locus. *Pharmacogenet Genomics* **26:**197-207.

Shang W, Liu J, Chen R, Ning R, Xiong J, Liu W, Mao Z, Hu G, and Yang J (2016) Fluoxetine reduces CES1, CES2, and CYP3A4 expression through decreasing PXR and increasing DEC1 in HepG2 cells. *Xenobiotica* **46:**393-405.

Shi J, Wang X, Eyler RF, Liang Y, Liu L, Mueller BA, and Zhu HJ (2016a) Association of Oseltamivir Activation with Gender and Carboxylesterase 1 Genetic Polymorphisms. *Basic Clin Pharmacol Toxicol* **119:**555-561.

Shi J, Wang X, Nguyen JH, Bleske BE, Liang Y, Liu L, and Zhu HJ (2016b) Dabigatran etexilate activation is affected by the CES1 genetic polymorphism G143E (rs71647871) and gender. *Biochem Pharmacol* **119:**76-84.

Shibahara N, Masunaga Y, Iwano S, Yamazaki H, Kiyotani K, and Kamataki T (2011) Human cytochrome P450 1A1 is a novel target gene of liver X receptor α. *Drug Metab Pharmacokinet* **26:**451-457.

Song JC and White CM (2002) Clinical pharmacokinetics and selective pharmacodynamics of new angiotensin converting enzyme inhibitors: an update. *Clin Pharmacokinet* **41:**207-224.

Sun Z, Murry DJ, Sanghani SP, Davis WI, Kedishvili NY, Zou Q, Hurley TD, and Bosron WF (2004) Methylphenidate is stereoselectively hydrolyzed by human carboxylesterase CES1A1. *J Pharmacol Exp Ther* **310:**469-476.

Thomsen R, Rasmussen HB, Linnet K, and Consortium I (2014) In vitro drug metabolism by human carboxylesterase 1: focus on angiotensin-converting enzyme inhibitors. *Drug Metab Dispos* **42:**126-133.

Wang D, Lu R, Rempala G, and Sadee W (2019) Ligand-Free Estrogen Receptor alpha (ESR1) as master regulator for the expression of CYP3A4 and other cytochrome P450 enzymes in the human liver. *Mol Pharmacol* **96:**430-440.

Wu L, Hafiz MZ, Guan Y, He S, Xiong J, Liu W, Yan B, Li X, and Yang J (2018) 17β-estradiol suppresses carboxylesterases by activating c-Jun/AP-1 pathway in primary human and mouse hepatocytes. *Eur J Pharmacol* **819:**98-107.

Xu J, Xu Y, Li Y, Jadhav K, You M, Yin L, and Zhang Y (2016) Carboxylesterase 1 Is Regulated by Hepatocyte Nuclear Factor 4α and Protects Against Alcohol- and MCD diet-induced Liver Injury. *Sci Rep* **6:**24277.

Yang X, Zhang B, Molony C, Chudin E, Hao K, Zhu J, Gaedigk A, Suver C, Zhong H, Leeder JS, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, Ulrich RG,

Slatter JG, Schadt EE, Kasarskis A, and Lum PY (2010) Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. *Genome Res* **20:**1020-1036.

Yang X, Zhang X, Liu Y, Xi T, and Xiong J (2019) Insulin transcriptionally down-regulates carboxylesterases through pregnane X receptor in an Akt-dependent manner. *Toxicology* **422:**60-68.

Zhang Y, Cheng X, Aleksunes L, and Klaassen CD (2012) Transcription factor-mediated regulation of carboxylesterase enzymes in livers of mice. *Drug Metab Dispos* **40:**1191-1197.

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Footnotes

Financial support: This work was supported by National Institute of Health National Heart Lung and Blood Institute [R01 HL126969] and National Institute of General Medical Science [R01 GM120396, R35 GM140845]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors declare no conflicts of interest.

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

Figure legends

Figure 1. Transcription factors identified using SSI analysis and their predicted interactions affecting CES1 expression. Dot sizes represent the effect of each TF on CES1 expression: the larger the dot, the greater the predicted regulatory effect of that TF on CES1 expression. The connecting lines illustrate predicted interactions occurring between the TFs and/or CES1, with the length of the line indicating the overall impact of each interaction (the shorter the distance, the higher the SSI value of the interaction). TFs measured by more than one microarray probe (for example, NR1I3 and NR1I3.1; HNF4A.2 and HNF4A.4) yielded similar results.

Figure 2. *NR1H3* knockdown by siRNA and the effect on expression of *CES1* in a) Huh7 and b) HepG2 cells. Mean ± SD, n=4. Compared to negative control (NC), * $p<0.05$; ** $p<0.01$; *** $p<0.001$, ANOVA with Bonferroni: compared selected pairs post hoc test.

Figure 3. Effect of *NR1H3* transcriptional activation (TA) on the expression of *CES1* in a) Huh7 and b) HepG2 cells. Mean ± SD, n=4. Compared to negative control (NC), *$p<0.05$; ** $p<0.01$; *** $p<0.001$, ANOVA with Bonferroni: compared selected pairs post hoc test. Note: The measured NR1H3-211 level is the sum of NR1H3-211 and NR1H3-217 due to the lack of qPCR primer specificity.

Figure 4. Gene expression changes during iPSC to hepatocyte differentiation. (a & b) Expression levels of genes were measured at day 0, day 14, and day 32 using qRTPCR with GAPDH as an internal control. (c) Relative expression level of each *NR1H3* splicing isoform at the different stages of differentiation. Data expressed as % of the

DMD Fast Forward. Published on October 25, 2021 as DOI: 10.1124/dmd.121.000649
This article has not been copyedited and formatted. The final version may differ from this version.

total *NR1H3* level. Note: NR1H3-235 expression is too low (<1%) to be visible in the graph.

Figure 5. Correlation between the levels of mRNA (a & b) and protein (c and d) of *CES1* and two *NR1H3* isoforms. (a & c) *NR1H3-211* and (b & d) *NR1H3-201*.

Figure 6. Effect of the *NR1H3* agonist T0901317 on gene expression in HepG2 cells. (a) *CES1* (24hrs & 48hrs) (b) ABCG1 and FAS (24hrs). Panel b is in log10 scale. Mean ± SD, n=4. *** compared to DMSO, p<0.0001, ANOVA with Bonferroni: compared selected pairs post hoc test.

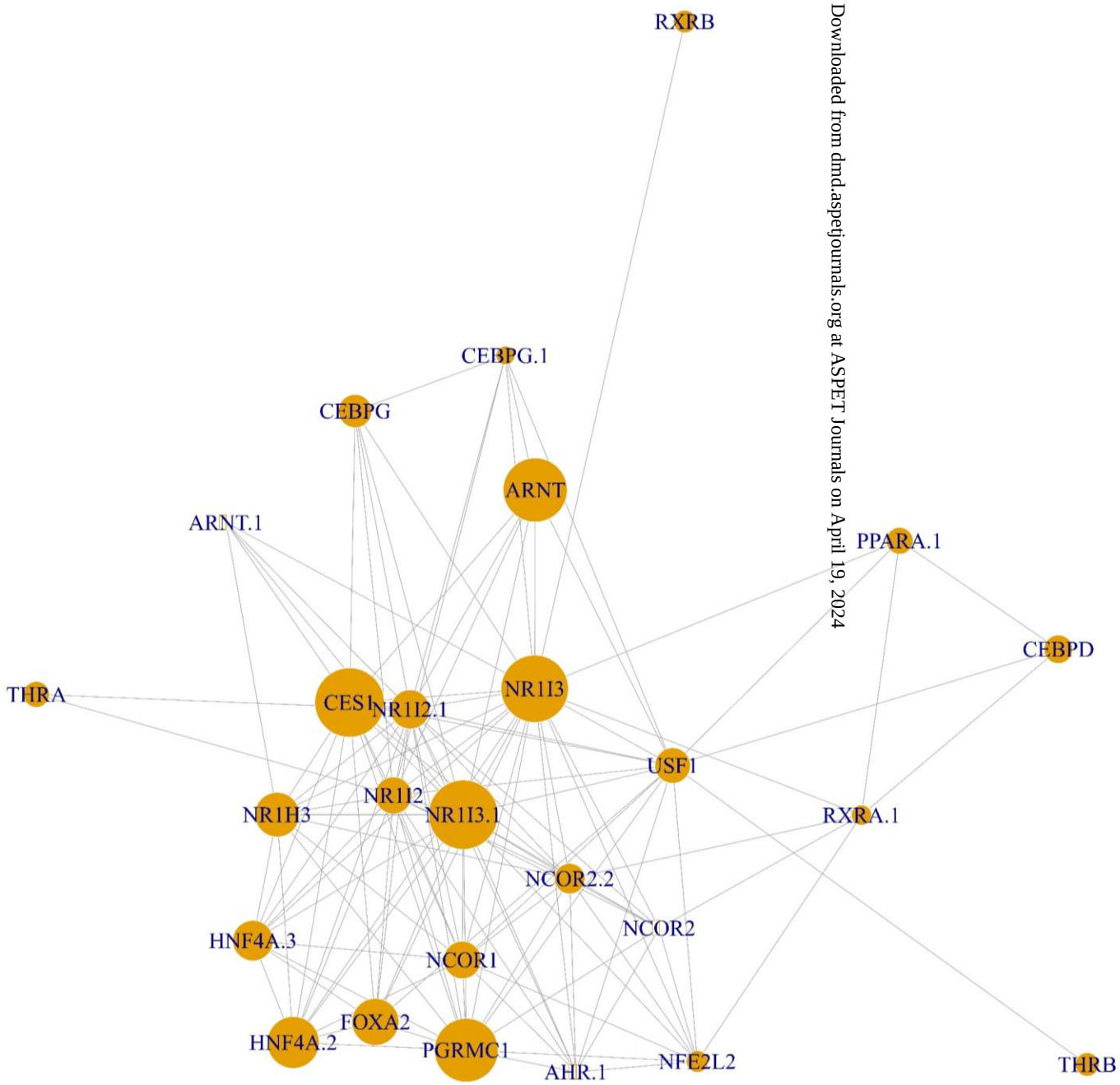


Figure 1.

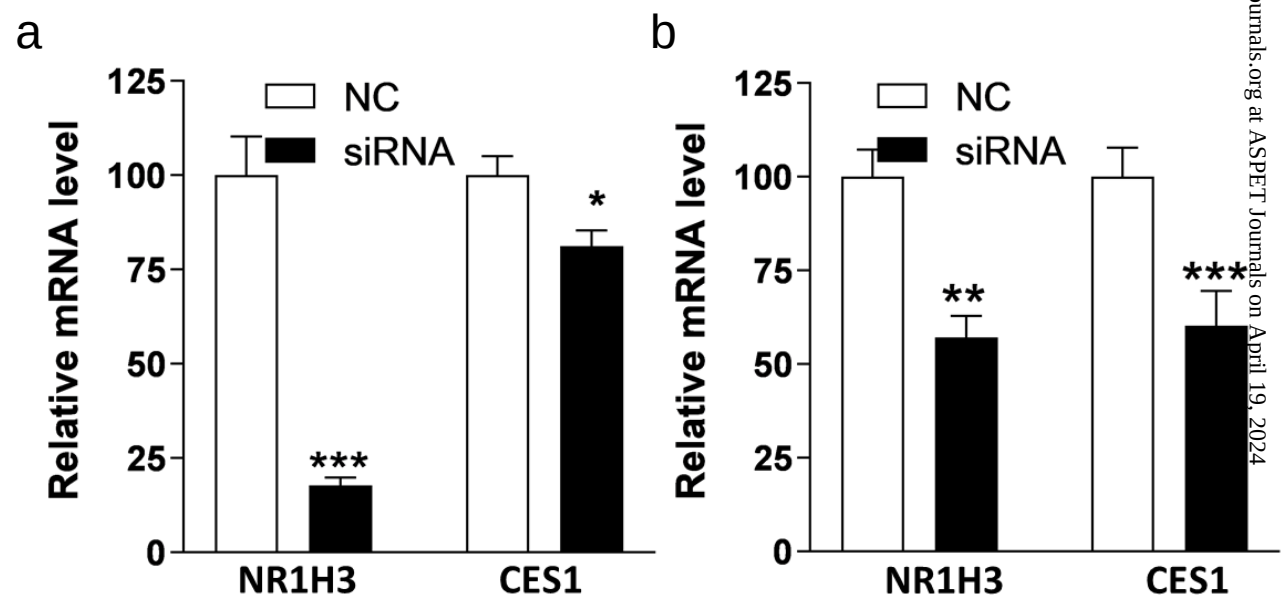


Figure 2.

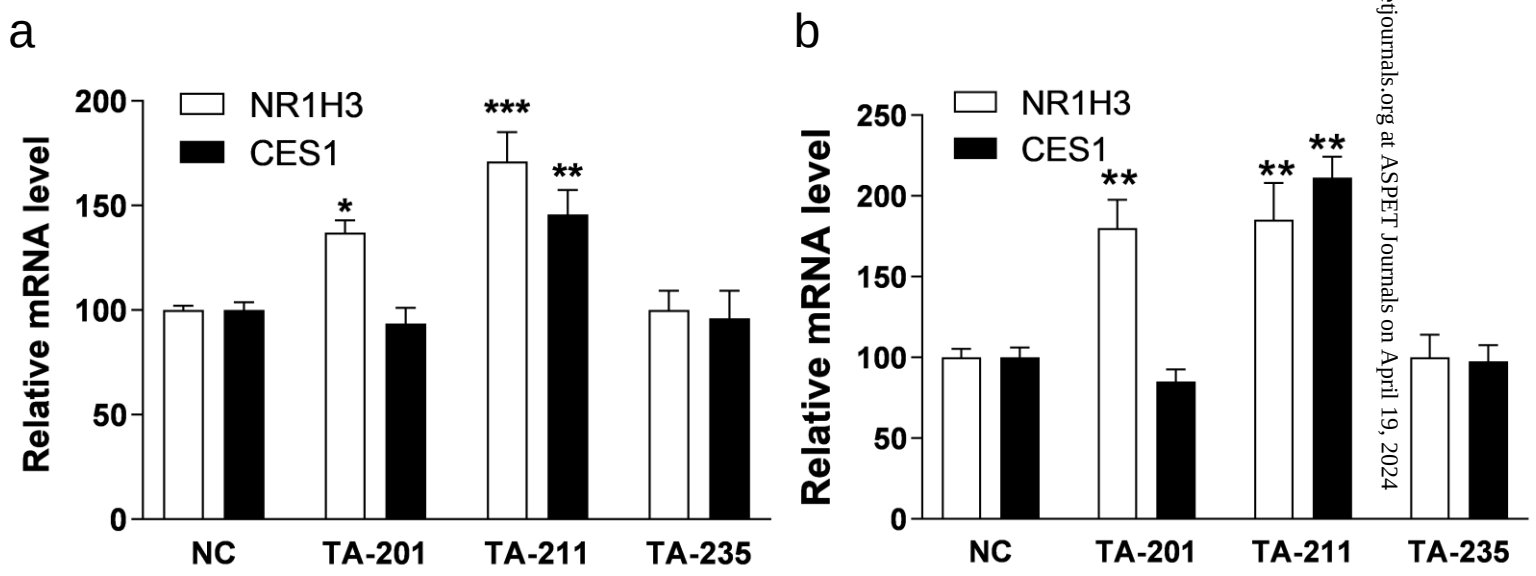


Figure 3.

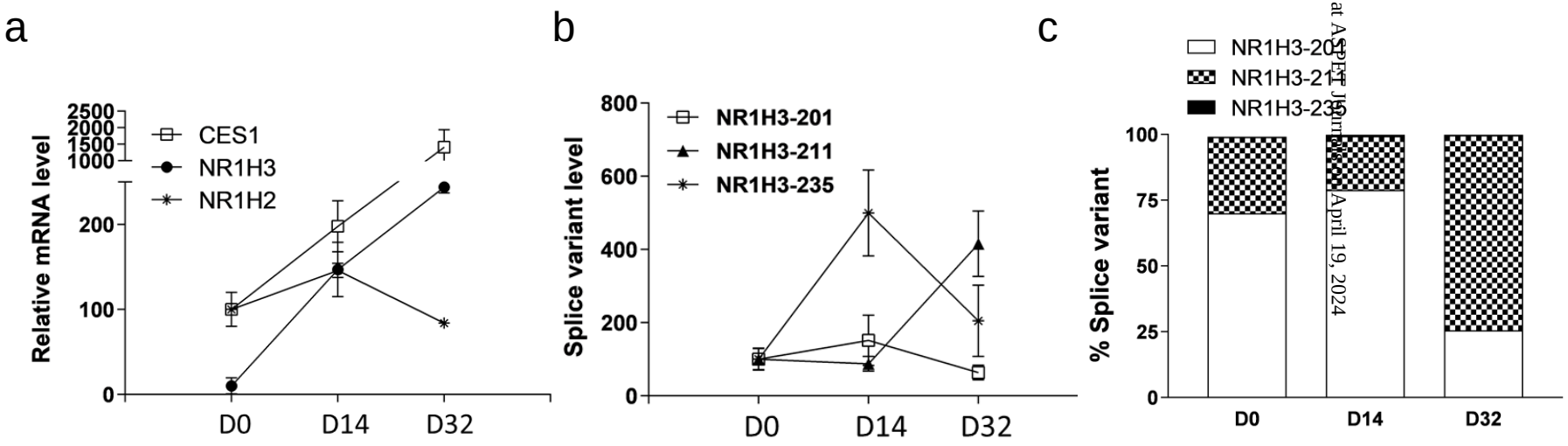


Figure 4

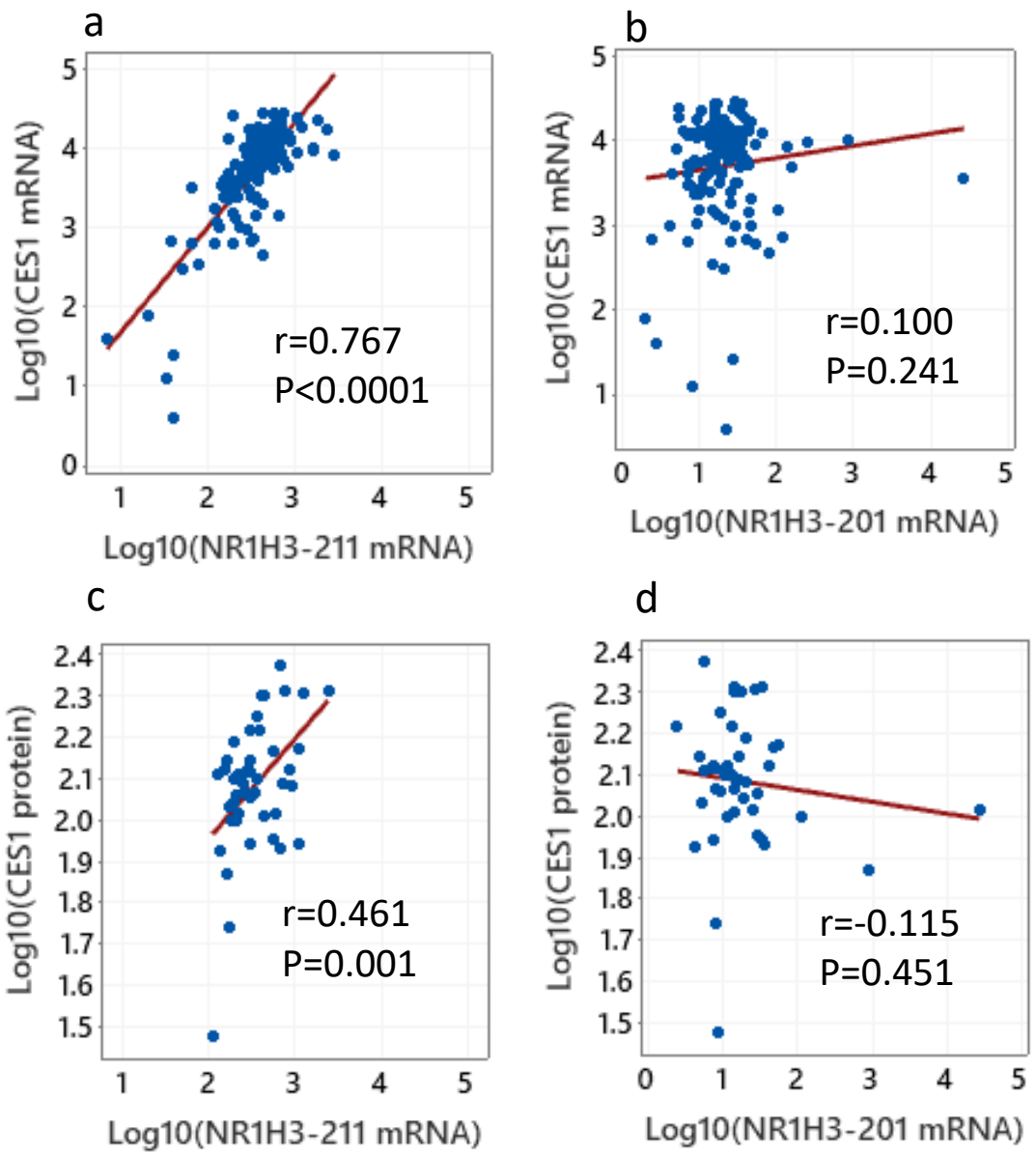


Figure 5

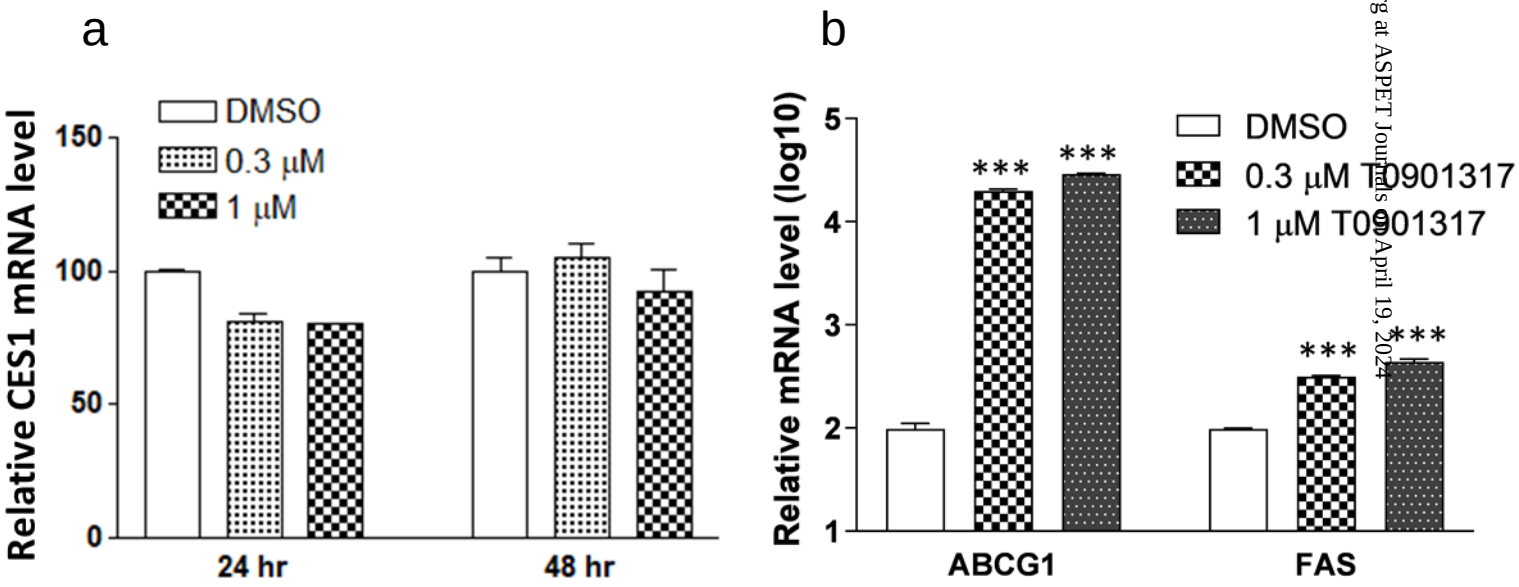


Figure 6.

Supplemental Table 1. Liver enriched transcription factors and their probe IDs

TFs	probe ID	Term used in analysis
AHR	10023817519	AHR
AHR	10033668916	AHR.1
AHRR	10025905866	AHR.2
ARNT	10025907990	ARNT
ARNT	10033668940	ARNT.1
CEBPA	10025909440	CEBPA
CEBPB	10025909776	CEBPB
CEBPD	10023822610	CEBPD
CEBPG	10023808900	CEBPG
CEBPG	10023828264	CEBPG.1
DBP	10025908201	DBP
ESR1	10023819651	ESR1
ESR1	10025920920	ESR1.1
ESR1	10033668534	ESR1.2
FOXA1	10025904606	FOXA1
FOXA1	10025929951	FOXA1.1
FOXA2	10025911776	FOXA2
FOXA3	10023807919	FOXA3
HNF4A	10025905066	HNF4A
HNF4A	10025906346	HNF4A.1
HNF4A	10025910259	HNF4A.2
HNF4A	10033668798	HNF4A.3
HNF4G	10023806639	HNF4G
HNF4G	10023808807	HNF4G.1
NCOA1	10025911463	NCOA1
NCOA2	10025907334	NCOA2
NCOA3	10023805195	NCOA3
NCOA3	10025902252	NCOA3.1
NCOR1	10025909797	NCOR1
NCOR2	10023805826	NCOR2
NCOR2	10025917191	NCOR2.1
NCOR2	10033668825	NCOR2.2
NFE2L2	10023818257	NFE2L2
NR0B1	10025902984	NR0B1
NR0B2	10023849951	NR0B2
NR1D2	10025907049	NR1D2
NR1D2	10025928478	NR1D2.1
NR1H2	10025910072	NR1H2
NR1H2	10033668639	NR1H2.1
NR1H3	10023812516	NR1H3
NR1H4	10023822564	NR1H4
NR1I2	10025905041	NR1I2
NR1I2	10033668879	NR1I2.1

NR1I3	10023822733	NR1I3
NR1I3	10033668893	NR1I3.1
NR2F1	10025912092	NR2F1
NR2F1	10033668482	NR2F1.1
NR2F2	10025904230	NR2F2
NR2F2	10033669012	NR2F2.1
NR3C1	10023821395	NR3C1
NR5A2	10025913330	NR5A2
ONECUT1	10025908639	ONECUT1
PGRMC1	10023811817	PGRMC1
PPARA	10023821315	PPARA
PPARA	10023830863	PPARA.1
PPARA	10023836179	PPARA.2
PPARA	10023836937	PPARA.3
PPARA	10023849782	PPARA.4
PPARA	10033668496	PPARA.5
PPARD	10023810148	PPARD
PPARD	10025902985	PPARD.1
PPARD	10033668779	PPARD.2
PPARG	10023809100	PPARG
PPARG	10026391591	PPARG.1
PPARG	10033668769	PPARG.2
RXRA	10023824148	RXRA
RXRA	10033668836	RXRA.1
RXRB	10025910971	RXRB
RXRB	10033668596	RXRB.1
RXRG	10025909577	RXRG
RXRG	10033668565	RXRG.1
THRA	10025906409	THRA
THRB	10023820789	THRB
USF1	10025911390	USF1
VDR	10023815431	VDR
VDR	10023822267	VDR.1
VDR	10033668477	VDR.2
YY1	10023813066	YY1

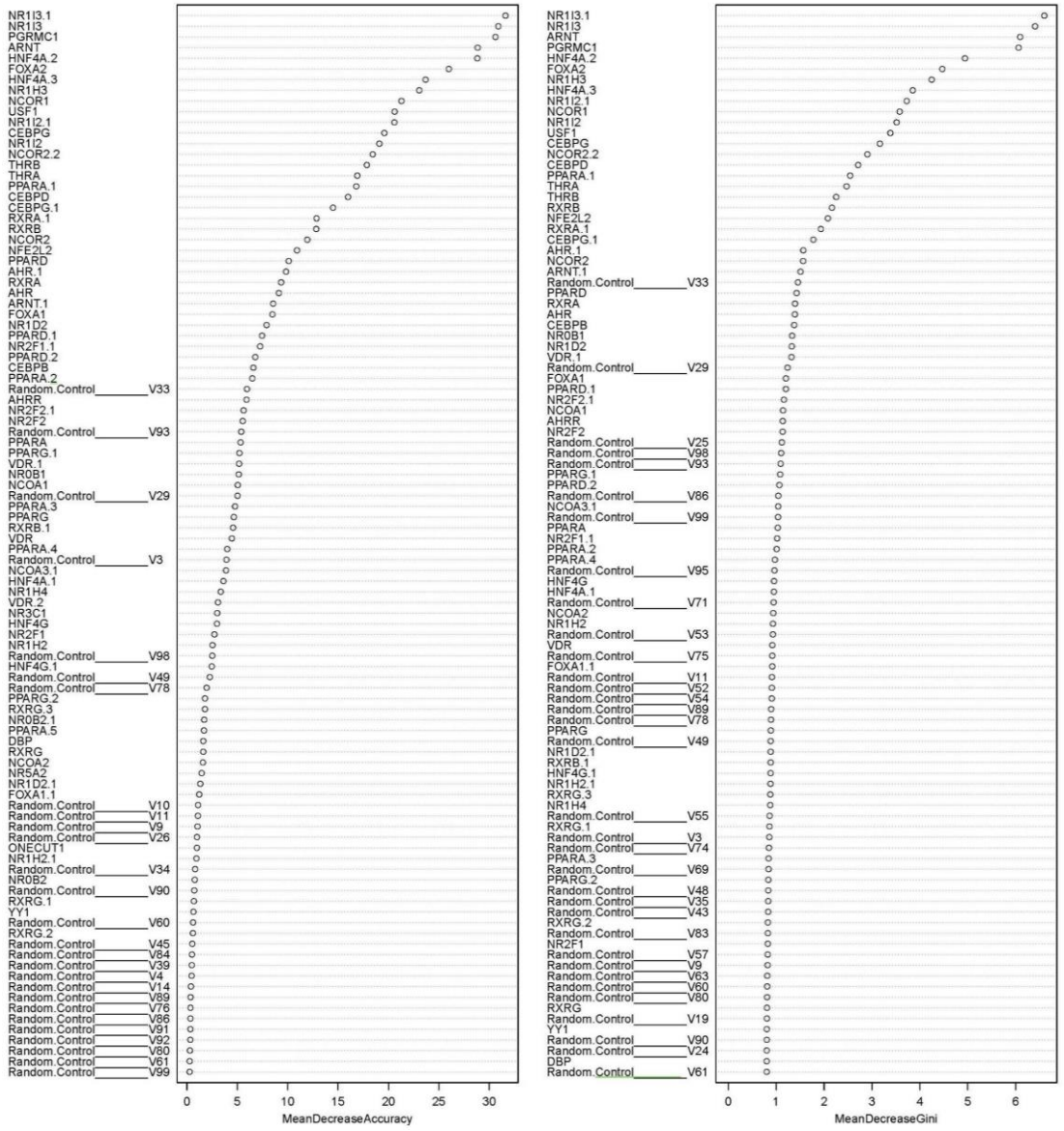
Supplemental Table 2. Sequence of primers and gRNA

gRNA for transcription activation	**Sequence**
NR1H3-201	
gRNA # 1	CCACCAGGTTCACGCCGAGA
gRNA #2	GCAAGCGGTCCGGCTGGAGC
gRNA #3	TGCATAGATTACAACGGTGA
NR1H3-211	
gRNA # 1	CGCTGGGTAAGGAGAGGAAG
gRNA #2	TGGAACTTGGCTGGTCTGCA
gRNA #3	TTTGGCCGGGAGTAGGGGGC
NR1H3-235	
gRNA # 1	CTGGTGAACGGTCTCCATGG
gRNA #2	GAATAAAATGGTTTGCCTAT
gRNA #3	TAGGCTCTGGGTCCCTATCA
negative control	ACGGAGGCTAAGCGTCGCAA
Real time PCR primers	**Sequence**
NR1H3 total	F: ATAACCGGGAAGACTTTGCCA
	R: GGTTGATGAATTCCACTTGCAG
NR1H3-211	F: TGTGCCTGACATTCCTCCTG
	R: CTTCCACAGCTCCACCGC
NR1H3-201	F: GGGTCGTGGTCTGGCTGT
	R: CTTCCACAGCTCCACCGC
NR1H3-235	F: GCCATCACCGTTGTAATCTATGC
	R: CTTCCACAGCTCCACCGC
NR1H2	F: TGCAGTGCAACAA ACGCTC
	R: GGCCAGGGCGTGACTTT
GAPDH	F: ACTCCTCCACCTTTGACGCT
	R: GGTCCACCACCCTGTTGC
POU5F1	F: CGAACCAGTATCGAGAACCGAG
	R: TTCTGGCGCCGGTTACAG
CER1	F: ATCTTGCCCATCAAAAGCCA
	R: CGGCTCCAGGAAAATGAACA

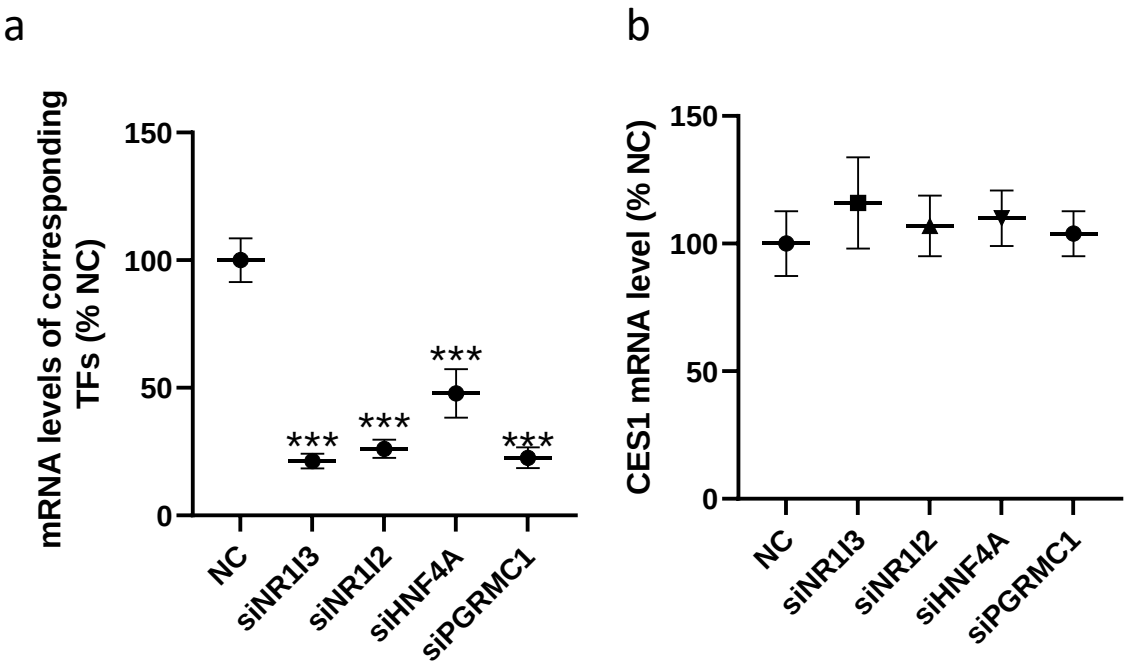
CYP3A4	F: CTCTCATCCCAGACTTGGCCA R: ACAGGCTGTTGACCATCATAAAAG
ABCG1	F: CTTCGTCAGCTTCGACACCA R: CTGGAAGTGGCACGTCTCG
FAS	F: AGCAGTACACACCCAAGGCC R: TGGTCACCCTCGATGACGT
NR1I2	F: ATGTGCTGATGCAGGCCAT R: AGCACACCTGGGCGGTC
NR1I3	F: CACATGGGCACCATGTTTGA R: AAGGGCTGGTGATGGATGAA
HNF4A	F: ACATGGACATGGCCGACTAC R: CTCGAGGCACCGTAGTGTTT

Supplemental Table 3. Sobol's indices for top 25 TFs

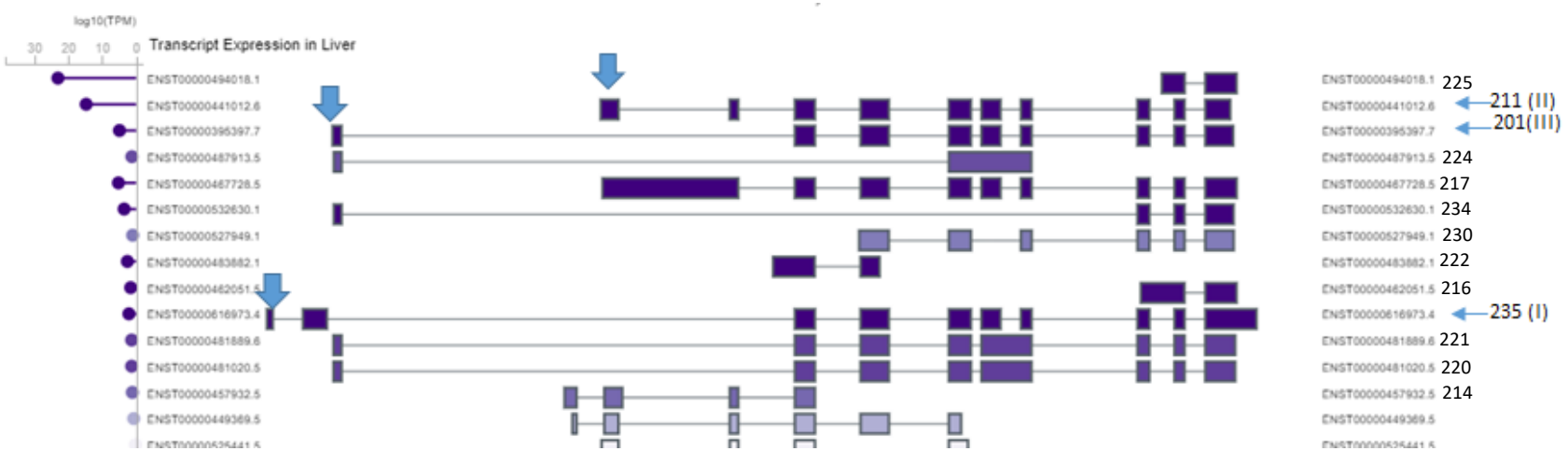
Gene	M.SI	M.ResDev.Pe	M.NullDev	Gene.Y
PGRMC1	0.016135	0.69064673	16.48143282	CES1
NR1I3	0.015792	0.6972257	16.48143282	CES1
NR1I3.1	0.015278	0.70706683	16.48143282	CES1
HNF4A.2	0.012495	0.76043716	16.48143282	CES1
HNF4A.3	0.012475	0.76081931	16.48143282	CES1
NR1H3	0.010807	0.79280343	16.48143282	CES1
ARNT.1	0.010325	0.80204274	16.48143282	CES1
NCOR1	0.010162	0.80516875	16.48143282	CES1
NR1I2.1	0.009311	0.82148384	16.48143282	CES1
NR1I2	0.009116	0.82521387	16.48143282	CES1
FOXA2	0.009089	0.8257423	16.48143282	CES1
ARNT	0.008906	0.82924152	16.48143282	CES1
NCOR2.2	0.008138	0.84396318	16.48143282	CES1
USF1	0.006826	0.86912378	16.48143282	CES1
CEBPG	0.006752	0.87054263	16.48143282	CES1
NCOR2	0.005389	0.89667583	16.48143282	CES1
PPARA.1	0.005353	0.89737277	16.48143282	CES1
NFE2L2	0.004439	0.91489536	16.48143282	CES1
RXRB	0.004325	0.91707452	16.48143282	CES1
AHR.1	0.004227	0.91896123	16.48143282	CES1
THRA	0.003778	0.92756596	16.48143282	CES1
CEBPG.1	0.003702	0.92902221	16.48143282	CES1
CEBPD	0.003099	0.94057905	16.48143282	CES1
THRB	0.00216	0.95858105	16.48143282	CES1
RXRA.1	0.001839	0.96473711	16.48143282	CES1



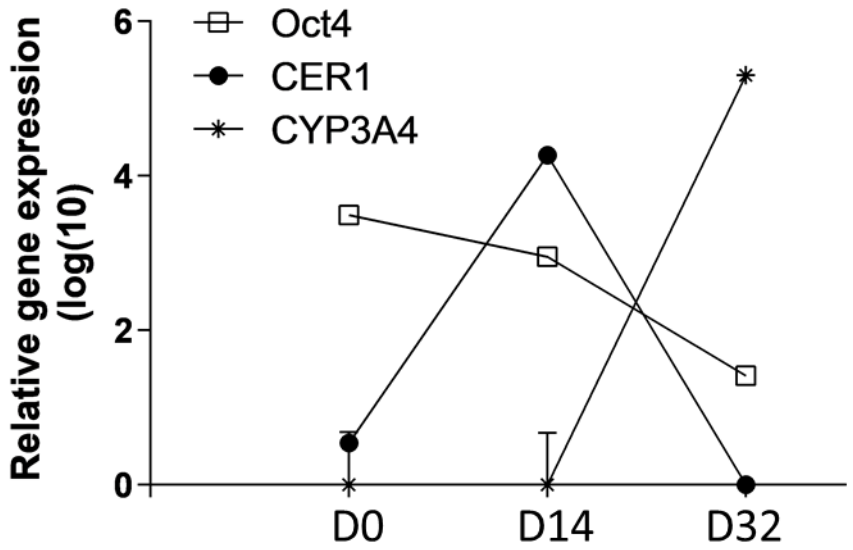
Supplemental Figure 1. TF mean decreases in Gini by fitting a random forest classifier of CES1. All TF effects on CES1 were simultaneously tested against 100 random controls (randomly generated Gaussian noise), using the mean decrease in Gini index of the random forest classifier of CES1 (dichotomized by CES1 median in GSE9588). Note: Gene names followed by a number represent results from different microarray probes, for example, NR1I3.1 and NR1I3 represent NR1I3 measured by two different probes.



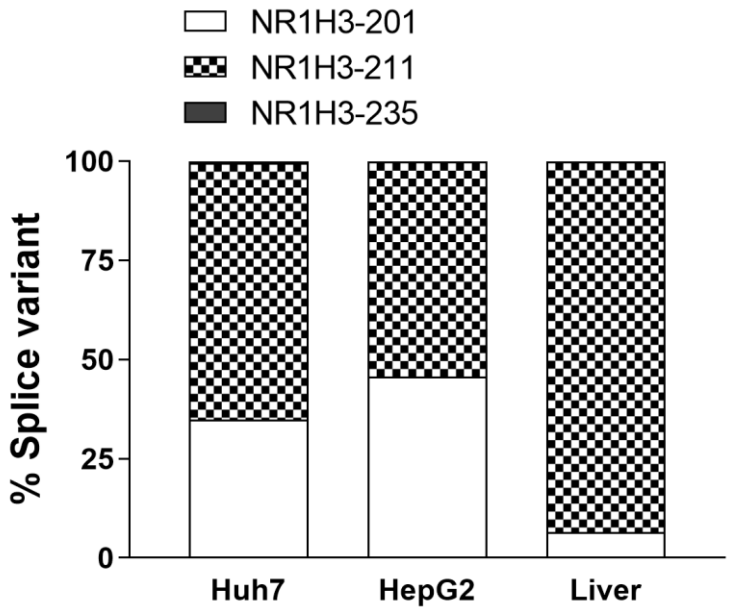
Supplemental Figure 2. The effects of siRNA knockdown of NR1I3, NR1I2, HNF4A and PGRMC1 on the expression of corresponding TFs (a) and CES1 (b) in Huh7 cells. Mean ± SD. Compared to negative control (NC), *** $p<0.0001$, ANOVA with Dunnett comparison with control posttest.



Supplemental Figure 3. The expression of different NR1H3 transcripts in human liver (data from GTEx portal (GTEx.portal). Downward arrows indicate the different promoters and the potential first exons. Horizontal arrows indicate the three transcripts that were targeted for transcriptional activation using CRISPR technology. Isoforms 211, 201 217, 235 and 221 are expected to be protein coding. Others are either processed intermediate transcripts or are expected to undergo nonsense mediated RNA decay.



Supplemental Figure 4. Changes in expression of marker genes during iPSC to hepatocyte differentiation. D0, day0; D14, day14; D32, day32.



Supplemental Figure 5. Relative expression of three NR1H3 splice isoforms in Huh7, HepG2 and liver tissue. Data are expressed as % of total NR1H3 level. NR1H3-235 level is too low (<1%) to be visible in the graph.