Received: 26 April 2022 | Revised: 11 August 2022 | Accepted: 23 August 2022

DOI: 10.1111/cts.13398

ARTICLE

Regulatory variants in a novel distal enhancer regulate the expression of CYP3A4 and CYP3A5

Joseph M. Collins[1] | **Adaeze C. Nworu**[1] | **Somayya J. Mohammad**[2] | **Liang Li**[3] | **Chengcheng Li**[3] | **Chuanjiang Li**[4] | **Ethan Schwendeman**[2] | **Mattew Cefalu**[2] | **Mahmoud Abdel-Rasoul**[5] | **Jessie W. Sun**[2,6] | **Sakima A. Smith**[2] | **Danxin Wang**[1]

[1]Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics, College of Pharmacy, University of Florida, Gainesville, Florida, USA

[2]Department of Internal Medicine, Division of Cardiology, College of Medicine, The Ohio State University, Columbus, Ohio, USA

[3]Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China

[4]Division of Hepatobiliopancreatic Surgery, Department of General Surgery, Nanfang Hospital, Southern Medical University, Guangzhou, China

[5]Center for Biostatistics, Department of Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, Ohio, USA

[6]School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

Correspondence
Danxin Wang, Department of Pharmacotherapy and Translational Research, College of Pharmacy and Center for Pharmacogenomics, University of Florida, 1345 Center Drive MSB PG-05B, Gainesville, Florida 32610, USA.
Email: danxin.wang@cop.ufl.edu

Funding information
National Institute of Health (NIH), Grant/Award Number: HHSN276201200017C, R01 MD011307, R35 GM140845 and R01GM120396

Abstract

The cytochrome P450 3As (CYP3As) are abundantly expressed in the liver and metabolize many commonly prescribed medications. Their expression is highly variable between individuals with little known genetic cause. Despite extensive investigation, cis-acting genetic elements that control the expression of the CYP3As remain uncharacterized. Using chromatin conformation capture (4C assays), we detected reciprocal interaction between a distal regulatory region (DRR) and the CYP3A4 promoter. The DRR colocalizes with a variety of enhancer marks and was found to promote transcription in reporter assays. CRISPR-mediated deletion of the DRR decreased expression of CYP3A4, CYP3A5, and CYP3A7, supporting its role as a shared enhancer regulating the expression of three CYP3A genes. Using reporter gene assays, we identified two single-nucleotide polymorphisms (rs115025140 and rs776744/rs776742) that increased DRR-driven luciferase reporter expression. In a liver cohort ($n = 246$), rs115025140 was associated with increased expression of CYP3A4 mRNA (1.8-fold) and protein (1.6-fold) and rs776744/rs776742 was associated with 1.39-fold increased expression of CYP3A5 mRNA. The rs115025140 is unique to the African population and in a clinical cohort of African Americans taking statins for lipid control rs115025140 carriers showed a trend toward reduced statin-mediated lipid reduction. In addition, using a published cohort of Chinese patients who underwent renal transplantation taking tacrolimus, rs776744/rs776742 carriers were associated with reduced tacrolimus concentration after adjusting for CYP3A5*3. Our results elucidate a complex regulatory network controlling expression of three CYP3A genes and identify two novel regulatory variants with potential clinical relevance for predicting CYP3A4 and CYP3A5 expression.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

There is limited knowledge concerning distal cis-acting regulatory elements and genetic variants controlling expression of the CYP3As.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Clinical and Translational Science* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

WHAT QUESTION DID THIS STUDY ADDRESS?

This study investigated whether undiscovered distal regulatory regions and/or variants may be contributing to variation in expression of the CYP3As.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

Our results identified a distal regulatory region (DRR) that controls the expression of CYP3A4, CYP3A5, and CYP3A7. This is the first report of a shared, distal enhancer regulating expression of multiple cytochrome P450 genes located in a gene cluster. In addition, two single-nucleotide polymorphisms in the DRR (rs115025140 and rs776744/rs746742) were identified that alter DRR transcriptional activity and are associated with increased expression of CYP3A4 and CYP3A5, respectively.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

These results are another step toward identifying biomarkers for predicting CYP3A expression/activity. Notably, if validated, rs115025140 may become the first frequent genetic biomarker for predicting CYP3A4 expression in African Americans. In addition, rs776744/rs746742 may explain additional inter-person variability of CYP3A5 not explained by CYP3A5*3, *6, and *7.

INTRODUCTION

The cytochrome P450 3As (CYP3As) are important drug-metabolizing enzymes as they are responsible for metabolism of ~50% of medications.[1] Particularly, CYP3A4 is the most abundant drug-metabolizing enzyme in the liver[2] and is also highly expressed in the intestines,[3] playing an important role in drug exposure and treatment outcomes. Activity of the CYP3As is impacted by several environmental factors (see review by ref. 4), their expression is highly variable between individuals,[2,5] and heritability of their activity is well-documented.[6,7] Despite this, the underlying genetic cause for CYP3A4 variation remains poorly understood, and current biomarkers appear to be better suited for predicting induced CYP3A expression as opposed to constitutive CYP3A activity.[8] CYP3A4 coding variants are rare and incapable of explaining the discrepancy in CYP3A4 activity between individuals, directly contributing to the difficulty in predicting CYP3A activity based on genotype. Our previously identified intronic regulatory variant CYP3A4*22 (rs35599367)[9] significantly reduces CYP3A4 expression in individuals carrying the variant and has been included in some pharmacogenomics genotyping panels, but CYP3A4*22 cannot explain large inter-person CYP3A4 variability in the population due to its low allele frequency (2%–8%). Therefore, the underlying genetic cause of variable CYP3A4 expression and activity remains largely unknown.

Three genetic elements that control gene transcription are the promoter, and proximal and distal enhancers. The CYP3A4 gene locus, including both its promoter and proximal enhancers, has been extensively studied.[10,11] Despite this, much of the variability in CYP3A4 remains unexplained, suggesting that many of the regulatory processes controlling CYP3A4 are unresolved. Furthermore, intestinal and hepatic CYP3A4 appears to be regulated differently because there is differential expression and distribution of CYP3A4 splicing isoforms between the two tissues (based on GTEx[12]) and CYP3A4*22 does not affect CYP3A4 expression in the intestines.[9,13] Because the liver is the main organ involved in drug metabolism, we focused on searching for regulatory domains/variants controlling the CYP3As in the human liver for this study. We hypothesized that distal enhancers regulate transcription of the four CYP3A genes in the CYP3A locus, and that this regulation may be dependent on varying interactions with their respective promoters. Previously, using circular chromatin conformation capture assays (4C assays), we identified three regions (R1, R2, and R4) located throughout the CYP3A locus that had enhancer-like features and interacted with the promoters of the four *CYP3A* genes in the human liver.[14] R2 overlaps with the previously identified CLEM4 enhancer region located ~10 kb upstream of CYP3A4 promoter,[15] whereas R1 and R4 are novel. These chromatin interactions appear inter-dependent, as the R2 and R4 have opposing effects on CYP3A4 and CYP3A43 expression, likely due to their competition for interaction with the different CYP3A promoters.[14] Conversely, the role of the R1 region was not clear. The R1 is located ~90 kb downstream of CYP3A4 promoter and ~13 kb upstream of CYP3A5 promoter, has enhancer-like features (based on chromatin immunoprecipitation-quantitative polymerase

chain reaction [qPCR] and reporter gene assays), yet initial CRISPR-mediated deletion of the R1 failed to change expression of CYP3A4.[14] Based on its enhancer-like features, we decided to investigate the R1 in more detail to determine whether it has a regulatory role that we may have previously missed.

We discovered that the previously identified R1 comprises only half of an enhancer region (Figure S1), which may explain why it did not previously appear to regulate CYP3A4. Here, we report that this extended R1 (~4 kb in length) functions as an enhancer for CYP3A4, CYP3A5, and CYP3A7. Based on these results, we have renamed this region as the CYP3A distal regulatory region (DRR). Furthermore, two single-nucleotide polymorphisms (SNPs) located within the DRR (rs115025140 and rs776744/rs776742) are associated with increased expression of CYP3A4 and CYP3A5, respectively. In clinical cohorts, we observed that carriers of these SNPs have higher CYP3A activity, as indicated by reduced statin-mediated LDL reduction in rs115025140 carriers and lower steady-state tacrolimus dose in rs776744/rs776742 carriers.

METHODS

See Supplemental File for materials and methods.

RESULTS

The distal regulatory region regulates expression of CYP3A4, CYP3A5, and CYP3A7

We used online tools to further evaluate the R1 and its surrounding regions. Visualizing ENCODE data[16,17] on the University of California Santa Cruz (UCSC) genome browser[18] showed that the R1 region overlapped with several features indicative of enhancers (Figure S1). This qualitative approach also revealed a second region nearby the R1 as a potential enhancer. Thus, we named the whole region as the CYP3A DRR and renamed the original R1 region "DRRα" and the new region "DRRβ." The ENCODE data[16,17] indicates that both the DRRα and DRRβ colocalize with transcription factor (TF) binding clusters, DNase hypersensitivity sites, H3K4me1/H3K27ac histone marks, and cap analysis gene expression (CAGE) data[19] support multiple enhancer RNAs across the region (Figure S1). CistromeDB[20] analysis of the DRR also identified over a hundred TF binding sites, H3K4me1/H3K27ac marks, and chromatin accessibility across the region. Furthermore, using chromatin immunoprecipitation (ChIP)-qPCR in hepatocytes, we detected strong enrichment of the P300 histone acetyltransferase protein across the DRRβ region, to an even greater degree than we previously observed at the DRRα[14] (fold enrichments: 68 ± 7, 78 ± 28, and 130 ± 24, in hepatocytes from three donors, respectively). These results support that the DRRα and DRRβ, either alone or together, are potential active enhancers.

To test for physical contact between the DRR and promoters of the *CYP3A* genes, we performed 4C assays in human primary hepatocytes. We decided to use hepatocytes from both White and Black donors (pooled samples from three donors per race) to test whether there were any obvious racial differences in the chromatin interaction maps. In both samples, we observed a clear 4C peak at the DRR when using the CYP3A4 promoter as a viewpoint (Figure 1a, top two panels, indicated by a black arrow), confirming our previous results.[14] In addition, in a reciprocal experiment using the DRR as the bait, we observed a peak near the CYP3A4 promoter (Figure 1a, middle panels), further indicating that the DRR physically interacts with the CYP3A4 promoter. However, due to the spatial proximity of the CYP3A5 promoter and DRR (~13 kb) (Figure 2a), when using either region as the 4C bait, strong 4C signals made identification of clear interaction peaks difficult (Figure 1a, bottom panels), possibly resulting from frequent 4C self-ligation events. An interaction peak was also found near the DRR when using the CYP3A7 promoter as the viewpoint, but not when the DRR itself was used as the viewpoint (Figure S2). Similarly, the CYP3A43 promoter also showed interaction with the DRR, although there was no apparent peak in the reciprocal experiment (Figure S2).

We then conducted Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq) in the same primary hepatocytes to identify chromatin-accessible regions that are indicative of regulatory elements.[21] We identified ATAC-Seq peaks at the DRR region (Figure 1b). We also analyzed published hepatocyte ChIP-Seq datasets for H3K4me1, K3K27ac, and p300,[22] all of which displayed peaks covering the DRR (Figure 1c). These results indicate that the DRR is an active enhancer region, consistent with our in silico results. We did not observe any differences in 4C peak and ATAC-Seq signals between hepatocytes from White and Black donors occurring within the CYP3A locus.

To determine whether the DRR regulates the *CYP3A* genes, we used CRISPR-mediated genome editing to target and delete each DRR subregion and then measured CYP3A mRNA expression. These experiments revealed that loss of the DRRβ decreased expression of CYP3A4, CYP3A5, and CYP3A7 (Figure 2b–e). In contrast, deletion of DRRα did not significantly alter expression of CYP3A4 or CYP3A5, but did decrease expression of CYP3A7 (Figure 2d), in agreement with our previous results.[14] Expression of CYP3A43

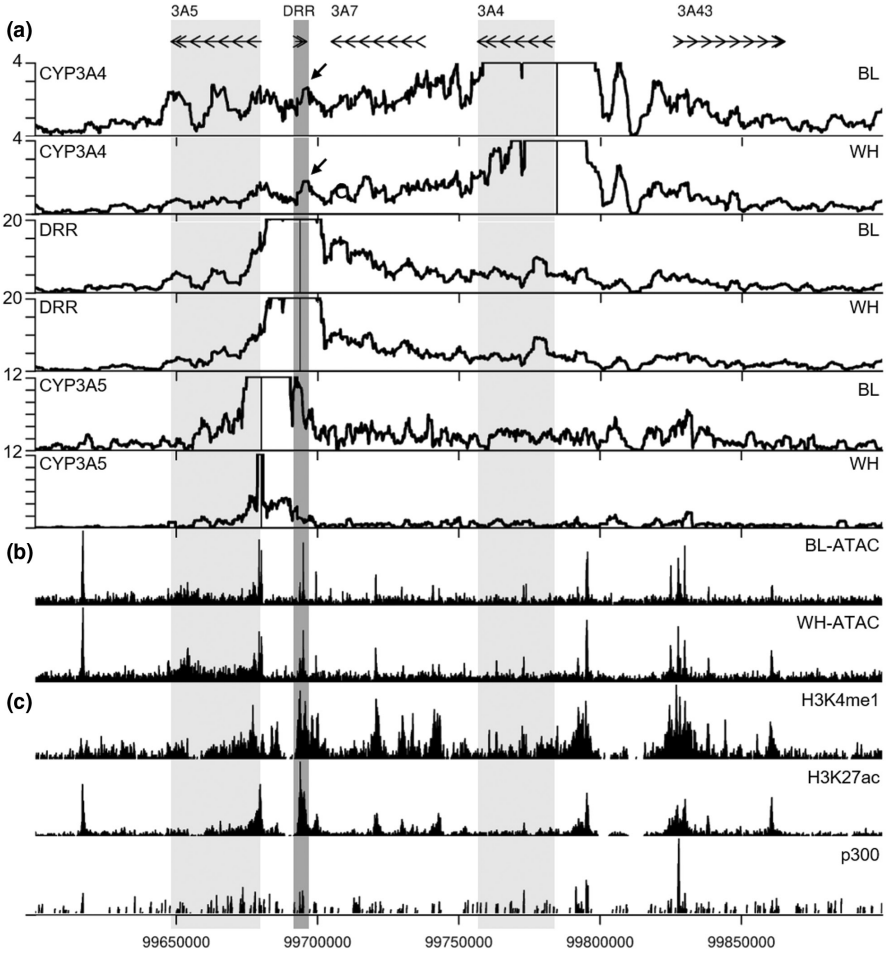

FIGURE 1 Chromatin interactions, enhancer marks, and open chromatin regions across the CYP3A locus. (a) The top six panels illustrate 4C peak data conducted in human primary hepatocytes using different viewpoints, specifically CYP3A4, the DRR, and CYP3A5 (indicated in the upper left). These 4C panels are arranged in pairs, with results from both Black (BL) and White (WH) donor samples shown (indicated in the upper right). The vertical black line in each panel indicates the location of the viewpoint. Values on the far left indicate the "maximum y-plot" setting used to generate each plot in the 4SEE package.[39] A black arrow on both 3A4 viewpoint plots (top two panels) indicates the peak occurring at the DRR region. Similarly, the arrow on both DRR viewpoint plots indicates the peak near the CYP3A4 promoter. A CYP3A locus diagram is included at the top of the plot and the shaded vertical boxes highlight each CYP3A region: light gray indicates the CYP3A4 and CYP3A5 gene regions and dark gray indicates the DRR. Genomic position on chromosome 7 is indicated at the bottom of the figure. (b) ATAC-Seq signal in human primary hepatocytes. Results from both Black and White donor samples are shown as indicated on the far right (BL-ATAC and WH-ATAC, respectively). (c) ChIP-Seq data for H3K4me1, H3K27ac, and p300 analyzed from publicly available data.[22] 4C, circular chromatin conformation capture assays; ATAC-Seq, Assay for Transposase-Accessible Chromatin using sequencing; DRR, distal regulatory region.

was unaffected by loss of either DRR subregion (Figure 2e). These results indicate that the DRR regulates expression of CYP3A4, CYP3A5, and CYP3A7, and it appears that the DRRβ plays a more important role than the DRRα. We also attempted to delete the entire 4 kb DRR, but cell viability was significantly reduced, which made assessing its impact on *CYP3A* gene expression unreliable.

We then used reporter gene assays to test for transcriptional activity of the DRR. Both of the individual DRRα and DRRβ regions showed increased expression of the reporter, but inclusion of the entire DRR led to significantly higher expression (Figure 3a). These results indicated that the entire DRR functions as stronger enhancer than either subregion alone, similar to the CYP7A1 enhancer that we previously discovered.[23] Taken together, the combined results of our in silico, ChIP-qPCR, ATAC-seq, 4C, and reporter gene assays support that the DRR functions as a shared enhancer controlling the expression of CYP3A4, CYP3A5, and CYP3A7.

SNPs located within the DRR alter expression of CYP3A4 and CYP3A5

We scanned the DRR for SNPs and found seven: rs118168183, rs776745, rs111266634, rs776744, rs55830753, rs115025140, and rs776742 (Figure 3b, Table 1). Two

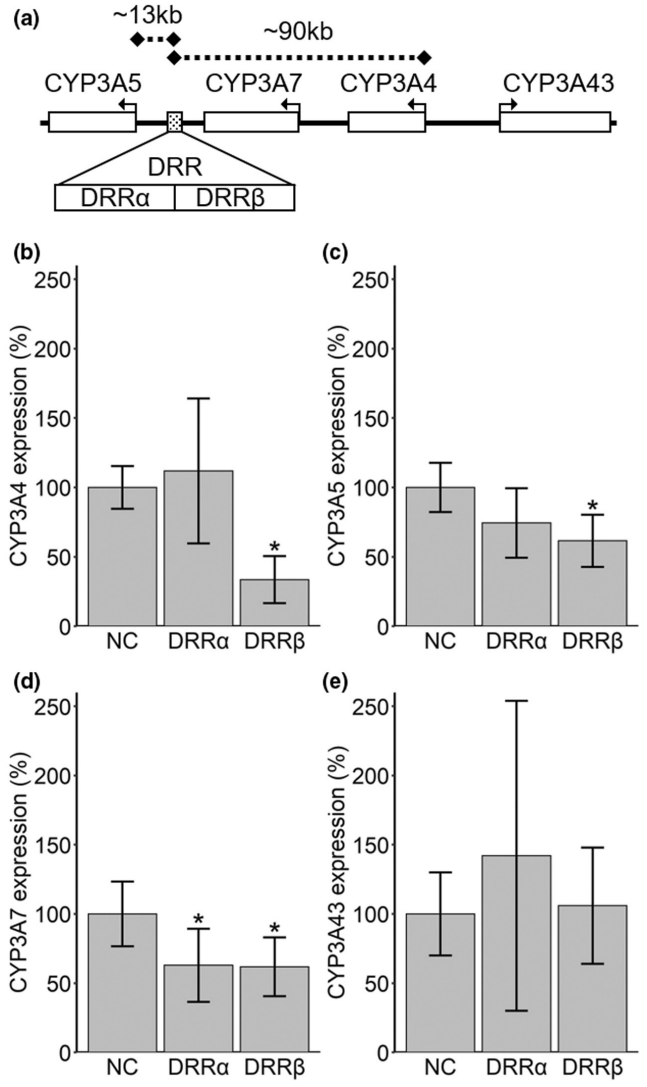

FIGURE 2 CRISPR-mediated deletion of the DRR and subsequent effect on expression of the CYP3A genes. (a) CYP3A locus diagram illustrating location of the DRR. Effect of CRISPR-mediated deletion of the DRRα or DRRβ on expression of (b) CYP3A4, (c) CYP3A5, (d) CYP3A7, and (e) CYP3A43. NC, negative control; *Compared to negative control, $p < 0.05$. DRR, distal regulatory region.

are unique to individuals of African ancestry (AFR; rs111266634 and rs115025140) and three highly linked SNPs (rs776745, rs776744, and rs776742) have very different allele frequencies between European (EUR) and AFR individuals (Table 1). Thus, the frequencies of the six common haplotypes (DRR-H1 through DRR-H6) formed by these seven SNPs varies greatly in different ancestral backgrounds (Table 1). DRR-H1 is the predominant reference haplotype in both EUR and East Asian (EAS) populations, with frequencies of 90% and 71%, respectively. DRR-H3 through DRR-H6 are common haplotypes occurring in AFR individuals with a combined frequency of 69%. We cloned these six haplotypes and tested their ability to promote expression of the luciferase reporter (Figure 3c). DRR-H1 showed the highest luciferase activity, whereas the haplotypes DRR-H2 through DRR-H5 all led to a drastic decrease in luciferase expression (Figure 3c). DRR-H6, which contains rs115025140 and is an alternate AFR haplotype, showed luciferase activity approaching DRR-H1. These results indicate that variants modify transcriptional activity of the DRR and that rs115025140 in DRR-H6 promotes expression when compared to the other common AFR haplotypes.

We also tested the DRRα and DRRβ separately to evaluate their independent effects (haplotypes in Table 1). Again, constructs were generated based on common haplotypes occurring in either EUR or AFR individuals. Of the three SNPs occurring in the DRRα (rs118168183, rs776745, and rs111266634), only the haplotype containing rs111266634 (only occurs in AFR individuals), was found to significantly alter reporter expression (~50%–77% decrease; Figure 3d). In the DRRβ, rs776744 and rs776742 (in complete linkage disequilibrium [LD]) led to roughly three-fold increased luciferase reporter activity regardless of the presence or absence of the other two SNPs (rs55830753 and rs115025140; Figure 3e). Based on these results from testing the entire DRR and the DRRα/DRRβ separately, rs118168183, rs115025140, rs111266634, and rs776744/rs776742 have potential regulatory roles controlling expression of *CYP3A4* and *CYP3A5* genes.

We then tested for association among these four SNP genotypes and expression of CYP3A4 and CYP3A5 in a 246-sample liver cohort (see Table 2 for demographic information). The rs115025140 was associated with increased CYP3A4 mRNA expression, with each variant G allele being associated with 1.8-fold increased CYP3A4 mRNA, after adjusting for CYP3A4*22[9] and the expression of several key TFs known to affect CYP3A4 expression[24] (Figure 4a). Consistently, the rs115025140 G allele is also associated with higher CYP3A4 protein expression (1.6-fold per allele) after adjusting for the same covariates (Figure 4b). Because rs776744/rs776742 are in perfect LD, we chose to genotype rs776744 as a representative of the two linked SNPs. The rs776744/rs776742 carriers were associated with increased CYP3A5 mRNA expression (1.36-fold; Figure 4c) after adjusting for expression of key TFs and the CYP3A5*3, *6, and, *7 genotypes, known to affect CYP3A5 expression.[11] In contrast, there was no association between rs111266634 or rs118168183 genotype and expression of either CYP3A4 or CYP3A5, despite these SNPs appearing to reduce transcriptional activity in reporter gene assays (Figure 3). Overall, these results suggest that the DRR SNPs rs115025140 and rs776744/rs776742 modify the expression of *CYP3A4* and *CYP3A5*, respectively.

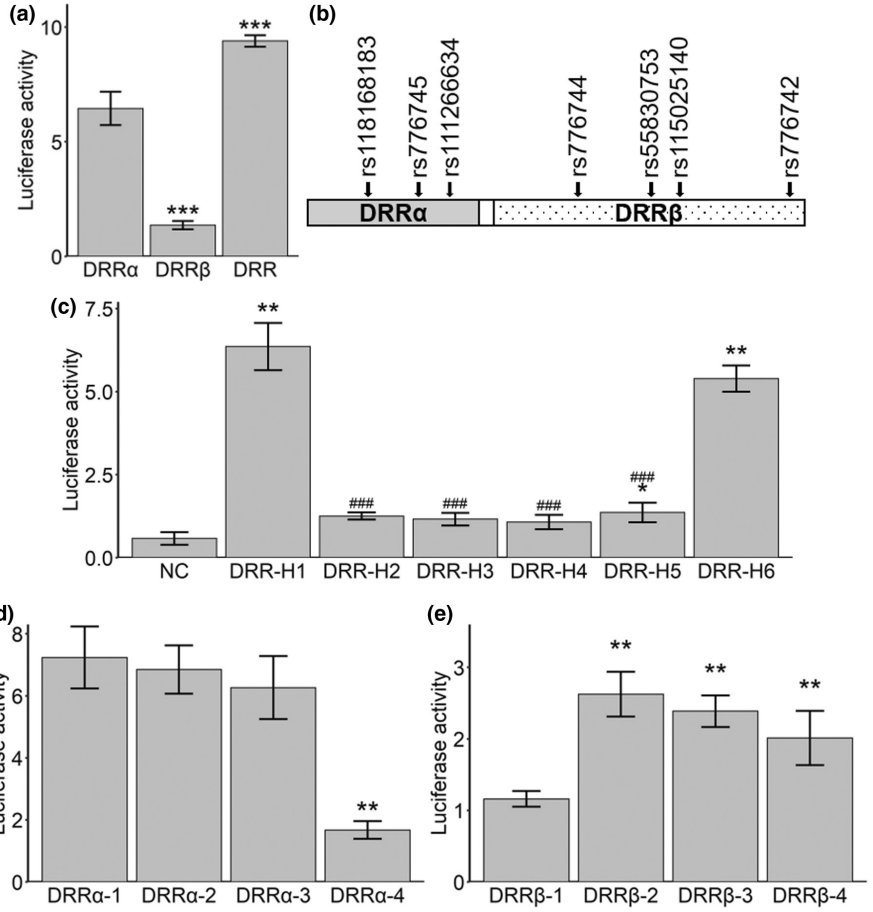

FIGURE 3 Impact of DRR haplotypes on luciferase reporter activity. (a) Luciferase reporter assay using the DRR or isolated DRRα or DRRβ region. Results are shown relative to the pGL3 empty vector, which is set at 1. ***Compared to DRRα, $p < 0.001$. (b) Diagram of the DRR and location of SNPs tested. (c) Luciferase reporter activity of the various DRR haplotypes in the full-length DRR construct. Compared to pGL3 (NC), **$p < 0.01$, ***$p < 0.001$. Compared to DRR-H1, $^{###}p < 0.001$. Luciferase reporter activity when the various haplotypes are tested in the (d) isolated DRRα and (e) isolated DRRβ. Results are shown relative to the pGL3 empty vector, which is set at 1. Compared to DRRα-1 or DRRβ-1, **$p < 0.01$. DRR, distal regulatory region; SNPs, single-nucleotide polymorphisms.

Clinical association studies support that the DRR variants rs115025150 and rs776744/rs776742 alter CYP3A expression

The in vivo effects of rs115025140 and rs776744/rs776742 on CYP3A metabolism were tested in two clinical cohorts. Because rs115025140 is unique to the AFR population, we tested this SNP in a cohort of African Americans recruited from the community and medical clinics (Columbus, OH; $n = 230$) who were taking atorvastatin, simvastatin, or lovastatin for lipid control. Patient characteristics are shown in Table 2. Although not statistically significant, rs115025140 variant G allele carriers showed a trend toward having less statin-mediated lipid reduction than non-carriers (-24.8 ± 17.8 vs. -41.9 ± 21.2, $p = 0.163$) after adjusting for diabetes, hypertension, age, body mass index, and sex, consistent with rs115025140 increasing reporter expression and the allele being associated with higher mRNA and protein expression of CYP3A4 in the liver (Figures 3 and 4).

We used a published cohort of Chinese patients who underwent renal transplantation who were taking tacrolimus to prevent graft rejection[25] to test the association between rs776744/rs776742 genotype and steady-state concentration of tacrolimus (C0) adjusted by body weight (W) and dosage (D; Log(C0/D*W). Patient characteristics were the same as previously reported.[25] To control for the influence of CYP3A5*3 (rs776746), we analyzed the association of rs776744/rs776742 in different CYP3A5*3 genotype groups separately. The rs776744/rs776742 variants were associated with reduced tacrolimus concentration in CYP3A5*3/*1 and *3/*3 groups (Figure 5). In CYP3A5*1/*1 group, rs776744 CC and CT genotypes appear to have higher tacrolimus concentration, but this is likely driven by one individual homozygous for CYP3A5*1 and the rs776744/rs776742 reference allele who also had very low tacrolimus concentration. These results are consistent with rs776744/rs776742 causing increased transcriptional activity and its association with increased CYP3A5 expression in the liver (Figures 3e and 4c).

DISCUSSION

We characterized a distal regulatory region DRR within the CYP3A locus as a shared enhancer for multiple CYP3A genes using CRISPR-mediated deletion (Figure 2), chromatin accessibility assays (Figure 1), and reporter gene assays (Figure 3). This is further supported by direct physical contacts that occur between the DRR and promoters

TABLE 1 DRR haplotypes tested and frequencies in relevant populations

DRR haplotypes and frequencies						
rs#	**DRR-H1**	**DRR-H2**	**DRR-H3**	**DRR-H4**	**DRR-H5**	**DRR-H6**
rs118168183	G	***A***	G	G	G	G
rs776745	T	T	***G***	***G***	***G***	***G***
rs111266634	G	G	G	G	***A***	G
rs776744	C	C	***T***	***T***	***T***	***T***
rs55830753	A	A	A	***G***	A	A
rs115025140	A	A	A	A	A	***G***
rs776742	C	C	***T***	***T***	***T***	***T***
Population						
EUR	0.8956	0.0318	0.0239	0.0487	0	0
AFR	0.3071	0	0.3888	0.1415	0.0787	0.0794
EAS	0.7133	0	0.2579	0.0288	0	0

DRRα Haplotypes				
rs#	**DRRα-1**	**DRRα-2**	**DRRα-3**	**DRRα-4**
rs118168183	G	***A***	G	G
rs776745	T	T	***G***	***G***
rs111266634	G	G	G	***A***

DRRβ Haplotypes				
rs#	**DRRβ-1**	**DRRβ-2**	**DRRβ-3**	**DRRβ-4**
rs776744	C	***T***	***T***	***T***
rs55830753	A	A	A	***G***
rs115025140	A	A	***G***	A
rs776742	C	***T***	***T***	***T***

Note: The bold italic letters indicate the base pair changes in variant alleles.

Abbreviations: AFR, African ancestry; DRR, distal regulatory region; EAS, East Asian ancestry; EUR, European ancestry.

of the CYP3A genes in 4C assays (Figure 1 and Figure S2). We then identified two variants located within the DRR (rs115025140 and rs776744/rs776742) that increased transcription of the luciferase reporter (Figure 3). Using a liver sample cohort, we determined that the variant rs115025140 is associated with higher CYP3A4 mRNA and protein expression (Figure 4a,b) and that rs776744/rs776742 are associated with higher CYP3A5 mRNA expression (Figure 4c). Finally, we observed a trend toward reduced statin-mediated LDL reduction in rs115025140 carriers and lower steady-state tacrolimus concentration in rs776744/rs776742 carriers (Figure 5). As the AFR population has low allele frequency for CYP3A4*22 (0.08% vs. 5% in EUR individuals), the only common CYP3A4 genetic biomarker available to date, validation of rs115025140 (6%–8% in AFR individuals) may help predict CYP3A4 activity in those of AFR descent. Moreover, some diseases (e.g., atherosclerotic cardiovascular disease) disproportionally impact AFR populations and clinically it is important to identify genetic biomarkers that can help guide statin-mediated therapy (or other CYP3A4 substrate drugs) in this population. In addition, the two variants rs776744/rs776742 are in high LD with the functional CYP3A5*1 allele and likely explain additional CYP3A5 variability not captured by the known variants CYP3A5*3, *6, and *7.

The DRR is a shared enhancer for the *CYP3A* genes

By leveraging publicly available ENCODE data[16,17] paired with the UCSC genome browser,[18] we identified that the DRR consists of two adjacent regions (DRRα and DRRβ) that both colocalize with known enhancer features: clusters of TF-binding sites, DNase hypersensitivity, H3K4me1/H3K27ac marks, and CAGE signals (Figure S1). In our previous experiments, deletion of the DRRα (previously named the R1) did not affect CYP3A4 expression[14] as confirmed here in our replicate experiment (Figure 2b).

TABLE 2 Liver, hepatocyte, and patient characteristics

Liver donors	**AFR, *n* = 113**	**EUR, *n* = 133**	***p* value**
Sex (female), *n*, %	55, 48.6%	70, 52.6%	0.61
Age	55.8 ± 20.6	60.5 ± 13.2	0.04
Primary hepatocyte information			
Lonza lot #	**Gender**	**Age**	**BMI**
White donor pool			
HUM182391	Male	46	32.5
HUM182701	Male	31	22.1
HUM190131	Female	38	31.2
Black donor pool			
HUM183121	Male	49	30.7
HUM4056B	Male	55	28.9
HUM4225	Female	24	21.9
Stain cohort characteristics	**rs115025140 genotype**		
	CC, *n* = 215	**CT, *n* = 15**	***p* value**
Age (mean + SD)	60.48 (9.67)	61.51 (9.55)	0.665
Gender (female %)	112 (52.83%)	7 (38.89%)	0.256
BMI (Mean + SD)	33.60 (8.04)	35.01 (8.62)	0.477
History of hypertension (yes %)	197 (92.92%)	17 (94.44%)	>0.999
History of diabetes (yes %)	116 (54.72%)	9 (50.00%)	0.807
Smoking status (active %)	49 (23.56%)	7 (38.89%)	0.160
Statin			
Atorvastatin	181 (85.38%	16 (88.89%)	>0.999
Lovastatin	1 (0.47%)	0 (0.00%)	
Simvastatin	30 (15.15%)	2 (11.11%)	
Pre-statin cholesterol			
≤125 (low)	25 (11.79%)	2 (11.11%)	0.497
125–200 (ideal)	109 (51.42%)	7 (38.89%)	
≥200 (high)	78 (36.79%)	9 (50.00%)	
Post-statin cholesterol			
≤125 (low)	57 (26.89%)	8 (44.44%)	0.312
125–200 (ideal)	126 (59.43%)	8 (44.44%)	
≥200 (high)	29 (13.68%)	2 (11.11%)	
Pre-statin A1C			
≤5.7 (ideal)	43 (22.28%)	6 (42.86%)	0.169
5.7–6.4 (pre-diabetes)	71 (36.79%)	5 (35.71%)	
≥6.4 (diabetes)	79 (40.93%)	3 (21.43%)	
Post-statin A1C			
≤5.7 (ideal)	39 (20.42%)	7 (43.75%)	0.100
5.7–6.4 (pre-diabetes)	53 (27.75%)	4 (25.00%)	
≥6.4 (diabetes)	99 (51.83%)	5 (31.25%)	

Abbreviations: AFR, African ancestry; BMI, body mass index; EUR, European ancestry.

Instead, loss of the DRRβ led to decreased expression of CYP3A4, CYP3A5, and CYP3A7 (Figure 2). These results indicate that the DRR regulates the constitutive expression of three CYP3A genes and that the DRRβ is a necessary component for the DRR's enhancer activity. However, luciferase reporter assay results showed that the full-length

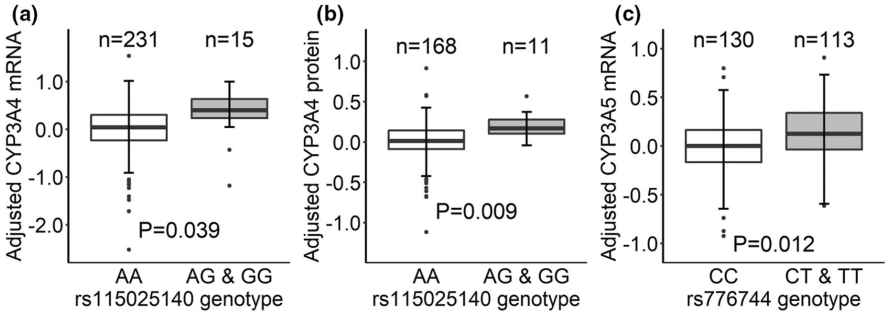

FIGURE 4 Association between two DRR SNPs and CYP3A4 or CYP3A5 expression. Association between CYP3A4 expression (after adjusting for CYP3A4*22 and expression levels of TFs, see Methods) and rs115025140 genotype at the (a) mRNA and (b) protein levels. (c) Association between CYP3A5 mRNA expression and rs776744 genotype after adjusting for covariates (CYP3A5*3, *6, *7 and expression levels of TFs, see Methods). DRR, distal regulatory region; SNPs, single-nucleotide polymorphisms; TF, transcription factor.

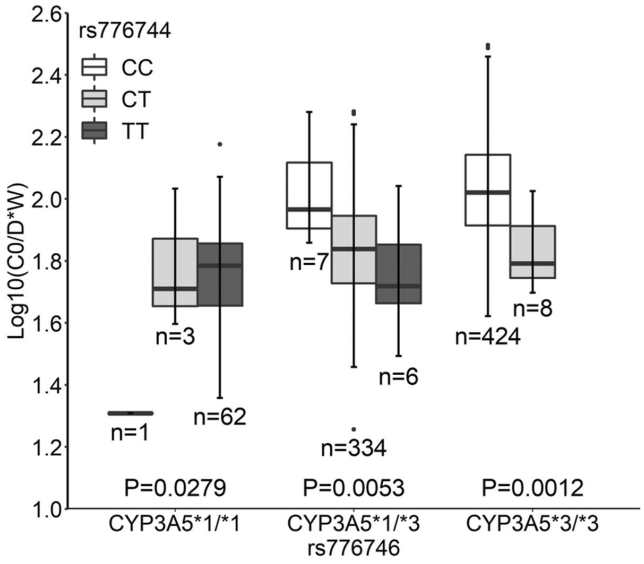

FIGURE 5 Association between rs776744 genotype and dose adjusted tacrolimus concentration in patients who undergo renal transplantation stratified by CYP3A5*3 genotype.

DRR has higher overall enhancer activity compared to either separate region alone (Figure 3a), suggesting the entire DRR is required for maximal enhancer activity. This is consistent with our previous study of a CYP7A1 enhancer, where two enhancer regions separated by ~2 kb had much stronger activity together than either alone.[23] Our results demonstrate for the first time that a distal enhancer within the CYP3A locus regulates the expression of three CYP3A genes, indicating complex gene regulation occurring in the CYP3A locus. The interaction network appears to consist of at least four enhancers identified to date: the previously identified inducible xenobiotic-responsive enhancer module (XREM)[26] and constitutive enhancer CLEM4,[15] our recently identified R2 (overlaps with CLEM4) and R4 regions,[14] and the DRR identified here. It appears then that regulation of the CYP3As behave similarly to other genomic loci, where multiple genes are controlled by the same enhancer and multiple enhancers regulate the same gene.[27]

Discovery of regulatory variants controlling expression of CYP3A4 and CYP3A5

Six common haplotypes formed by seven SNPs in the DRR show varying haplotype frequencies in each ancestral background (Table 1). In liver samples, we identified that rs115025140 is associated with increased CYP3A4 expression (Figure 4a,b) and that rs776744/rs776742 are associated with higher CYP3A5 expression (Figure 4c). Both of these variants are located in the DRRβ, consistent with the DRRβ as critical component in CYP3A regulation. It is unclear how these SNPs modify either the expression of CYP3A4 or CYP3A5 but not both. In the liver, expression of CYP3A4 and CYP3A5 is correlated,[24] but there is evidence for differences in their regulation.[28] For instance, the key CYP3A regulator ESR1 plays a larger role in CYP3A4 expression than that of CYP3A5 in the liver[29] and CYP3A5 is more broadly expressed across tissues than CYP3A4 (based on GTEx[12]). Therefore, the CYP3A-specific effects of rs115025140 and rs776744/rs776742 indicate that the DRR may contribute to the differential expression patterns of CYP3A4 and CYP3A5.

Although the underlying mechanism is not clear, it is possible that these SNPs alter TF binding motifs that contribute differently to the expression of each gene. As TFs are known to mediate interactions between enhancers and promoters[30] (thereby promoting transcription), this hypothetical change to TF binding would in turn affect DRR-CYP3A promoter contacts. Indeed, rs115025140, rs776744, and rs776742 are all predicted in HaploReg (version 4.1)[31] to change binding motifs for several different TFs. Of

note, the rs115025140 G allele generates a better motif for ZNF143, a chromatin-looping factor.[32] Previously, we found that in multiple cells types (including hepatocytes and the liver) ESR1 chromatin binding often colocalizes with the ZNF143 motif and that the interaction of these two proteins may play an important role in gene regulation through long-range chromatin interactions.[33] Therefore, it is possible that the rs115025140 variant promotes interaction of the DRR with CYP3A4 through increased binding of ZNF143 and/or ZNF143/ESR1, thereby increasing expression of CYP3A4.

The effect of the three highly linked SNPs (rs776745, rs776744, and rs776742) are not as clear as rs115025140. In reporter assays, the full-length DRR indicated that AFR haplotypes containing rs776745/rs776744/rs776742 typically have reduced luciferase activity compared to DRR-H1 (Figure 3c). However, analysis of each independent DRR fragment revealed the opposite, as rs776744/rs776742 (located in the DRRβ) instead resulted in increased reporter expression (Figure 3e), whereas rs776745 (DRRα) did not have a significant effect (Figure 3d). In our liver cohort, we found that these SNPs were associated with increased CYP3A5 expression (Figure 4c), supporting their positive role on DRR-mediated CYP3A5 regulation. These results indicate that additional factors coordinating DRR enhancer activity (e.g., TF binding and chromatin context) in vivo are not fully recapitulated in the reporter gene assay, a well-characterized limitation of the approach.[34] Notably, rs776742 (G>T) ablates a CpG site, and therefore may alter DNA methylation in vivo, another well-characterized epigenetic mode of gene regulation.[35] Based on the positive association between rs776745/rs776744/rs776742 genotype and CYP3A5 expression, it seems likely that either rs776744 or rs776742 (or both) are the functional variant, as these two promote DRRβ-driven luciferase expression.

Clinical relevance of the DRR regulatory variants

CYP3A4*22 is the only common and clinically relevant variant in CYP3A4. However, the allele frequency of CYP3A4*22 is very low in AFR (minor allele frequency [MAF] ~0.9%), and therefore cannot explain the observed variability of CYP3A4 expression and activity in this population. We found that the AFR-specific SNP rs115025140 (MAF 6%–8%) increases transcriptional activity in reporter gene assays (Figure 3) and is associated with increased expression of CYP3A4 mRNA and protein in liver samples (Figure 4). Thus, due to its frequency, rs115025140 is the first functional SNP that can explain a portion of CYP3A4 variability in AFR individuals. The CYP3As metabolize statins, decreasing their bioavailability and lipid-reducing effects[36] and, therefore, as CYP3A expression/activity is increased, statin exposure is expected to decrease. Our results indicate that rs115025140 carriers have reduced statin-mediated lipid reduction, although not at a statistically significant level. This lack of statistical significance is possibly driven by the small sample size and heterogeneity of the study cohort (demographics in Table 2). In addition, the cohort used for this analysis was not originally intended for studying the effect of rs115025140 and patients were not selected based on the presence of this SNP for inclusion. Furthermore, patients were enrolled from several outpatient clinics and confounding factors (e.g., statin dosage, type of statin, patient compliance, comedications, etc.) may have interfered with our ability to detect statin outcome differences based on rs115025140 genotype. Due to its potential to increase expression of CYP3A4, future studies are warranted to test the clinical impact of rs115025140.

CYP3A5*3, as well as CYP3A5*6 and *7 in AFR individuals, are currently the main genetic variants affecting CYP3A5 expression and enzyme activity. However, there remains variability in CYP3A5 expression and activity even after accounting for these three SNPs.[10,37] Here, we observed that rs776744/rs746742 increase expression of the luciferase reporter (Figure 3e) and are also associated with an increase of CYP3A5 expression in liver samples (Figure 4c). Consistent with these results, after adjusting for CYP3A5*3 genotype, rs776744/rs746742 are associated with reduced tacrolimus concentration in a large clinical cohort of Chinese patients who underwent renal transplantation (Figure 5). In the EAS population, rs776744/rs746742 are in high LD with rs4646450 ($R^2 = 0.9304$) and in the same clinical cohort used in this study, rs4646450 was associated with decreased tacrolimus concentration, although with unknown molecular mechanism.[38] Thus, due to their high LD with rs776744/rs746742 in EAS, rs4646450 may instead have been acting as a marker for rs776744/rs746742 in the previous study. In support of this, rs4646450 does not associate with CYP3A5 mRNA level in our liver cohort ($p = 0.113$) after adjusting for CYP3A5*3, *6, and *7 genotypes and relevant transcription factors, where the LD between these two SNPs is much lower (EUR, $R^2 = 0.0752$; AFR, $R^2 = 0.1966$). Therefore, rs776744/rs746742 appears to be better suited to predict CYP3A5 expression across different populations, increasing its potential to facilitate clinical translation.

In summary, we have identified a novel shared distal enhancer region DRR controlling expression of several CYP3As. Within the DRR we discovered two genetic variants, rs115025140 and rs776744/rs746742, that are associated with increased expression of CYP3A4 and CYP3A5, respectively. In clinical cohorts, carriers of rs115025140

and rs776744/rs746742 have higher CYP3A activity, showing reduced statin-mediated lipid reduction and reduced tacrolimus concentration, respectively. If validated, rs115025140 has the potential to help predict CYP3A4 activity for AFR populations and rs776744/rs746742 may better explain inter-person variability of CYP3A5 in addition to the known CYP3A5 variants. Overall, our results elucidate the complex chromatin interaction landscape controlling expression of the CYP3A gene cluster. It is worth noting that, like the CYP3As, many pharmacogene families exist in clusters (e.g., CYP2Cs and CYP1As) and their potential for similar shared enhancer co-regulation is a compelling future direction.

AUTHOR CONTRIBUTIONS

J.M.C. and D.W. wrote the manuscript. J.M.C. and D.W. designed the research. J.M.C., A.C.N., S.J.M., L.L., CC.L., CJ.L., E.S., M.C., J.W.S., S.A.S., and D.W. performed the research. J.M.C., M.A.R., and D.W. analyzed the data.

ACKNOWLEDGMENTS

LentiCRISPR was provided by Dr. Feng Zhang (Broad Institute, Cambridge, Massachusetts, USA); pCMV-VSV-G was provided by Dr. Bob Weinburg (Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, USA), and psPAX2 was provided by Dr. Didier Trono (Ecole Polytechnique Federale De Lausanne, Switzerland).

FUNDING INFORMATION

This work was supported by the National Institute of Health (NIH) (R01 GM120396 and R35 GM140845 to D.W.; R01 MD011307 to S.A.S.). The content is solely the responsibility of the authors and does not represent the official view of the National Institute of Health. Liver samples were provided by the Cooperative Human Tissue Network, which is funded by the National Cancer Institute. Human hepatocytes were provided by the Liver Tissue Cell Distribution System (Pittsburg, Pennsylvania, USA), which is funded by NIH contract No. HHSN276201200017C.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

ORCID

Liang Li https://orcid.org/0000-0002-0687-074X
Danxin Wang https://orcid.org/0000-0002-7134-3065

REFERENCES

1. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther*. 2013;138:103-141.
2. Achour B, Barber J, Rostami-Hodjegan A. Expression of hepatic drug-metabolizing cytochrome p450 enzymes and their intercorrelations: a meta-analysis. *Drug Metab Dispos*. 2014;42:1349-1356.
3. Paine MF, Hart HL, Ludington SS, Haining RL, Rettie AE, Zeldin DC. The human intestinal cytochrome P450 "pie". *Drug Metab Dispos*. 2006;34:880-886.
4. Klomp SD, Manson ML, Guchelaar HJ, Swen JJ. Phenoconversion of cytochrome P450 metabolism: a systematic review. *J Clin Med*. 2020;9:2890.
5. Rahmioglu N, Heaton J, Clement G, et al. Genetic epidemiology of induced CYP3A4 activity. *Pharmacogenet Genomics*. 2011;21:642-651.
6. Ozdemir V, Kalow W, Tang BK, et al. Evaluation of the genetic component of variability in CYP3A4 activity: a repeated drug administration method. *Pharmacogenetics*. 2000;10:373-388.
7. Penno MB, Dvorchik BH, Vesell ES. Genetic variation in rates of antipyrine metabolite formation: a study in uninduced twins. *Proc Natl Acad Sci USA*. 1981;78:5193-5196.
8. Matthaei J, Bonat WH, Kerb R, et al. Inherited and acquired determinants of hepatic CYP3A activity in humans. *Front Genet*. 2020;11:944.
9. Wang D, Guo Y, Wrighton SA, Cooke GE, Sadee W. Intronic polymorphism in CYP3A4 affects hepatic expression and response to statin drugs. *Pharmacogenomics J*. 2011;11:274-286.
10. Kuehl P, Zhang J, Lin Y, et al. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat Genet*. 2001;27:383-391.
11. Lamba JK, Lin YS, Thummel K, et al. Common allelic variants of cytochrome P4503A4 and their prevalence in different populations. *Pharmacogenetics*. 2002;12:121-132.
12. Carithers LJ, Ardlie K, Barcus M, et al. A novel approach to high-quality postmortem tissue procurement: the GTEx project. *Biopreserv Biobank*. 2015;13:311-319.
13. Wang D, Sadee W. CYP3A4 intronic SNP rs35599367 (CYP3A4*22) alters RNA splicing. *Pharmacogenet Genomics*. 2016;26:40-43.
14. Collins JM, Wang D. Cis-acting regulatory elements regulating CYP3A4 transcription in human liver. *Pharmacogenet Genomics*. 2020;30:107-116.
15. Matsumura K, Saito T, Takahashi Y, et al. Identification of a novel polymorphic enhancer of the human CYP3A4 gene. *Mol Pharmacol*. 2004;65:326-334.
16. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57-74.
17. Rosenbloom KR, Sloan CA, Malladi VS, et al. ENCODE data in the UCSC genome browser: year 5 update. *Nucleic Acids Res*. 2013;41:D56-D63.
18. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12:996-1006.
19. FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest ARR, Kawaji H, et al. A promoter-level mammalian expression atlas. *Nature*. 2014;507:462-470.
20. Zheng R, Wan C, Mei S, et al. Cistrome data browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res*. 2019;47:D729-D735.
21. Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin*. 2014;7:33.
22. Smith RP, Eckalbar WL, Morrissey KM, et al. Genome-wide discovery of drug-dependent human liver regulatory elements. *PLoS Genet*. 2014;10:e1004648.

23. Wang D, Hartmann K, Seweryn M, Sadee W. Interactions between regulatory variants in CYP7A1 (cholesterol 7α-hydroxylase) promoter and enhancer regions regulate CYP7A1 expression. *Circ Genom Precis Med*. 2018;11:e002082.
24. Collins JM, Wang D. 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*. 2021;36:205-214.
25. Lu H, Jiang H, Yang S, et al. Trans-eQTLs of the CYP3A4 and CYP3A5 associated with tacrolimus trough blood concentration in Chinese renal transplant patients. *Biomed Pharmacother*. 2022;145:112407.
26. Goodwin B, Hodgson E, Liddle C. The orphan human pregnane X receptor mediates the transcriptional activation of CYP3A4 by rifampicin through a distal enhancer module. *Mol Pharmacol*. 1999;56:1329-1339.
27. Fulco CP, Munschauer M, Anyoha R, et al. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science*. 2016;354:769-773.
28. Lolodi O, Wang YM, Wright WC, Chen T. Differential regulation of CYP3A4 and CYP3A5 and its implication in drug discovery. *Curr Drug Metab*. 2017;18:1095-1105.
29. Wang D, Lu R, Rempala G, Sadee W. Ligand-free estrogen receptor α (ESR1) as master regulator for the expression of CYP3A4 and other cytochrome P450 enzymes in the human liver. *Mol Pharmacol*. 2019;96:430-440.
30. Palstra RJ, Grosveld F. Transcription factor binding at enhancers: shaping a genomic regulatory landscape in flux. *Front Genet*. 2012;3:195.
31. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2012;40:D930-D934.
32. Wen Z, Huang ZT, Zhang R, Peng C. ZNF143 is a regulator of chromatin loop. *Cell Biol Toxicol*. 2018;34:471-478.
33. Collins JM, Huo Z, Wang D. ESR1 ChIP-seq identifies distinct ligand-free ESR1 genomic binding sites in human hepatocytes and liver tissue. *Int J Mol Sci*. 2021;22:1461.
34. Inoue F, Kircher M, Martin B, et al. A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome Res*. 2017;27:38-52.
35. Thurman RE, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome. *Nature*. 2012;489:75-82.
36. Neuvonen PJ, Niemi M, Backman JT. Drug interactions with lipid-lowering drugs: mechanisms and clinical relevance. *Clin Pharmacol Ther*. 2006;80:565-581.
37. Vyhlidal CA, Pearce RE, Gaedigk R, et al. Variability in expression of CYP3A5 in human fetal liver. *Drug Metab Dispos*. 2015;43:1286-1293.
38. Chen D, Lu H, Sui W, et al. Functional CYP3A variants affecting tacrolimus trough blood concentrations in Chinese renal transplant recipients. *Pharmacogenomics J*. 2021;21:376-389.
39. Ben Zouari Y, Platania A, Molitor AM, Sexton T. 4See: a flexible browser to explore 4C data. *Front Genet*. 2019;10:1372.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Collins JM, Nworu AC, Mohammad SJ, et al. Regulatory variants in a novel distal enhancer regulate the expression of CYP3A4 and CYP3A5. *Clin Transl Sci*. 2022;15:2720-2731. doi:10.1111/cts.13398