

MOL # 114439

Title Page

Coordinated Transcriptional Regulation of Cytochrome P450 3As by Nuclear Transcription Factor Y (NF-Y) and Specificity Protein 1 (Sp1)

Ruohong Chen, Jun Jiang, Zhangsheng Hu, Wenchu Ye, Qianqian Yuan, Mengyuan Li, Jikai Wen* and Yiqun Deng*

Guangdong Provincial Key Laboratory of Protein Function and Regulation in Agricultural Organisms, College of Life Sciences, South China Agricultural University, Tianhe District, Guangzhou, Guangdong 510642, P.R. China.

Key Laboratory of Zoonosis of Ministry of Agriculture and Rural Affairs, South China Agricultural University, Guangzhou, Guangdong, 510642, P.R. China.

MOL # 114439

Running Title Page

Running title: The conserved basal regulated mechanism of CYP3A.

Corresponding author:

* Prof. Yiqun Deng

Guangdong Provincial Key Laboratory of Protein Function and Regulation in Agricultural Organisms, College of Life Sciences, South China Agricultural University, Guangzhou, Guangdong, China.

Key Laboratory of Zoonosis of Ministry of Agriculture and Rural Affairs, South China Agricultural University, Guangzhou, Guangdong, China.

Phone: +86-20-38604967, Fax: +86-20-38604967.

Email: yqdeng@scau.edu.cn (Prof. Yiqun Deng)

Number of text pages: 47

Number of tables: 0

Number of figures: 6

Number of references: 41

Number of words: Abstract 191

Introduction 793

Discussion 1300

MOL # 114439

Abbreviations used

CCAR1, cell cycle and apoptosis regulator 1; C/EBP α , CCAAT enhancer binding protein alpha; CAR, constitutive androstane receptor; CYP, cytochrome P450; DAPA, DNA affinity precipitation assay; EMSA, Electrophoretic mobility shift assay; HNF1 α , hepatocyte nuclear factor 1 alpha; HNF3, hepatocyte nuclear factor 3; mut-Sp1, Sp1 binding site mutation, mut-NF-Y, the NF-Y binding site mutation; NF-Y, nuclear transcription factor Y; NF-YA, nuclear transcription factor Y subunit alpha; NF-YB, nuclear transcription factor Y subunit beta; NF-YC, nuclear transcription factor Y subunit gamma; PXR, pregnane X receptor; Sp1, Specificity Protein 1; Sp3, Specificity Protein 3; SV40, simian vacuolating virus 40; WT, wild type;.

MOL # 114439

Abstract

The cytochrome P450 (CYP) 3A subfamily plays vital roles in the metabolism of endogenous chemicals and xenobiotics. Understanding the basal expression of CYP3As in humans and pigs is crucial for drug evaluation. In this study, we demonstrated that the basal transcriptional regulation of *CYP3As* in hepatocytes is evolutionarily conserved between humans and pigs. The basal expression of *CYP3A* genes is transactivated by two *cis*-acting elements, the CCAAT box and GC box, located a constant distance apart in the proximal promoter region of six *CYP3A* genes. Mutation analysis of these two *cis*-acting elements suggested that they play important roles in mediating basal expression, but to different extents, because of the nucleotide variations in the elements. Two transcription factors, NF-Y and Sp1, directly bind to these *cis*-acting elements in *CYP3A* proximal promoters in HepG2 cells and porcine hepatocytes. Furthermore, changing the distance between the NF-Y and Sp1 binding sites resulted in decreases in the promoter activity of *CYP3As*. Conclusively, our results show that human and porcine *CYP3A* genes are regulated by NF-Y and Sp1 in a coordinated manner, and that the distance between these two *cis*-acting elements is crucial for constitutive *CYP3A* expression.

MOL # 114439

1. Introduction

Cytochrome P450s belong to a heme monooxygenase superfamily named according to their characteristic absorption at 450 nm (Omura and Sato, 1962). P450 genes are categorized into different families and subfamilies based on the divergent evolution and amino acid sequence homology of this superfamily (Nebert et al., 1991). CYP1-4 family members play crucial roles in the metabolism of endogenous substances, including sterols, fatty acids, and vitamins, and the biotransformation of exogenous compounds, including drugs, environmental pollutants, and toxins (Isin and Guengerich, 2007; Zanger and Schwab, 2013). In particular, approximately 50-60% of known clinical drugs can be metabolized by the human CYP3A4 enzyme (Kumar and Surapaneni, 2001; Li et al., 1995).

CYP3A4 and CYP3A5 are mainly expressed in the liver and gastrointestinal tract (Guengerich, 1999; Kivisto et al., 1996). In the liver and hepatocytes, the basal expression of *CYP3A4* is regulated by many transcription factors, such as the binding of CCAAT enhancer binding protein alpha (C/EBP α), hepatocyte nuclear factor 3 (HNF3), and pregnane X receptor (PXR) to the proximal promoter and distal regions of *CYP3A4* (Bombail et al., 2004; Matsumura et al., 2004; Rodriguez-Antona et al., 2003). Cross talk between transcription factors has been extensively studied. It has been reported that the interaction of cell cycle and apoptosis regulator 1 (CCAR1) with the constitutive androstane receptor (CAR) enhances the expression of CAR-induced genes, such as *CYP3A4* (Kanno et al., 2018). Proximal PXR, CAR, NF-Y, Sp1, and Sp3 cooperatively regulate the basal transcription of *CYP3A5* (Burk et al., 2004; Iwano et al., 2001).

MOL # 114439

Due to their physiological and anatomical similarity to humans, pigs represent an important animal model for drug evaluations in humans (Liu et al., 2008). Our previous report suggested that porcine CYP3A46 is related to human CYP3A4 based on the high amino acid sequence identity and similar oxidation activity (Jiang et al., 2011). In addition, our previous studies suggested that in the proximal promoter region, the GC box transactivates the expression of *CYP3A46* and that both the GC box and CCAAT box determine the basal transactivation of *CYP3A22* in porcine primary hepatocytes (Dong et al., 2015; Liu et al., 2016). This regulation of porcine *CYP3A22* is similar to the regulation of the basal transactivation of human *CYP3A5* (Iwano et al., 2001). This regulatory pattern remains to be addressed in the basal regulation of other *CYP3A* genes, including porcine *CYP3A29*, *CYP3A46*, and human *CYP3A4* (Rodriguez-Antona et al., 2003; Tirona et al., 2003).

Sp1 is the first identified promoter-specific factor required for the transcription of the simian vacuolating virus 40 (SV40) genome (Dyanan and Tjian, 1983). Sp1 binds to GC boxes through the three conserved C2H2-type zinc fingers in the C-terminus. NF-Y is a trimeric protein formed by NF-YA, NF-YB, and NF-YC, and exhibits sequence-specific binding to CCAAT boxes (Maity and de Crombrugghe, 1998). A genome-wide study indicated that CCAAT boxes are mainly located within a conserved distance of -150 to +50 bp from the transcription start sites in 18% of the 21,000 human promoters (Benner et al., 2013). Structural analysis revealed that NF-Y bound DNA in a histone-like mode (Ceribelli et al., 2008). Coimmunoprecipitation suggested the direct interaction between Sp1 and NF-Y in human cells (Lim and Chang, 2009; Roder et al., 1999). In the human genome, CCAAT boxes

MOL # 114439

and GC boxes are observed to have high tendency to co-occur in regions -150/+50 bp from transcription start sites (Benner et al., 2013). In addition, NF-Y has genome-wide partnerships with Sp1, and the distances between the *cis*-acting elements are conserved (Dolfini et al., 2016).

In this study, we aimed to determine whether porcine *CYP3A* genes have a regulatory mechanism of *CYP3A4* transactivation similar to that in human cells. We found that both NF-Y and Sp1 can bind the proximal CCAAT box and GC box in the promoters of *CYP3A29* and *CYP3A46* as previously reported for *CYP3A22* (Liu et al., 2016). However, in contrast to the regulation of *CYP3A22*, Sp1 is the main regulator of *CYP3A29* and *CYP3A46* basal expression, while NF-Y minimally contributes to the basal expression. A base-pair mismatch (ACTGG) in the CCAAT box is considered as a main cause of this discrepancy in the regulation of basal transcription. Then, we found similar *cis*-acting elements in the proximal promoters of *CYP3A4* and *CYP3A5*. A series of experiments showed that the corresponding proximal CCAAT box and GC box regulated the expression of *CYP3A4* and *CYP3A5*. This transcription model is similar to the model summarized in porcine *CYP3As* and unveils a conserved transcriptional activity pattern between human and porcine *CYP3A* isoforms. Changing the distance between the NF-Y and Sp1 binding sites repressed the transactivation activity of the proximal promoters in both human and porcine *CYP3As*, and the decreased promoter activity suggested that NF-Y and Sp1 coordinated the constitutive expression of human and porcine *CYP3As* in a spatially coordinated manner.

MOL # 114439

2. Materials and Methods

Ethics statement. All related experiments were performed in strict accordance with the regulations of the Administration of Affairs Concerning Experimental Animals of Guangdong Province, China, under approval number 2017-D007. All surgeries were performed under anesthesia, and all efforts were made to minimize suffering.

Animals and cell culture. Hepatocytes were isolated from a 9-day-old Danish Landrace × Yorkshire × Duroc cross-breed male pig purchased from the College of Veterinary Medicine, South China Agricultural University. A modified two-step *in situ* collagenase perfusion procedure was performed as previously described (Liu et al., 2016). Briefly, hepatocytes were isolated by collagenase digestion. After filtering through a cell strainer, the fresh hepatocytes were centrifuged and resuspended in William's E medium (Sigma-Aldrich; Merck KGaA, Darmstadt, Germany) supplemented with 10% fetal bovine serum (FBS; PAN-Biotech, Aidenbach, Germany), 1% penicillin G/streptomycin (Life Technologies, Inc., Invitrogen, CA, USA), $1 \times \text{SITE}$ (Sigma), $0.1 \mu\text{M}$ dexamethasone and $0.1 \mu\text{M}$ insulin (Invitrogen). The hepatocytes were seeded on a plate at a density of 1×10^5 cells/cm² and cultured in a humidified 37°C incubator with 5% CO₂. After 4 h, the medium was removed, and the adherent cells were maintained in William's E medium supplemented with 10% FBS and 1% penicillin G/streptomycin.

HepG2 cells (HB-8065, American Type Culture Collection, Manassas, VA, USA) and COS-7 cells (CRL-1651, ATCC) were cultured in Dulbecco's modified Eagle's medium

MOL # 114439

(DMEM) supplemented with 10% FBS and 1% penicillin G/streptomycin antibiotics (Invitrogen) at 37°C and 5% CO₂. Huh-7, QGY-7703, and QSG-7701 cells were kindly provided by Professor. Zhu Xiaofeng (State Key Laboratory of Oncology in South China, Sun Yat-Sen University Cancer Center, China). The HepaRG cell line was purchased from Shanghai Guandao Biological Engineering Co., Ltd. (Shanghai, China). All the hepatocytes were cultured in DMEM/F-12 supplemented with 10% FBS and 1% penicillin G/streptomycin antibiotics (Life Technologies, Inc., Invitrogen, CA, USA) at 37°C and 5% CO₂.

Plasmid constructs. Based on NCBI GenBank, we amplified the 5'-flanking regions from -4349 to -1 and the -707 to -1, -364 to -1, -244 to -1, -157 to -1, -131 to -1, and -120 to -1 deletion constructs (the translational start site was designated +1) of porcine *CYP3A29* (GenBank Accession Number: NC_010445.4) by PCR. Porcine genomic DNA from liver tissue was used as a template. The *Mlu* I (TaKaRa, Qingdao, China, #1071A) and *Xho* I (Takara, 1094A) digested PCR fragments were inserted upstream of pGL3-Basic (Promega Corp., Madison, WI, USA) to generate the luciferase reporter constructs. The 5'-flanking regions of *CYP3A5* (-180 to -1) and *CYP3A4* (-181 to -1) were cloned from genomic DNA from HepG2 cells. Then, the DNA fragments were inserted into the *Mlu* I/*Xho* I sites of pGL3-Basic (Promega). To further analyze the transcriptional activation of *cis*-acting elements, variant luciferase reporter constructs were generated using a QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). The NF-YA and Sp1 overexpression plasmids were generated as previously reported (Liu et al., 2016). All primer sequences are shown in Table S1.

MOL # 114439

KOD-FX DNA polymerase (Toyobo, Osaka, Japan) was used for the polymerase chain reaction. The plasmids were extracted with a TIANprep Mini Plasmid Kit II (TIANGEN, Beijing, China) and verified by DNA sequencing (Sangon Biotech, Shanghai, China). All corresponding constructs were prepared using an Endo-Free Plasmid Kit (Omega, Norcross, GA, USA) for transfection. The extraction and purification of the plasmids were performed according to the manufacturer's instructions. T4 DNA ligase was purchased from Takara (2011A).

Transient transfection and luciferase activity detection. Porcine primary hepatocytes were seeded into 24-well plates and cultured in DMEM to 80% confluence. Opti-MEM (Life Technologies) was used as the transfection reagent. The cells in each well of the 24-well plates were transfected with 0.4 μ g of the firefly luciferase reporter construct, 0.2 μ g of the *Renilla* luciferase reporter construct (pRL-TK) and 1.5 μ L of Lipofectamine 3000 (Invitrogen). HepG2 cells were seeded and cultured in 24-well plates in DMEM. For the *CYP3A4* and *CYP3A5* reporter constructs, 0.6 μ g of the firefly luciferase reporter plasmids and 60 ng of the *Renilla* luciferase reporter plasmid (pRL-TK) were incubated with 1.5 μ L of Lipofectamine 3000 (Invitrogen). The transfected cells were lysed after 24 h of incubation, and luciferase activity was measured with a dual luciferase reporter assay system (Promega) according to the manufacturer's instructions. The promoter transcription strength was represented by the relative luciferase activity, and the firefly luciferase activity for each construct was normalized to the *Renilla* luciferase activity as a transfection reference. In this quantitative

MOL # 114439

analysis, the fold changes in the activity of each construct were calculated relative to the activity of the empty vector (Basic). Each assay was performed in triplicate, and the obtained results represent those of three independent experiments.

To overexpress NF-YA/Sp1, COS-7 cells were seeded and cultured in 10 cm plates until reaching 80% confluence. Then, the cells were transfected with mixtures containing 6 μ g of the NF-YA/Sp1 overexpression vector. After 24 h of transfection, the cells were harvested, the protein level was determined, and the nuclear fractions were isolated.

For the RNA interference, siRNA targeting Sp1, NF-YA and the unspecific siRNA control were designed and synthesized by GenePharma (Suzhou, China). In total, 40 pmol of siRNA were transiently transfected into HepG2/porcine primary hepatocytes (a well in 12-well plates) using Lipofectamine 3000 (Invitrogen). After 24 h of transfection, the cells were lysed and subsequently analyzed by western blotting.

Electrophoretic mobility shift assays (EMSA). The nuclear fractions from NF-YA/Sp1 overexpressed COS-7 or HepG2 cells were isolated using a Nuclear Extraction Kit (#P0027, Beyotime, Haimen, China). Then, 5' biotin labeled probes corresponding to putative NF-Y or Sp1 binding sites of porcine *CYP3A29* and human *CYP3A4* and *CYP3A5* were prepared using an EMSA Probe Biotin Labeling Kit (Beyotime, #GS008). EMSA was performed with an EMSA Assay Kit (Beyotime, #GS008) according to a previously described method (Dong et al., 2015; Yang et al., 2014). Briefly, 5 μ g nuclear extract protein from NF-YA/Sp1 overexpressed COS-7 cells or HepG2 cells were pre-incubated in binding buffer for 10 min at

MOL # 114439

20°C. For the specific and nonspecific competition assays, 100-fold unlabeled wildtype/mutant probes were added. For the super-shift assay, 1 µg anti-NF-YA antibody (sc-17753, Santa Cruz Biotechnology, Santa Cruz, CA, USA) or anti-Sp1 (ab13370, Abcam, Cambridge, MA, USA) was added. Then, 0.1 pmol 5'-biotin-labeled double stranded probe were added in a total volume of 10 µL for further incubation for 20 min at 20°C.

RNA isolation and real-time PCR. Total RNA was isolated from HepG2 and porcine primary hepatocytes using TRIzol reagent (Invitrogen) and reverse transcribed to cDNA using M-MLV reverse transcriptase (Promega) as previously described (Jiang et al., 2018). Real-time PCR was performed on a CFX Connect™ Real-Time System (Bio-Rad, Hercules, CA, USA). Reactions were performed using GoTaq qPCR Master Mix (Promega) in 20 µL volumes according to the manufacturer's recommendations. The primers used were as follows:

Porcine GAPDH-F: 5'-GTCGGTTGTGGACCTGAC-3';

Porcine GAPDH-R: 5'-TGGTCGTTGAGGGCAATG-3';

Porcine *CYP3A22*-F: 5'-TACCGTAAGGGCATGTGGCA-3';

Porcine *CYP3A22*-R: 5'-GGTTGGAGACAGCAATGTTCGT-3';

Porcine *CYP3A29*-F: 5'-CCTGAAATTAACCACGCAAGGGCT-3';

Porcine *CYP3A29*-R: 5'-TCTGGGATGCAGCTTTCTTGACCA-3';

Porcine *CYP3A46*-F: 5'-CGAATCTCTCAGAATATTCCCAGT-3';

Porcine *CYP3A46*-R: 5'-GCGCAAAGATTGGCACCATCAT-3';

MOL # 114439

Human GAPDH-F: 5'-CCTCAAGATCATCAGCAATGCC-3';

Human GAPDH-R: 5'-TGGTCATGAGTCCTTCCACGAT-3';

Human *CYP3A4*-F: 5'-GCCTGGTGCTCCTCTATCTA-3';

Human *CYP3A4*-R: 5'-GGCTGTTGACCATCATAAAAG-3';

Human *CYP3A5*-F: 5'-TGACCCAAAGTACTGGACAG -3';

Human *CYP3A5*-R: 5'-TGAAGAAGTCCTTGCGTGTC -3';

Western blotting. The cells were lysed in RIPA buffer as previously described (Ma et al., 2017). The protein expression levels were determined via western blotting as previously described (Liu et al., 2016). The antibodies used in this work included anti-GAPDH antibody (Santa Cruz, sc-32233), porcine CYP3A antibody (mouse polyclonal antibody, lab stock), anti-CYP3A4 antibody (Absin, abs136521), anti-CYP3A5 (Absin, abs137509), anti-NF- κ B antibody (Santa Cruz, sc-17753), and anti-Sp1 antibody (Abcam, ab13370).

DNA affinity precipitation assay (DAPA). The 5' biotinylated probes were synthesized by PCR amplification of the *CYP3A22/CYP3A29* luciferase reporter plasmids with the following primers: F: 5' biotin-AGCAGGTGGAGCACAGGGG-3' and R: 5' biotin-CTGGAGAGTCAC TGGGGCT-3' for *CYP3A22* and F: 5' biotin-GTGATGACTCGCTGAGGCT-3' and R: 5' biotin-CTGCAGGAAGAGCACTGGT-3' for *CYP3A29*. The assay was performed by incubating 0.5 μ g of biotin-labeled probe with 50 μ g of nuclear extract in 400 μ L binding

MOL # 114439

buffer containing 20 mM HEPES, pH 7.9, 10% glycerol, 50 mM KCl, 0.2 mM EDTA, 1.5 mM MgCl₂, 10 μM ZnCl₂, 1 mM dithiothreitol, 0.25% Triton-X100 and 1 x protease inhibitor (Biotool, Huston, USA) on ice. After 45 min, the DNA/protein complexes were incubated with 40 μL pre-equilibrated streptavidin-coated magnetic beads (from mTRAPtm Maxi) (Active motif, Carlsbad, CA, USA) for 1 h with rotation at 4°C. Then, the magnetic beads were pelleted by placing the tubes in a Magnetic Separation Rack, and 1 mL of binding buffer was mixed with the beads, rotated for 5 min, and washed 5 times to remove nonspecific binding. Finally, the binding proteins were eluted with 2 x SDS sample buffer with boiling for 5 min. Then, the protein samples were separated by SDS-PAGE and analyzed by western blotting with Sp1 or NF-Y antibody to confirm protein binding to the applied probes.

Chromatin immunoprecipitation assay (ChIP). The ChIP assays were performed according to the recommendations of the SimpleChIP® Enzymatic Chromatin IP Kit (#9003, Cell Signaling Technology Inc., Danvers, USA). In brief, the porcine primary hepatocytes were cross-linked via the addition of 1% final formaldehyde concentration at room temperature for 10 min. The addition of micrococcal nuclease and sonication were used to fragment the cross-linked chromatin to a length of approximately 150-900 bp. In total, 10 μg of digested chromatin were used for each immunoprecipitation. The cross-linked DNA-protein complex was immunoprecipitated with 1 μg of antibodies and protein G magnetic beads (Cell Signaling Technology, #9006). The anti-H3 antibody (Cell Signaling Technology, #4620) and normal rabbit IgG (Cell Signaling Technology, #2729) were added as a positive and negative

MOL # 114439

control, respectively, and chromatin was subjected to immunocomplexes with the anti-NF- κ B (Santa Cruz, sc-17753) or anti-Sp1 (Abcam, ab13370) antibodies. The DNA was eluted, the cross-linking was reversed, and the DNA was purified according to the protocol. Then, the immunoprecipitated DNA was used as a template for real-time PCR with special primers. The primers used for the design of the proximal NF- κ B and Sp1 binding sites of the *CYP3A29* promoter were F: 5'-GTCCATGGTGCTTTGCCAACTA-3' and R: 5'-CGCTTTGCCTGCTGCTTTCAC-3'. For *CYP3A4* and *CYP3A5*, real-time PCR was performed with the following primers: *CYP3A4* (F 5'-CAGTGAGGCTGTTGGATTGT-3'; R 5'-TGTGTGATTCTTTGCCAACTTC-3') and *CYP3A5* (F 5'-TGGACTCCCCGATAACACTGAT-3'; R 5'-AGCCAAGCTGCTGAAAGATTTA-3').

Statistical analysis. All statistical procedures were analyzed using SPSS 16.0 (IBM, USA). All figures were plotted using GraphPad Prism 5.0 (USA). Statistically significant differences among more than two groups were determined using one-way analysis of variance (ANOVA) or two-way ANOVA followed by Dunnett's multiple comparison tests. For two groups independent samples Student's t-tests (two-tailed) were used. The correlations between two factors were evaluated by Pearson correlation analysis. The data are presented as the means \pm SD of at least three independent experiments. And significance was defined as *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

MOL # 114439

3. Results

The conserved proximal CCAAT boxes and GC boxes in the promoters of porcine and human CYP3As affected the mRNA expression of *CYP3A* genes. To investigate the mechanism underlying the regulation of *CYP3A* genes basal transcription in both porcine and human hepatocytes, sequence alignment was performed to analyze the conservation of *cis*-acting elements in the proximal promoters of different *CYP3A* isoforms. As shown in Fig. 1A, two *cis*-acting elements, i.e., a CCAAT box and GC box, were identified in the proximal promoter region of six *CYP3A* genes in both the human and porcine genome. As indicated, the *cis*-acting elements in the proximal promoter sequence of porcine *CYP3A22* were similar to those in human *CYP3A5*, and the *cis*-acting elements in the porcine *CYP3A29* and *CYP3A46* promoters are similar to those in human *CYP3A4* and *CYP3A7*. In the proximal promoters of human *CYP3A4* and *CYP3A7* and porcine *CYP3A29* and *CYP3A46*, the consensus ACTGG sequence in the CCAAT box is an inverted CCAAT box with a one base-pair mismatch.

To demonstrate the binding specificity of the transcription factors NF-Y and Sp1 to the proximal promoters of porcine *CYP3A* genes, a DAPA was used to verify the ability of these transcription factors to bind promoter DNA *in vitro*. As indicated in Fig. 1A, the sequences of the proximal promoters of *CYP3A29* and *CYP3A46* were nearly identical, suggesting that the *CYP3A29* probe was representative of both *CYP3A29* and *CYP3A46*. By incubating the nuclear fractions of porcine primary hepatocytes with the 5' biotinylated promoter oligonucleotide of *CYP3A29/CYP3A46* and *CYP3A22*, we demonstrated via the DAPA that the endogenous transcription factors NF-Y and Sp1 directly bound the labeled promoter DNA

MOL # 114439

of *CYP3A22* and *CYP3A29/CYP3A46* (Fig. 1B). Relative to *CYP3A22* promoter DNA binding, Sp1 was almost equally precipitated by the labeled *CYP3A29* probe, but the binding of NF-YA to this probe was reduced. This finding suggests that the GC boxes of these two *cis*-acting elements likely play a vital role in the basal transactivation of genes in the porcine and human *CYP3A* families. The ability of Sp1 and NF-Y to regulate porcine *CYP3A* expression *in vivo* was tested using siRNA targeting Sp1 and NF-YA, respectively. siRNA targeting the coding region of Sp1 and NF-YA was transfected into porcine primary hepatocytes, HepG2 (human hepatocellular carcinoma) cells, QSG-7701 cells (normal hepatocytes), and HepaRG cells (a human hepatocellular carcinoma-derived cell line with high CYP450 expression). Consequently, the mRNA levels of porcine and human CYP3As in si-Sp1 cells were decreased (Fig. 1C, and 1D, Supplementary Fig. 1A, and 1B). In addition, the mRNA levels of *CYP3A4* and *CYP3A5* were increased in cells overexpressing Sp1 (Supplementary Fig. 2). We further analyzed the efficiency in knocking down protein expression, as shown in Supplementary Fig. 3. Taken together, these results indicate that knockdown of Sp1 and NF-YA clearly decreased the mRNA expression levels of porcine and human CYP3A genes.

NF-Y and Sp1 bind the CCAAT box and GC box, respectively, in the proximal core transactivation region of *CYP3A29/CYP3A46*. We generated a series of 5' deletion firefly luciferase reporter constructs using the pGL3-Basic vector to determine whether this proximal promoter is a core regulatory region in porcine *CYP3A29* and *CYP3A46*. The promoter

MOL # 114439

sequence was identical to the *CYP3A29* genome sequence (accession no. NC_010445.4). These reporters were individually transfected into porcine primary hepatocytes with the reference plasmid pRL-TK containing the *Renilla* luciferase gene. In this study, the start codon in the *CYP3A29* ORF was defined as 0. As shown in Fig. 2A, the promoter region -244 to -1 had the strongest transactivation activity among all constructs, as follows: 31-fold higher than that of the control (pGL3-Basic transfected group). In contrast with the -244 to -1 region, the region from -4349 to -244 did not strongly contribute to basal transactivation and led to a 1.4-fold decrease (Fig. 2A). This finding suggests that consistent with *CYP3A22* and *CYP3A46*, the proximal DNA region of the *CYP3A29/CYP3A46* promoter shows the highest promoter activity. More interestingly, the relative luciferase activity gradually decreased from 31-fold (-244 to -1), 21-fold (-157 to -1), and 11-fold (-131 to -1) to 1.6-fold (-120 to -1) compared with that in the control (Fig. 2A). These results strongly suggest that the minimal and essential region responsible for the basal transactivation of *CYP3A29/CYP3A46* is a 36-bp region from -157 to -120 and that some *cis*-acting elements are likely located in this region (Fig. 2A). As indicated in the sequence alignment shown in Fig. 1A, these *cis*-acting elements are likely the binding sites of NF-Y and Sp1.

An EMSA and a DAPA were used to further demonstrate the binding specificity of the transcription factors NF-Y and Sp1 to the *CYP3A29/CYP3A46* proximal promoter *in vitro*. As shown in Fig. 2B, two shifted bands formed in the Sp1-overexpressing COS-7 nuclear fractions incubated with the 5' biotin-labeled *CYP3A29/CYP3A46* probe (Sp1-WT). The addition of a 100-fold excess of unlabeled WT probe but not the cold mutant probe (Sp1-mut)

MOL # 114439

competed out the shifted bands (lanes 3 and 4, Fig. 2B). The shifted bands were slowed by protein specifically binding the GC box in the WT *CYP3A29/CYP3A46* probe. The DNA-protein complex was super-shifted following incubation with anti-Sp1 antibody, and the density of the upper binding band was markedly reduced (lane 5, Fig. 2B). Thus, Sp1 specifically bound the proximal promoter of *CYP3A29/CYP3A46*, and the upper shifted band was formed by Sp1 binding. The DAPA was performed using wildtype oligonucleotides for the *CYP3A29/CYP3A46* promoter and mutated oligonucleotides for Sp1 or NF-Y to further confirm the binding specificity of Sp1 and NF-Y to the *CYP3A29/CYP3A46* proximal promoter region (Fig. 2C). Moreover, the ChIP assays demonstrated that NF-YA and Sp1 were abundantly recruited to the proximal promoter region from nucleotides -212 to -54 of *CYP3A29/CYP3A46* *in vivo* (Fig. 2D).

Sp1 and NF-Y directly transactivate the transcription of *CYP3A29/CYP3A46*. As shown in Fig. 1A, certain proximal promoter regions in *CYP3A29* and *CYP3A46* have only one mismatch that is unrelated to the two *cis*-acting elements. We hypothesized that porcine *CYP3As* have a unified mechanism that regulates basal transcription and that the CCAAT boxes in *CYP3A29/CYP3A46* and *CYP3A22* make various contributions to transcription. To investigate the activation capability of the GC box and CCAAT box, we used site-directed mutagenesis to mutate or remove these binding sites to disrupt the binding of these transcription factors to the promoter DNA. The mutation and deletion sites are underlined in the top panel of Fig. 3A. Both the mutation and deletion of the Sp1 binding site resulted in a

MOL # 114439

90% decrease in luciferase activity compared with that driven by the wildtype promoter. However, relative to the activity driver by the CCAAT box in *CYP3A22*, the change in luciferase activity was not dramatic, but a decrease was observed when the NF-Y binding site was mutated (Fig. 3A). These data suggest that the proximal GC box dominates the basal transactivation of porcine *CYP3A29* and *CYP3A46*. The difference between *CYP3A29/CYP3A46* and *CYP3A22* may be due to the single-base-pair mutation in the CCAAT box of the *CYP3A29/CYP3A46* proximal promoter.

To explore whether this variation in the NF-Y binding site resulted in different effects on the transcriptional activation of *CYP3A22* and *CYP3A29*, we replaced the NF-Y and Sp1 binding sites for *CYP3A29* with the binding site for *CYP3A22* in the basal regulatory region of *CYP3A29* and *CYP3A46*. The mutated/exchanged region is underlined in the top panel of Fig. 3B. Mutation of the NF-Y binding site from ACTGG to ATTGG, similar to the *cis*-acting element in the *CYP3A22* proximal promoter, caused a 17% increase in luciferase activity. Unlike the wildtype promoter, the Sp1 binding site contained a CCACCT sequence (the GC box sequence of *CYP3A22*) rather than a CCACCC sequence (the GC box sequence of *CYP3A29*), resulting in a 37% decrease in *CYP3A29* promoter activity. Similarly, in the proximal promoter of *CYP3A46*, the replacement of *cis*-acting elements from the *CYP3A22* proximal promoter resulted in a 45% decrease in activity for the GC box, and a 38% increase in activity for the CCAAT box. As Supplementary Fig. 4 indicates, we replaced the NF-YA and Sp1 binding sites in *CYP3A22* with the binding sites in *CYP3A29/CYP3A46*, and the promoter activity of *CYP3A22* decreased for mut 29NF and increased for mut 29Sp1. These

MOL # 114439

data indicate that Sp1 and NF-Y play important roles in porcine *CYP3A* basal expression, and the variations in the *cis*-acting elements altered their individual contributions to the basal transcription levels.

Both NF-Y and Sp1 activate *CYP3A5* and *CYP3A4* transcription by binding the proximal CCAAT box and GC box. The proximal promoter regions of six *CYP3A* genes were highly conserved between humans and pigs (Fig. 1A). These six *CYP3As* contained an inverted CCAAT box and one proximal GC box. To determine whether these two conserved motifs truly activate human *CYP3A5* and *CYP3A4*, we amplified the proximal 5' upstream sequence in human *CYP3A5* (-180 to -1) and *CYP3A4* (-181 to -1). These regions were similar to the -157 to -1 region in *CYP3A29*, and the genomic sequence was confirmed (accession no. NC_000007.14). Based on the proximal sequences, we generated a series of reporter constructs with mutations or deletions of these two motifs. These luciferase reporters were transfected into HepG2 cells individually with the reference plasmid pRL-TK. As shown in Fig. 4A, the relative luciferase activity decreased by 17-fold for the constructs with deletion of either the CCAAT box, or GC box in the *CYP3A5* proximal promoter compared with that driven by the constructs containing the region from -180 to -1. For the constructs with mutation of a *cis*-acting element, the decrease in luciferase activity was similar to the decrease in the deletion constructs. These reductions in luciferase activity when each *cis*-acting element was dysfunctional demonstrated that both the CCAAT box and GC box are vital for the basal transactivation of *CYP3A5*. These results are consistent with a previous report of the

MOL # 114439

molecular mechanism underlying basal *CYP3A5* expression (Iwano et al., 2001).

Another Sp1 binding site was previously identified in the proximal promoter of *CYP3A4*, and this site was deduced to be irrelevant to the promoter activity of *CYP3A4* (Bombail et al., 2004). However, according to the sequence alignment shown in Fig. 1A, the previously reported Sp1 site is 32 bp upstream of our identified Sp1 binding site. Whether this conserved CCAAT box/GC box is involved in the transactivation of *CYP3A4* remains unclear. To elucidate the role of the NF-Y and Sp1 binding sites in the transcriptional activation of human *CYP3A4*, we performed a dual luciferase assay with the construct containing the promoter region from -181 to -1 and constructs containing binding site mutations. As shown in Fig. 4B, the relative luciferase activity was reduced by 54% and 60% in HepG2 cells with NF-Y binding site deletion and mutation, respectively, compared with that driven by the wildtype promoter. The deletion and mutation of the Sp1 binding site resulted in luciferase activity decreases of 53% and 77%, respectively. Due to the low expression level of *CYP3A4* and *CYP3A5* in HepG2 cells, we performed luciferase assays in Huh-7, QGY-7703, QSG-7701, and HepaRG cells (Supplementary Fig. 5), and obtained the same results. These results suggest that the conserved CCAAT box and GC box are important elements responsible for the proximal transactivation of human *CYP3A4*.

NF-Y and Sp1 bind the proximal CCAAT box and GC box, respectively, of *CYP3A5/CYP3A4*. The deletion and mutation constructs demonstrated that the CCAAT box and GC box in the proximal promoters of *CYP3A4* and *3A5* were involved in the basal

MOL # 114439

transactivation of *CYP3A4* and *3A5*. To further confirm that NF-Y and Sp1 activate transcription by directly binding the CCAAT box and GC box, respectively, the EMSA assay were performed using nuclear protein prepared from HepG2 cells. The sequences of the probes are shown in the top panel of Fig. 5A. As shown in the bottom panel of Fig. 5A lane2, a shift band was detected when the HepG2 nuclear fractions were incubated with the 5' biotin labeled wildtype probe (corresponding to the CCAAT box in the proximity of *CYP3A5*). The protein/DNA complex could be completed by the addition of 100-fold unlabeled wildtype probe, but not by the addition of 100-fold unlabeled binding site mutant probe. This result suggests that the protein/DNA binding complex was formed through specific binding (Fig. 5A lane 3 and lane 4). The addition of the NF-YA antibody resulted in a shift band that was super-shifted (Fig. 5A, lane 5). Our results reveal that NF-Y directly binds the CCAAT box to activate the transcription of *CYP3A5*. Meanwhile, the EMSA was performed to confirm that Sp1 binds the GC box. The sequence of the 5' biotin-labeled probe corresponded to the proximal GC box of *CYP3A5*. The labeled probe was incubated with nuclear fractions from HepG2, and two shift bands were detected in the gel (Fig. 5B lane 2). The addition of 100-fold unlabeled probe and unlabeled mutant probe further confirmed that the shift bands were specifically formed by DNA/protein binding (Fig. 5B lane 3 and lane 4). The upper band was abolished when the complex was super-shifted by the Sp1 antibody, confirming that the upper band was formed by the specific binding of Sp1 to the GC box. Taken together, NF-Y and Sp1 are directly associated with the two motifs identified in the proximal promoter of *CYP3A5 in vitro*.

MOL # 114439

To determine whether Sp1 is recruited to the proximal GC box of *CYP3A4*, an EMSA assay was performed with nuclear protein from HepG2 cells and a probe corresponding to the GC box of *CYP3A4*. After incubation, two shift bands of DNA/protein complexes were detected (Fig. 5C lane2). The upper band was super-shifted against the Sp1 antibody (Fig. 5C lane5). As shown in Fig. 5C lane 3, the shift bands could be competed with the addition of 100-fold unlabeled wildtype probe containing the Sp1 binding site. The addition of 100-fold unlabeled binding site mutant probe did not affect the formation of the shift band. This finding further confirms that Sp1 directly and specifically binds the proximal GC box of *CYP3A4* *in vitro*.

We corroborated the direct binding *in vivo* through a ChIP assay with anti-NF-YA and anti-Sp1 antibodies. As shown in Fig. 5D, the NF-Y and Sp1 binding is enriched at the *CYP3A4* and *CYP3A5* proximal promoters. Consistent with our previous results, compared with *CYP3A5*, the *CYP3A4* proximal promoter with the mutated CCAAT box exhibits weaker binding of NF-YA.

Coordinated regulation of human and porcine *CYP3As* by NF-Y and Sp1. Notably, we observed that the distance between the two *cis*-acting elements is constant at 11 bps (approximately 1 helix of the DNA double helix). To explore whether NF-Y and Sp1 coordinately participate in the transcriptional activation of human *CYP3A* genes, we generated a series of constructs that alter the distance between the NF-Y and Sp1 binding sites of *CYP3A5* and *CYP3A4* as 1 bp, 6 bp, and 16 bp without changing the binding motifs then

MOL # 114439

transfected into HepG2 cells. As shown in Fig. 6A, the activity of the construct containing the region -180 to -1 of *CYP3A5* was 65-fold higher than that of the empty pGL3-basic vector. When we deleted the 10 bases (-10) between the NF-Y and Sp1 binding sites, the promoter activity decreased by 11-fold compared to that with the Basic vector. The 4-base deletion (-4) and 5-base insertion (+5) decreased the activity by 33-fold and 25-fold, respectively. Combined with the data from the element mutation/deletion analysis by the dual luciferase assay (Fig. 4A), these results indicate that these two factors may play a synergistic role in the transcriptional activation of *CYP3A5*. As shown in Fig. 6B, the proximal promoter activity of the -10, -4 and +5 constructs were reduced by 75%, 75% and 54% for *CYP3A4*. And we further confirmed the results in Huh-7 and QGY-7703, QSG-7701 and HepaRG cell lines (supplementary figure 6).

To further investigate whether this coordinated relationship between NF-Y and Sp1 is conserved between the human and porcine *CYP3A* family, we generated similar distance-changed constructs for *CYP3A22* (Fig. 6C) and *CYP3A29/CYP3A46* (Fig. 6D) then transfected into porcine primary hepatocytes. The effect of the different distances between NF-Y and Sp1 on the promoter activity was tested via a dual luciferase assay. The proximal promoter activity of the -10, -4, and +5 constructs was reduced by 93%, 90%, and 63% for *CYP3A22* (Fig. 6C) and 91%, 43%, and 31% for *CYP3A29/CYP3A46* (Fig. 6D). Taken together, without changing the binding sites of NF-Y and Sp1, the spatial distance determined the proximal promoter activities of *CYP3A5* and *CYP3A22* harboring the canonical CCAAT box and *CYP3A29* harboring the mutated CCAAT box, which might weaken the binding of

MOL # 114439

NF-Y, including *CYP4A46* and *3A4*, although to different extents. Overall, these data indicate that Sp1 and NF-Y probably have a conserved synergistic relationship in the transcriptional activity of human and porcine *CYP3A* genes.

MOL # 114439

4. Discussion

The CYP family is a key metabolic enzyme superfamily that is responsible for numerous metabolic pathways of endogenous and xenobiotic compounds in plants, animals, and microorganisms (Nelson et al., 1996). The porcine CYP3A subfamily consists of *CYP3A22* (also known as *CYP3A39*), *CYP3A29*, and *CYP3A46* (Sakuma et al., 2004). All three porcine CYP3As isoforms perform the same function as CYP3A4 in the metabolism of endogenous and xenobiotic compounds (Wang et al., 2011; Wu et al., 2016). Indeed, we found that the distal HNF1 and proximal GC box function in the basal transcriptional regulation of *CYP3A46* (Dong et al., 2015), which is similar to the regulation of *CYP3A4* (Bombail et al., 2004; Matsumura et al., 2004; Tirona et al., 2003). We recently demonstrated that both the CCAAT box and GC box in the proximal promoter play important roles in the basal expression of *CYP3A22* (Liu et al., 2016), which is similar to the regulation of *CYP3A5* in humans (Iwano et al., 2001). These studies indicate that similar to human *CYP3As*, the porcine *CYP3A* isoforms might be regulated by different mechanisms rather than by the conserved regulation of *CYP3A4*.

The *CYP3A* genes are clustered on human chromosome 7 and porcine chromosome 3. Evolutionarily, these genes are potentially regulated by a conserved pattern due to gene duplication (McArthur et al., 2003; Yan and Cai, 2010), but regulation by diversified DNA alterations, including in the promoter region and coding sequence, is possible. To confirm this hypothesis, the alignment of the proximal promoter sequences between the human *CYP3A* genes, including *CYP3A4*, *3A5*, and *3A7*, and the porcine *CYP3A* genes, including *CYP3A22*,

MOL # 114439

3A29, and 3A46, were apparently two conserved proximal *cis*-acting elements (a GC box and a CCAAT box at a constant 11 bp distance). Similarly, the CCAAT box in the proximal promoter regions of *CYP3A4* and *CYP3A7* contained one base-pair mutation, which is identical to that in porcine *CYP3A29* and *CYP3A46* (Fig. 1A). These aspects of the proximal promoter sequences support our hypothesis that the porcine *CYP3As* and human *CYP3As* may share the same core mechanism for the regulation of basal transactivation.

Subsequently, we explored whether these two proximal *cis*-acting elements play determinant roles in the basal expression of *CYP3A29* and *CYP3A46* as previously reported in *CYP3A22* (Liu et al., 2016). The EMSA and DAPA assays confirmed that Sp1 directly binds the canonical GC box of *CYP3A29/CYP3A46* *in vitro* (Fig. 1B, Fig. 2B and Fig. 2C). The mutated CCAAT box in the proximal region of the *CYP3A29/CYP3A46* promoter recruited the NF-Y complex but much weaker, and in contrast to NF-Y, apparently activated the basal expression of *CYP3A22*. Both direct bindings were shown in the ChIP assay *in vivo*, and thus, Sp1 strongly, while NF-Y weakly, contribute to the basal expression of *CYP3A29/CYP3A46* (Fig. 2D). This unequal contribution of the two transcription factors to the basal transactivation of porcine *CYP3As* was further supported by the truncate promoter analysis of *CYP3A29* and binding-site mutagenesis analysis of porcine primary hepatocytes (Fig. 2A and Fig. 3A). Together with our previous data (Dong et al., 2015; Liu et al., 2016), we conclude that the core transactivation region of porcine *CYP3A* genes is in the proximal promoter region, and Sp1 plays a significant role in the basal transcription of *CYP3A22*, *CYP3A29*, and *CYP3A46*. However, the function of NF-Y in the basal expression of *CYP3A29/CYP3A46*

MOL # 114439

remains to be further elucidated due to its contribution to the basal transactivation of *CYP3A22*, which is much stronger than that to the transactivation of *CYP3A29* and *CYP3A46*. To further determine whether the weak binding of NF-Y to the mutated CCAAT box has an impact on its basal expression, we replaced the mutated CCAAT box with either a typical CCAAT box from *CYP3A22* or another null site without NF-Y binding. The luciferase activity apparently increased following the canonical CCAAT box replacement and decreased with the null site construct. Additionally, the DAPA assay revealed that the binding capabilities of NF-Y at the *CYP3A29/CYP3A46* probe are much weaker than those at the *CYP3A22* probe in the proximal promoter region (Fig. 1B). Taken together, NF-Y also regulates porcine *CYP3A* expression to different extents, because the single base variation of the CCAAT site caused alterations in the binding capacity of NF-Y in the proximal promoter region of *CYP3A29*, *CYP3A46*, and *CYP3A22*.

As previously mentioned, the same *cis*-acting elements were identified in the proximal promoter regions of three human *CYP3A* genes. Thus, we were interested to explore the regulatory functions of these two *cis*-acting elements in human *CYP3A* genes. As the predominant CYP in the human adult liver (Lacroix et al., 1997), *CYP3A4* is directly regulated by many transcription factors, and its expression has been extensively studied. However, few studies have investigated the proximal CCAAT box and GC box, and only one study reported an Sp1 site previously identified in the *CYP3A4* promoter region with no functions in transcriptional activation (Bombail et al., 2004). However, this reported Sp1 binding site completely differs from the identified GC box in our study and is 32 bp upstream.

MOL # 114439

Thus, the basal regulation of *CYP3A4*, especially the importance of its proximal promoter region with these two *cis*-acting elements, must be reconsidered, and the *cis*-acting elements must be further elucidated. Regarding the study strategy used to investigate the porcine *CYP3A* genes, the mutation analysis of the proximal binding sites of the *CYP3A5* (Fig. 4A) and *CYP3A4* (Fig. 4B) promoters revealed that both NF-Y and Sp1 participated in the basal transcription activity of human *CYP3As* isoforms. The EMSA and ChIP assays confirmed that both NF-Y and Sp1 directly bind the *CYP3A5* and *CYP3A4* proximal promoters *in vitro* and *in vivo* (Fig. 5A-C and 5D).

The interaction between NF-YA and Sp1 has been shown using a yeast two-hybrid system and GST pull-down (Roder et al., 1999). Sp1 and NF-Y may coordinate in the basal regulation of *CYP3A* genes in humans and pigs. Interestingly, we found that these two *cis*-acting elements in the proximal promoter regions of six *CYP3A* genes maintain a constant distance at 11 bp as a DNA double helix. This coordination was confirmed by a global promoter analysis of the human genome (Benner et al., 2013; Dolfini et al., 2016) as follows: 27% and 18% of 21,000 human promoters have a GC box and CCAAT box in the -150 to +50 region, and the CCAAT box and GC box co-occur with high tendencies in the same promoter with a co-binding pattern. A previous report showed that the heterotrimeric transcription factor NF-Y consists of a minor-groove-binding subunit NF-YA and two histone-fold domains (HFDs) containing subunits NF-YB/NF-YC (Nardini et al., 2013; Romier et al., 2003). The C-terminus of Sp1 contains a canonical DNA binding domain composed of major-groove-binding C₂H₂-type zinc fingers (Thiesen and Schroder, 1991). Based on the

MOL # 114439

known DNA binding structure of NF-Y (PDB ID:4AWL) (Nardini et al., 2013) and Sp1 (PDB ID:1AAY) (Elrod-Erickson et al., 1996; Oka et al., 2004), the spatial configuration of NF-Y, Sp1-DNA binding to promoters with 1 helix distance and the direction of the binding could provide an exact space for NF-Y and Sp1 to bind the *cis*-acting elements. Changing the distance between the NF-Y and Sp1 binding sites repressed the transactivation activities in the proximal promoters of the porcine and human *CYP3A* genes (Fig. 6A-D). These results support the hypothesis that the spatial co-localization of two *trans*-acting factors is required for the basal transcription of *CYP3A* genes in an evolutionarily conserved manner.

In conclusion, our study suggests that porcine and human *CYP3A* genes are constitutively expressed under the control of their proximal promoters. The basal expression of these *CYP3A* genes depends on the recruitment of Sp1 and NF-Y via *cis*-acting elements in the proximal region. Regarding the *CYP3A* genes, the CCAAT box and the GC box contribute to the conserved regulation mechanism but to different extents due to a single base DNA variation.

MOL # 114439

Acknowledgements

We thank Professor. Zhu Xiaofeng (State Key Laboratory of Oncology in South China, Sun Yat-Sen University Cancer Center, China) for kindly provided the Huh-7, QGY-7703, QSG-7701 cell lines.

MOL # 114439

Authorship Contributions

Participated in research design: Ruohong chen, Jun Jiang, Jikai Wen, and Yiqun Deng.

Conducted experiments: Ruohong chen, Zhangsheng Hu, Wenchu Ye and Mengyuan Li.

Performed data analysis: Ruohong chen and Qianqian Yuan.

Wrote or contributed to the writing of the manuscript: Ruohong chen, Jun Jiang, Jikai Wen and Yiqun Deng.

MOL # 114439

References

- Benner C, Konovalov S, Mackintosh C, Hutt KR, Stunnenberg R and Garcia-Bassets I (2013) Decoding a signature-based model of transcription cofactor recruitment dictated by cardinal cis-regulatory elements in proximal promoter regions. *PLoS genetics* **9**(11): e1003906.
- Bombail V, Taylor K, Gibson GG and Plant N (2004) Role of Sp1, C/EBP alpha, HNF3, and PXR in the basal- and xenobiotic-mediated regulation of the CYP3A4 gene. *Drug metabolism and disposition: the biological fate of chemicals* **32**(5): 525-535.
- Burk O, Koch I, Raucy J, Hustert E, Eichelbaum M, Brockmoller J, Zanger UM and Wojnowski L (2004) The induction of cytochrome P450 3A5 (CYP3A5) in the human liver and intestine is mediated by the xenobiotic sensors pregnane X receptor (PXR) and constitutively activated receptor (CAR). *The Journal of biological chemistry* **279**(37): 38379-38385.
- Ceribelli M, Dolfini D, Merico D, Gatta R, Vigano AM, Pavesi G and Mantovani R (2008) The histone-like NF-Y is a bifunctional transcription factor. *Molecular and cellular biology* **28**(6): 2047-2058.
- Dolfini D, Zambelli F, Pedrazzoli M, Mantovani R and Pavesi G (2016) A high definition look at the NF-Y regulome reveals genome-wide associations with selected transcription factors. *Nucleic acids research* **44**(10): 4684-4702.
- Dong L, Chen Q, Liu X, Wen J, Jiang J and Deng Y (2015) Role of Specificity Protein 1, Hepatocyte Nuclear Factor 1alpha, and Pregnane X Receptor in the Basal and Rifampicin-Induced Transcriptional Regulation of Porcine Cytochrome P450 3A4. *Drug metabolism and disposition: the biological fate of chemicals* **43**(10): 1458-1467.
- Dynan WS and Tjian R (1983) Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* **32**(3): 669-680.
- Elrod-Erickson M, Rould MA, Nekludova L and Pabo CO (1996) Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure* **4**(10): 1171-1180.
- Guengerich FP (1999) Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annual review of pharmacology and toxicology* **39**: 1-17.
- Isin EM and Guengerich FP (2007) Complex reactions catalyzed by cytochrome P450 enzymes. *Biochimica et biophysica acta* **1770**(3): 314-329.
- Iwano S, Saito T, Takahashi Y, Fujita K and Kamataki T (2001) Cooperative regulation of CYP3A5 gene transcription by NF-Y and Sp family members. *Biochemical and biophysical research communications* **286**(1): 55-60.
- Jiang J, Tang S, Xia J, Wen J, Chen S, Shu X, Huen MSY and Deng Y (2018) C9orf140, a novel Axin1-interacting protein, mediates the negative feedback loop of Wnt/beta-catenin signaling. *Oncogene* **37**(22): 2992-3005.
- Jiang J, Wang JP, Cai H, Li KB and Deng YQ (2011) CYP3As catalyze nifedipine oxidation in pig liver microsomes: Enzyme kinetics, inhibition and functional expression. *Catal Commun* **12**(8): 694-697.

MOL # 114439

- Kanno Y, Zhao S, Yamashita N, Saito N, Ujiie A, Iijima R, Kikawa N, Nemoto K and Inouye Y (2018) Cell cycle and apoptosis regulator 1, CCAR1, regulates enhancer-dependent nuclear receptor CAR transactivation. *Molecular pharmacology*.
- Kivisto KT, Bookjans G, Fromm MF, Griese EU, Munzel P and Kroemer HK (1996) Expression of CYP3A4, CYP3A5 and CYP3A7 in human duodenal tissue. *British journal of clinical pharmacology* **42**(3): 387-389.
- Kumar GN and Surapaneni S (2001) Role of drug metabolism in drug discovery and development. *Medicinal research reviews* **21**(5): 397-411.
- Lacroix D, Sonnier M, Moncion A, Cheron G and Cresteil T (1997) Expression of CYP3A in the human liver--evidence that the shift between CYP3A7 and CYP3A4 occurs immediately after birth. *European journal of biochemistry* **247**(2): 625-634.
- Li AP, Kaminski DL and Rasmussen A (1995) Substrates of human hepatic cytochrome P450 3A4. *Toxicology* **104**(1-3): 1-8.
- Lim K and Chang HI (2009) O-GlcNAcylation of Sp1 interrupts Sp1 interaction with NF-Y. *Biochemical and biophysical research communications* **382**(3): 593-597.
- Liu X, Wen J, Chen R, Zhang T, Jiang J and Deng Y (2016) T-2 toxin induces the expression of porcine CYP3A22 via the upregulation of the transcription factor, NF-Y. *Biochimica et biophysica acta* **1860**(10): 2191-2201.
- Liu Y, Zeng BH, Shang HT, Cen YY and Wei H (2008) Bama miniature pigs (*Sus scrofa domestica*) as a model for drug evaluation for humans: comparison of in vitro metabolism and in vivo pharmacokinetics of lovastatin. *Comparative medicine* **58**(6): 580-587.
- Ma S, Zhao Y, Sun J, Mu P and Deng Y (2017) miR449a/SIRT1/PGC-1alpha Is Necessary for Mitochondrial Biogenesis Induced by T-2 Toxin. *Frontiers in pharmacology* **8**: 954.
- Maity SN and de Crombrughe B (1998) Role of the CCAAT-binding protein CBF/NF-Y in transcription. *Trends in biochemical sciences* **23**(5): 174-178.
- Matsumura K, Saito T, Takahashi Y, Ozeki T, Kiyotani K, Fujieda M, Yamazaki H, Kunitoh H and Kamataki T (2004) Identification of a novel polymorphic enhancer of the human CYP3A4 gene. *Molecular pharmacology* **65**(2): 326-334.
- McArthur AG, Hegelund T, Cox RL, Stegeman JJ, Liljenberg M, Olsson U, Sundberg P and Celandier MC (2003) Phylogenetic analysis of the cytochrome P450 3 (CYP3) gene family. *Journal of molecular evolution* **57**(2): 200-211.
- Nardini M, Gnesutta N, Donati G, Gatta R, Forni C, Fossati A, Vonrhein C, Moras D, Romier C, Bolognesi M and Mantovani R (2013) Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. *Cell* **152**(1-2): 132-143.
- Nebert DW, Nelson DR, Coon MJ, Estabrook RW, Feyereisen R, Fujii-Kuriyama Y, Gonzalez FJ, Guengerich FP, Gunsalus IC, Johnson EF and et al. (1991) The P450 superfamily: update on new sequences, gene mapping, and recommended nomenclature. *DNA and cell biology* **10**(1): 1-14.
- Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW, Gunsalus IC and Nebert DW (1996) P450

MOL # 114439

- superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**(1): 1-42.
- Oka S, Shiraishi Y, Yoshida T, Ohkubo T, Sugiura Y and Kobayashi Y (2004) NMR structure of transcription factor Sp1 DNA binding domain. *Biochemistry* **43**(51): 16027-16035.
- Omura T and Sato R (1962) A new cytochrome in liver microsomes. *The Journal of biological chemistry* **237**: 1375-1376.
- Roder K, Wolf SS, Larkin KJ and Schweizer M (1999) Interaction between the two ubiquitously expressed transcription factors NF-Y and Sp1. *Gene* **234**(1): 61-69.
- Rodriguez-Antona C, Bort R, Jover R, Tindberg N, Ingelman-Sundberg M, Gomez-Lechon MJ and Castell JV (2003) Transcriptional regulation of human CYP3A4 basal expression by CCAAT enhancer-binding protein alpha and hepatocyte nuclear factor-3 gamma. *Molecular pharmacology* **63**(5): 1180-1189.
- Romier C, Cocchiarella F, Mantovani R and Moras D (2003) The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *The Journal of biological chemistry* **278**(2): 1336-1345.
- Sakuma T, Shimojima T, Miwa K and Kamataki T (2004) Cloning CYP2D21 and CYP3A22 cDNAs from liver of miniature pigs. *Drug metabolism and disposition: the biological fate of chemicals* **32**(4): 376-378.
- Thiesen HJ and Schroder B (1991) Amino acid substitutions in the SP1 zinc finger domain alter the DNA binding affinity to cognate SP1 target site. *Biochemical and biophysical research communications* **175**(1): 333-338.
- Tirona RG, Lee W, Leake BF, Lan LB, Cline CB, Lamba V, Parviz F, Duncan SA, Inoue Y, Gonzalez FJ, Schuetz EG and Kim RB (2003) The orphan nuclear receptor HNF4alpha determines PXR- and CAR-mediated xenobiotic induction of CYP3A4. *Nature medicine* **9**(2): 220-224.
- Wang J, Jiang J, Zhang H, Wang J, Cai H, Li C, Li K, Liu J, Guo X, Zou G, Wang D, Deng Y and Dai J (2011) Integrated transcriptional and proteomic analysis with in vitro biochemical assay reveal the important role of CYP3A46 in T-2 toxin hydroxylation in porcine primary hepatocytes. *Molecular & cellular proteomics : MCP* **10**(9): M111 008748.
- Wu J, Chen R, Zhang C, Li K, Xu W, Wang L, Chen Q, Mu P, Jiang J, Wen J and Deng Y (2016) Bioactivation and Regioselectivity of Pig Cytochrome P450 3A29 towards Aflatoxin B(1). *Toxins* **8**(9).
- Yan J and Cai Z (2010) Molecular evolution and functional divergence of the cytochrome P450 3 (CYP3) Family in Actinopterygii (ray-finned fish). *PloS one* **5**(12): e14276.
- Yang Q, Tang S, Dong L, Chen Q, Liu X, Jiang J and Deng Y (2014) Transcriptional regulation of chicken cytochrome P450 2D49 basal expression by CCAAT/enhancer-binding protein alpha and hepatocyte nuclear factor 4alpha. *The FEBS journal* **281**(5): 1379-1392.
- Zanger UM and Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacology & therapeutics* **138**(1): 103-141.

MOL # 114439

Footnotes

This work was supported by the National Nature Science Foundation of China [Grant 31471234]. Natural Science Foundation of Guangdong Province [Grant 2015A030312005], the Science and Technology Program of Guangzhou [Grant 201804020067, 201607010177], the Department of Education of Guangdong Province [2014GKXM020, 2017KCXTD001] and the Pearl River S&T Nova Program of Guangzhou [Grant 201610010063].

MOL # 114439

Legends for Figures

Fig. 1 The conserved proximal CCAAT boxes and GC boxes in the promoters of porcine and human CYP3As affected the mRNA expression of CYP3A genes. (A) Sequence alignment showing the transcription factor binding sites in the proximal promoter region of typical human and porcine *CYP3A* family isoforms. The conserved NF-Y and Sp1 binding sites are underlined. (B) Binding of the transcription factors NF-Y and Sp1 to the proximal promoter of porcine *CYP3A* genes *in vitro*. A DAPA was performed, and the precipitates were analyzed by western blotting with anti-Sp1 and anti-NF-YA antibodies. Histograms representing the DNA affinity precipitated Sp1/Input Sp1 ratio and DNA affinity precipitated NF-YA/Input NF-YA ratio (n=3). (C) Effects of Sp1 and NF-Y knockdown on the mRNA levels of p*CYP3A* genes in porcine primary hepatocytes. Real-time PCR to determine *CYP3A22*, *CYP3A29*, and *CYP3A46* expression normalized to GAPDH expression in NC, si-Sp1, and si-NF-YA cells (n=3). (D) Effects of Sp1 and NF-Y knockdown on the mRNA level of h*CYP3A* genes in HepG2 cells. Real-time PCR to determine *CYP3A4* and *CYP3A5* expression normalized to GAPDH expression in NC, si-Sp1, and si-NF-YA cells (n=3). Statistical significance was assessed using Student's t-test (two-tailed) (1B) and two-way analysis of variance followed by Dunnett's multiple comparison tests(1C and 1D). The results are shown as the means \pm SD of at least three independent experiments and significance was defined as ***p < 0.001, **p < 0.01, or *p < 0.05.

Fig. 2 NF-Y and Sp1 bind the CCAAT box and GC box, respectively, in the proximal

MOL # 114439

core transactivation region of *CYP3A29/CYP3A46*.

(A) Deletion analysis of the *CYP3A29* promoter in porcine primary hepatocytes. Porcine primary hepatocytes were transiently transfected with the luciferase reporter plasmid pGL3-Basic, which contained different lengths of the *CYP3A29* promoter. The left panel shows the position and length of the constructs. The start codon in the *CYP3A29* ORF was defined as +1. The *Renilla* luciferase reporter plasmid pRL-TK was co-transfected into cells as a normalization control. The right panel shows the relative activity of each construct with respect to that of empty pGL3-Basic vector under the same conditions (n = 3). (B) An EMSA was used to confirm the binding of Sp1 to the *CYP3A29/CYP3A46* proximal promoter. The nuclear fraction was extracted from Sp1 overexpressing COS-7 cells. The top panel shows the sequences of the *CYP3A29* Sp1 probe and mutant probe. The predicted binding sites and nonbinding mutant sequences are underlined. A 100-fold excess of the unlabeled *CYP3A29* Sp1 probe and mutant probe were added to the reactions in lane 3 and lane 4 as competition experiments. The formation of DNA/Sp1 complexes (arrow) was then monitored by 5% native gel electrophoresis. (C) A DAPA was used to further confirm the direct binding of Sp1 to the *CYP3A29/CYP3A46* proximal promoter. The DNA/protein complexes were analyzed by western blotting with anti-Sp1 and anti-NF-Y antibodies (n=3). mut-Sp1 represents Sp1 with a binding site mutation, and mut-NF-Y represents the NF-Y with a binding site mutation. The histograms show the DNA affinity precipitated Sp1/Input Sp1 ratio and DNA affinity precipitated NF-YA/Input NF-YA ratio, normalized to that of the wildtype (*CYP3A29*). (D) A ChIP assay was performed to quantitatively analyze the binding of NF-YA and Sp1 to the

MOL # 114439

cis-acting elements in the *CYP3A29/CYP3A46* proximal promoter in porcine hepatocytes (n=3). Statistical significance was determined using one-way analysis of variance (ANOVA), followed by Dunnett's tests (2A and 2D) and two-way analysis of variance followed by Dunnett's tests (2C). The data are presented as the means \pm SD of at least three independent experiments, and significance was defined as ***p < 0.001, **p < 0.01, and *p < 0.05.

Fig. 3 Sp1 (NF-Y) directly transactivates the transcription of *CYP3A29* (*CYP3A46*).

(A) Mutation analysis of the *CYP3A29/CYP3A46* proximal promoters in porcine primary hepatocytes. The top panel shows the -151 to -115 promoter region; the CCAAT box, GC box and mutated bases are underlined. The ellipse and rectangle represent the binding sites of the transcription factors NF-Y and Sp1, respectively. The black and white shapes represent the wild-type and mutated transcription factor binding sites, respectively. The right panel shows the relative activity of each construct relative to that of the empty vector pGL3-Basic vector under the same conditions (n = 3). (B) Replacement of the NF-Y and Sp1 binding sites in *CYP3A29/CYP3A46* with the sequence in *CYP3A22*. The sequences of the wildtype promoter and mutated promoter used in the experiment are shown in the top panel. In the bottom panel, each column represents the fold-change in the luciferase activity of each mutated construct relative to that of Basic (the pGL3-Basic plasmid) (n=3). mut-Sp1 represents the Sp1 binding site mutation, and mut-NF-Y represents the NF-Y binding site mutation. Statistical significance was determined using one-way analysis of variance (ANOVA), followed by Dunnett's tests (3A) and two-way analysis of variance followed by Dunnett's tests (3B). The

MOL # 114439

data are presented as the means \pm SD of at least three independent experiments, and significance was defined as *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Fig. 4 Both NF-Y and Sp1 activate *CYP3A5* and *CYP3A4* transcription by binding to the proximal CCAAT box and GC box.

(A and B) Mutation analysis of the proximal *CYP3A5* and *CYP3A4* promoters in HepG2 cells. The top panels in A and B show the wildtype and mutated promoter sequences. The mutated region (underlined) destroyed the CCAAT box and GC box. The schematic maps on the left show the mutant constructs. The right panel shows the promoter activity relative to that of Basic (the pGL3-Basic plasmid) (n=3). mut-Sp1 represents the Sp1 binding site mutation, and the mut-NF-Y represents the NF-Y binding site mutation. Statistical significance was determined using one-way analysis of variance (ANOVA) followed by Dunnett's tests. The data are presented as the means \pm SD of at least three independent experiments and significance was defined as *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Fig. 5 NF-Y and Sp1 bind the proximal CCAAT box and GC box, respectively, of *CYP3A5/CYP3A4*.

(A and B) EMSA of *CYP3A5* NF-Y and Sp1 probes. Probes corresponding to the *CYP3A5* proximal CCAAT box (A) and GC box (B) are shown above the gel. Binding sites and mutated sequences are underlined. Probe incubated with nuclear fraction (5 μ g) from HepG2 cells. EMSAs were performed as described in Materials and methods. DNA/protein

MOL # 114439

complexes and supershift band are indicated with arrows. (C) EMSA of *CYP3A4* Sp1 probes. (D) ChIP assays were performed to quantitative analysis of the binding of NF-YA and Sp1 to *cis*-acting elements in the *CYP3A4/CYP3A5* proximal promoter in HepG2 cells. Statistical significance were determined using two-way analysis of variance (ANOVA), followed by Dunnett's multiple comparison tests. Data are presented as the means \pm SD of at least three independent experiments and significance was defined as *** $P < 0.001$, ** $P < 0.01$, and * $P < 0.05$.

Fig. 6 Coordinated regulation of human and porcine *CYP3As* by NF-Y and Sp1.

(A and B) Change in distance between the NF-Y and Sp1 binding sites in the *CYP3A5*(A) or *CYP3A4*(B) promoter region impacts transcript activation. As described in Materials and methods, a series of constructs with different distances between these two transcription factors were generated and transfected into HepG2 cells. The left panel shows the distance between the NF-Y and Sp1 binding sites before and after the changes. The right panel represents the fold change in luciferase activities relative to pGL3-Basic (Basic) (n=3). (C and D) Change in the distance between the NF-Y and Sp1 binding sites in the *CYP3A22* (C) and *CYP3A29* (D) promoter regions impacts transcript activation. Distance-changed constructs were generated as described in Materials and methods and then transfected into porcine primary hepatocytes. Statistical significance were determined using one-way analysis of variance (ANOVA), followed by Dunnett's multiple comparison tests. Data are presented as the means \pm SD of at least three independent experiments and significance was defined as *** $P < 0.001$, ** $P < 0.01$,

MOL # 114439

and $*P < 0.05$

FIGURES

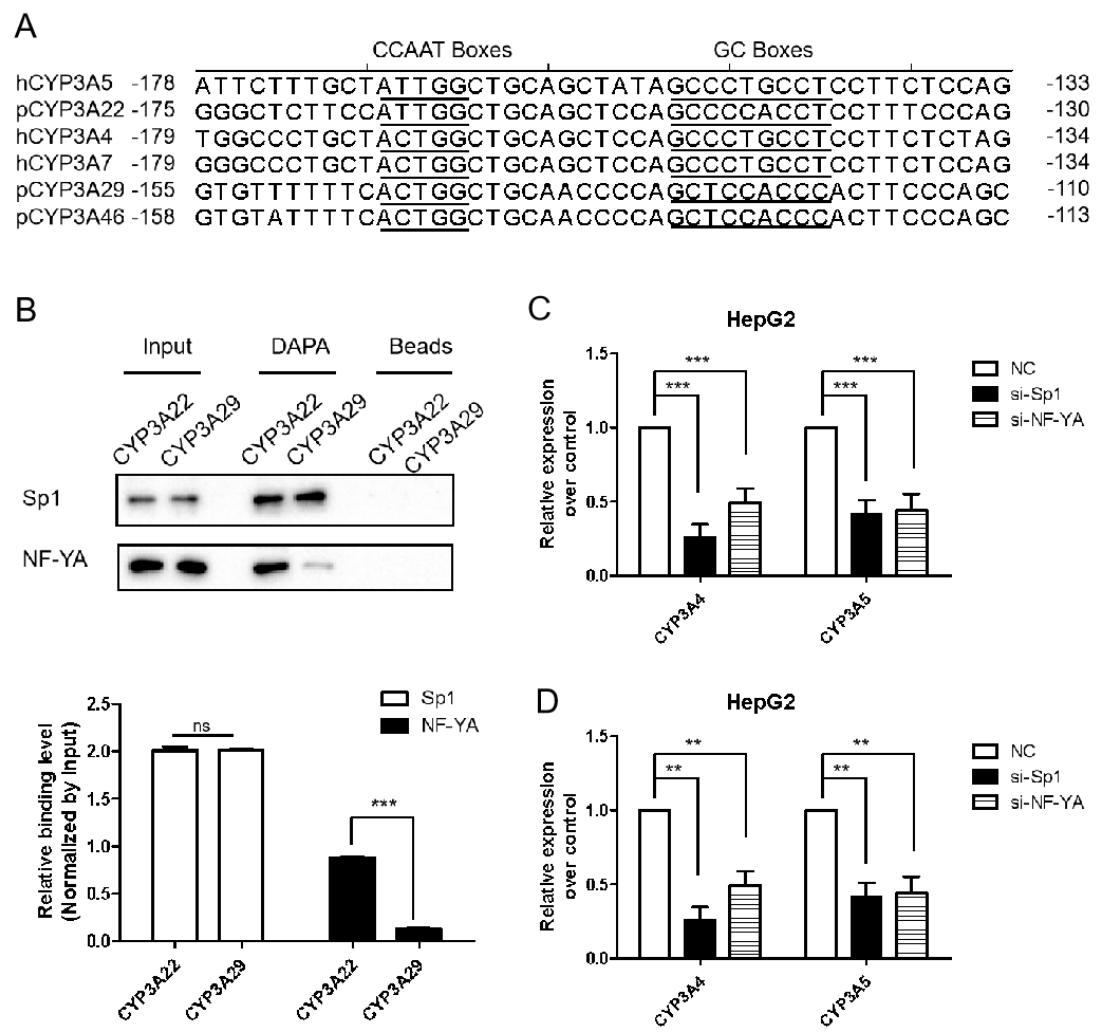


Figure 1



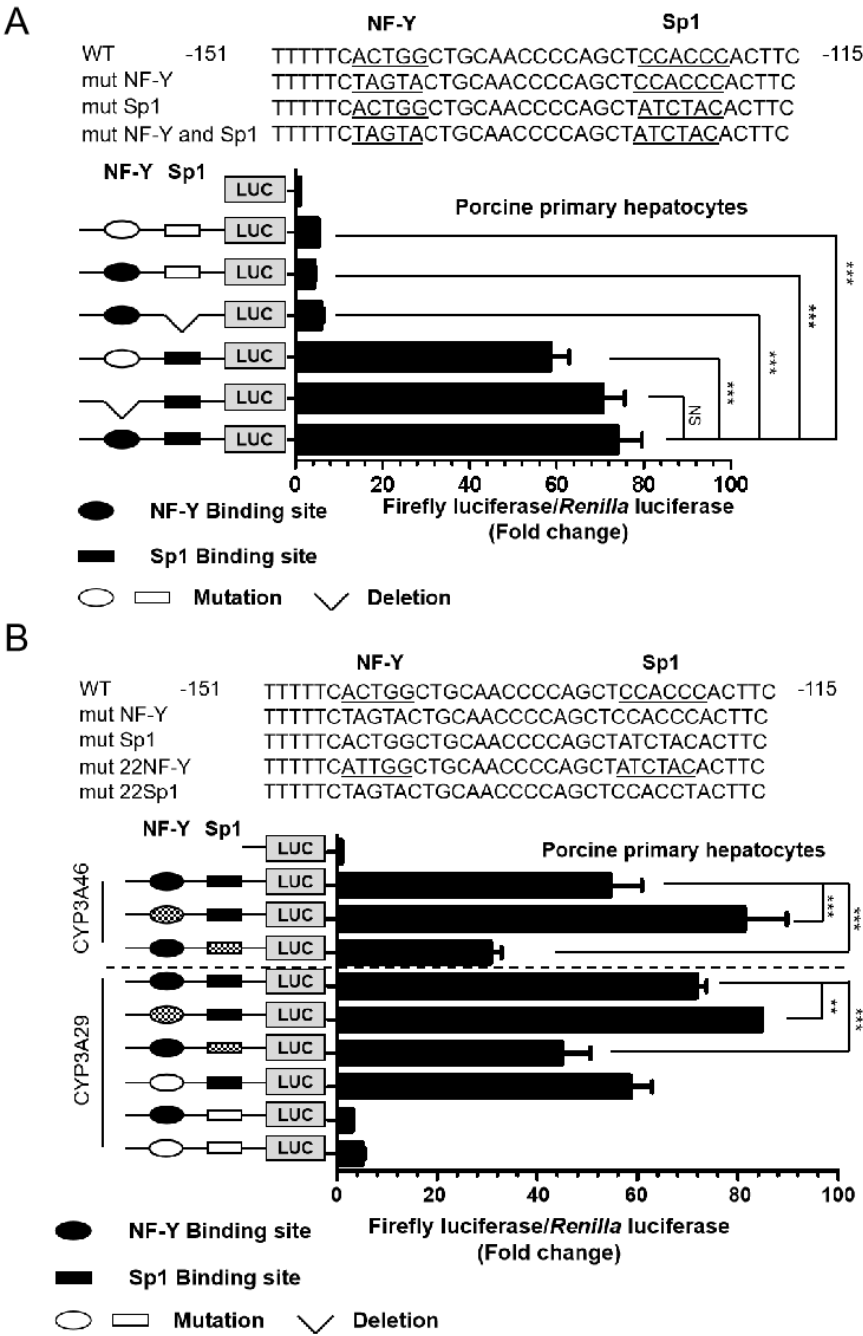


Figure 3

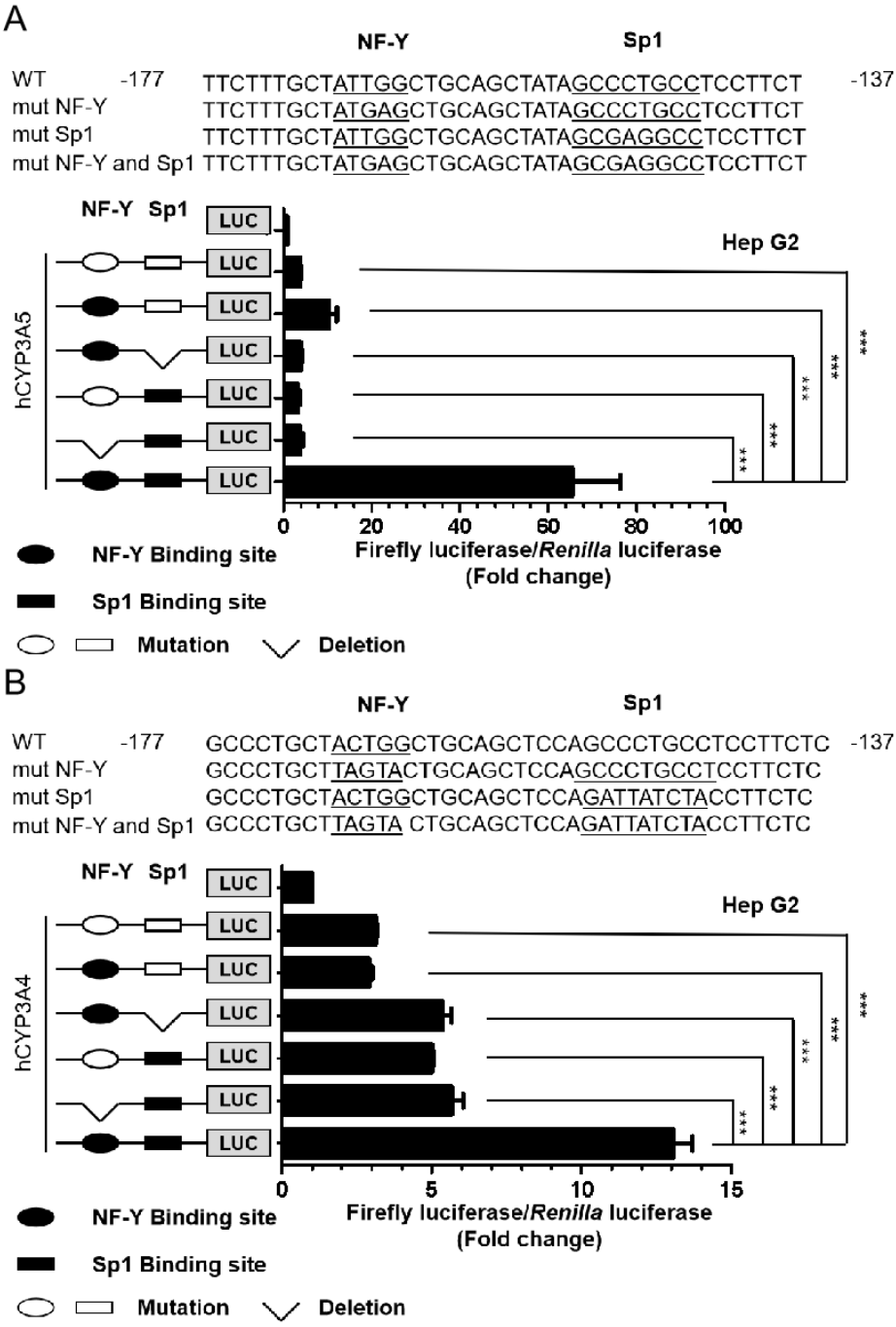


Figure 4

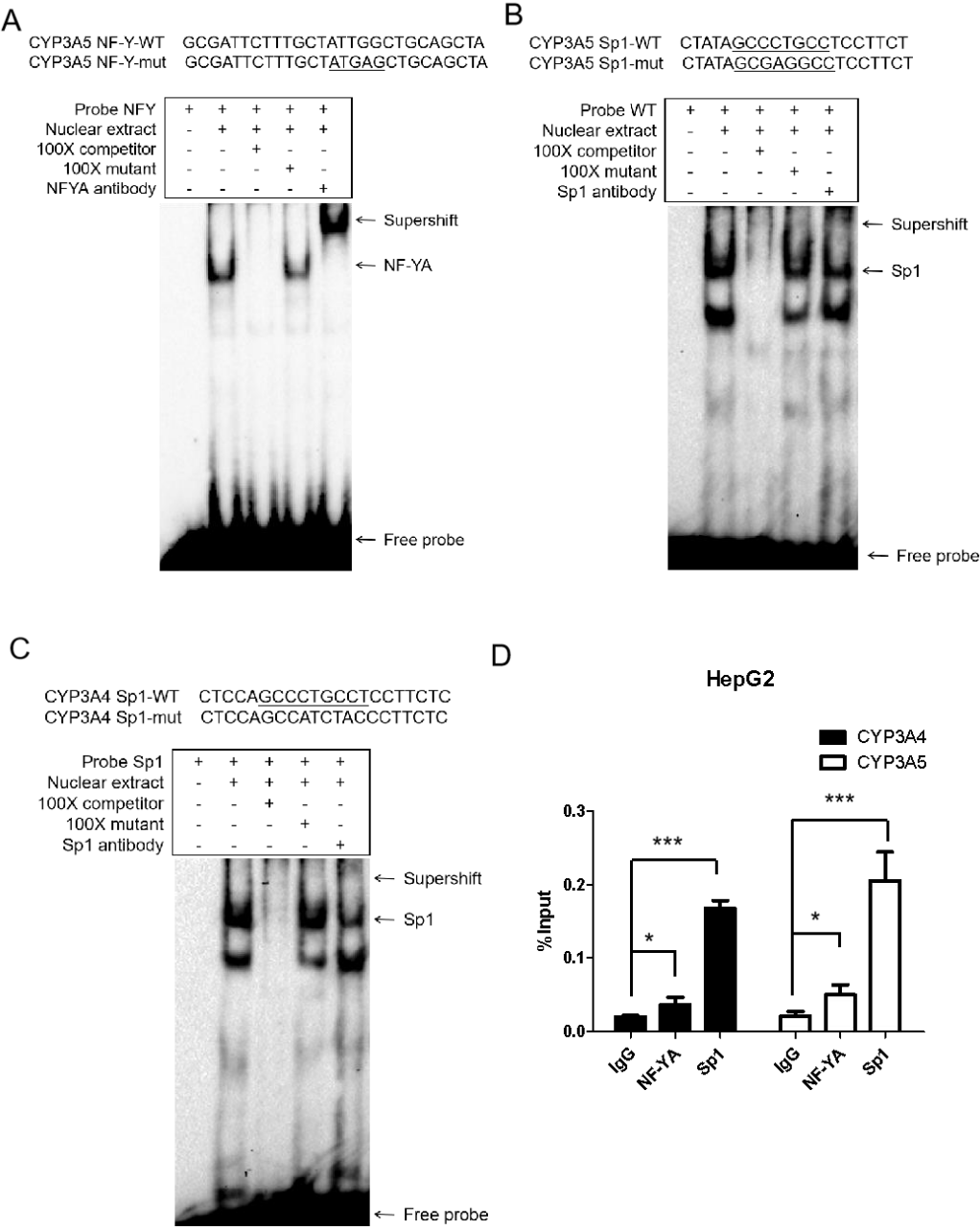


Figure 5

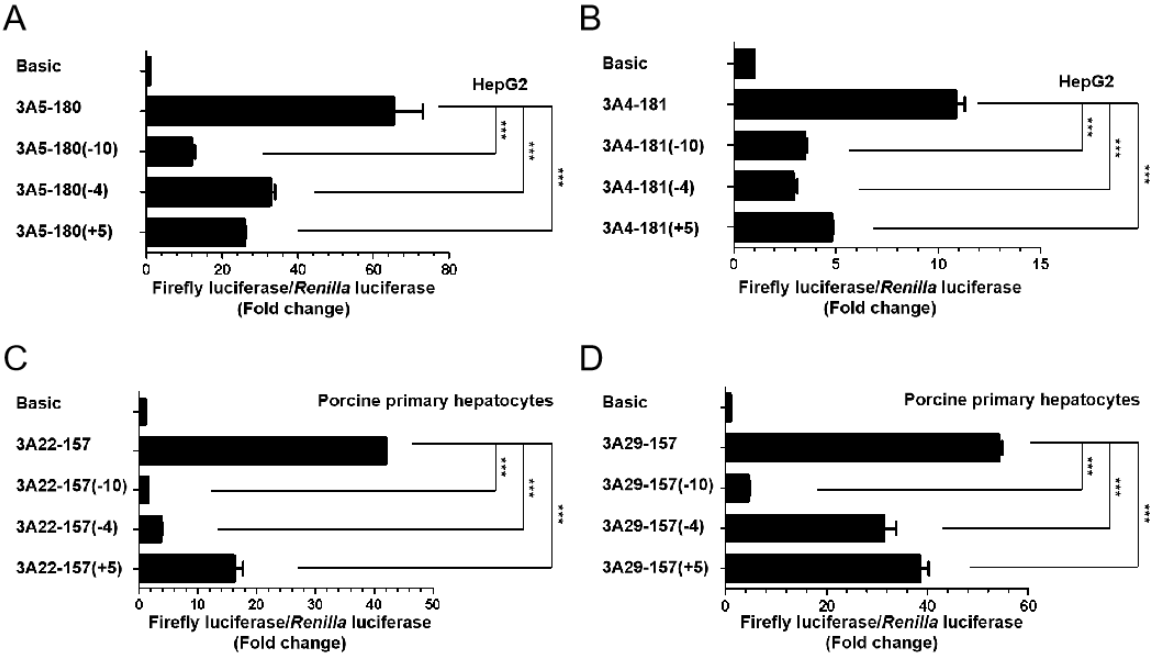


Figure 6