Foxa1 and Foxa2 together control developmental gene regulatory networks, and differentiation genes, in both human stem-cell derived liver progenitors and in a human liver cell line: evidence of a collapse of human liver differentiation

Iyan Warren[1], Mitchell Maloy[1], Daniel Guiggey[1], Ogechi Ogoke[1], Theodore Groth[1], Tala Mon[1], Saber Meamardoost[1], Xiaojun Liu[1], Antoni Szeglowski[4], Ryan Thompson[1], Peter Chen[3], Ramasamy Paulmurugan [4], Natesh Parashurama[1,2,3]

[1] Department of Chemical and Biological Engineering, University at Buffalo (State University of New York), Furnas Hall, Buffalo, NY 14260

[2] Clinical and Translation Research Center (CTRC), University at Buffalo (State University of New York), 875 Ellicott St., Buffalo, NY 14203

[3] Department of Biomedical Engineering, University at Buffalo (State University of New York), Furnas Hall, Buffalo, NY 14260

[4] Department of Radiology, Canary Center for Early Cancer Detection and the Molecular Imaging Program at Stanford, Stanford University, Palo Alto, CA 94304-5483

* Corresponding Author: Natesh Parashurama, 907 Furnas Hall, Buffalo, NY 14260; Tel: 716-645-1201; Fax: 716-645-3822; e-mail: nateshp@buffalo.edu
Running Title**:**

Running Title: Blocking human gut and liver differentiation with Foxa1/2

Keywords: Human pluripotent stem cells, gut tube, Foxa1, Foxa2, RNAi, endoderm, liver, hepatic progenitors, gene regulatory network, hepatic nuclear network, endoderm, stem cells, differentiation, shRNA, hepatic differentiation

ABSTRACT:

Foxa factors (Foxa1/2/3) play a critical role in liver organogenesis and initiating and maintaining developmental gene regulatory networks (GRN) in murine endoderm, gut, and liver. Nevertheless, it remains to be determined how Foxa2 together with Foxa1 (which compensates for Foxa2 loss) function during human liver differentiation/development. In novel, hypoxic protocols that induce hPSC-derived endoderm and liver progenitors, RNAi (siRNA and shRNA)-mediated knockdown of both Foxa1 and Foxa2 demonstrates widespread regulation of the endoderm and mesendoderm GRN, and albumin, which signifies liver differentiation block. The Foxa1/2 (-/-) phenotype in liver cells demonstrated significant downregulation of albumin, liver GRN (HNF4α, Hex, HNF1ß, and Tbx3), all consistent with a collapse of liver differentiation, and a significant upregulation of germ layer markers like Nanog (pluripotency), Pax6 (neurectoderm), GATA4 (endoderm/mesoderm), Cdx2 (intestine). RNA-seq and bioinformatics analysis of the shRNA Foxa1/2 (-/-) phenotype confirms global downregulation of liver-specific genes and unexpected upregulation of epithelial differentiation, mesodermal (cardiac), and neural genes. These data suggest time-dependent effects on endoderm and liver GRN, widespread effects on the liver transcriptome. Additionally, the Foxa1/2 (-/-) phenotype upregulates of cardiac, neural, and intestinal genes, and this data phenocopies data from human hepatocyte differentiation. These data suggest a novel role for Foxa1/2 in human liver differentiation.

INTRODUCTION:

Hepatic gene expression is precisely controlled by regulatory TFs which form a core GRN. These are well-studied, evolutionarily conserved, TFs that together form a GRN termed the hepatic nuclear factor network established in rodents and humans. The hepatic GRN include Foxa1, Foxa2, Foxa3, HNF1α, HNF1ß, HNF4α, HNF6, Hex, Tbx3, and Prox1 (Lau, Ng et al. 2018);(Harries, Brown et al. 2009). Despite being established for over 20 years, recent studies have heightened interest in the role of hepatic GRN-TF (Tauran, Poulain et al. 2019);(Wang, Wang et al. 2019). These core hepatic GRN-TF (Lau, Ng et al. 2018) function not only in differentiation and lineage choice (Ancey, Ecsedi et al. 2017), but also in tuning mature functions like metabolism (Wang and Wollheim 2005), and regeneration (Huck, Gunewardena et al. 2019). Adding their clinical significance, they play a major role in defining disease states like hepatic cancer (Wang, Zhu et al. 2014), fibrosis (Desai, Tung et al. 2016), and alcoholic hepatitis (Argemi, Latasa et al. 2019). Further intertwining these GRN with human health, is that they have been extremely successful in studies of transcription factor (TF)-based reprogramming experiments from fibroblasts to hepatocyte-like cells (Sekiya and Suzuki 2011). Improved understanding of GRN is therefore critical for an improved understanding of hepatocyte differentiation, reprogramming, physiology, and disease.

An important component of the core liver GRN is the Foxa family of TFs, which are a well-studied family of proteins that have important functions not only in liver physiology but also in liver development and organogenesis. Developmental defects in Foxa2-null mouse mutants result in an embryonic lethal phenotype with defective formation of the foregut endoderm (Cirillo, Lin et al. 2002); (Hallonet, Kaestner et al. 2002). Overexpression studies indicate a role in endoderm maintenance (Levinson-Dushnik and Benvenisty 1997). Seminal studies of the Foxa2 binding to the albumin promoter within endoderm revealed that it remodels compacted, normally inaccessible chromatin at the silent albumin promoter. Due to its unique structure (Shim, Woodcock et al. 1998);(Gadue, Gouon-Evans et al. 2009) which resembles histone tail proteins, Foxa2 can remodel chromatin by binding to histone binding sites at promoters/enhancers of hundreds of silent, liver specific genes, and its activity has been previously described as genetic potentiation or "pioneering" activity (Bossard and Zaret 1998); (Duncan, Navas et al. 1998); (Jung, Zheng et al. 1999); (Lee, Friedman et al. 2005). Consistent with this, in a conditional (Foxa3)-driven Foxa2 and Foxa1 double knockout (-/-) mouse, the early liver bud does not form, and

normal markers for endoderm and hepatic endoderm, like alpha-fetoprotein (Afp) and albumin (Alb), are absent (Lee, Friedman et al. 2005). In these studies, it was demonstrated that Foxa1 compensates for Foxa2, and therefore a double knockout was needed to observe the global effect on liver development. Indeed, this functional redundancy for both Foxa1 and Foxa2 knockout has been observed in studies investigating bile duct, pancreas, and lung development (Li, White et al. 2009); (Gao, LeLay et al. 2008); (Wan, Dingle et al. 2005) and intestinal cell differentiation (Ye and Kaestner 2009). During murine fetal liver development at day 18, microarray studies indicate that Foxa1 and Foxa2 are functionally redundant (Bochkis, Schug et al. 2012), in addition to the early liver bud studies. In the adult liver, studies of gene expression data strongly suggests that Foxa1 and Foxa2 do not compensate for each other (Bochkis, Schug et al. 2012), while in the same study, CHIP-seq co-occupancy studies indicate that ~1/4 to1/5 of the binding sites could be co-occupied by either Foxa1 or Foxa2. Co-occupied genes in adult liver included genes that regulate transcription, embryonic development, lipid metabolism, vesicular transport, and xenobiotic metabolism, as well as the HNF factors, HNF1α and HNF4α. Foxa1/2 co-occupied sequences also had increased recognition motifs for HNF1α and HNF4α, indicating an active developmental GRN in adult mouse liver, supported by another study (Watts, Zhang et al. 2011). Further strengthening the role of Foxa factors (Foxa1, Foxa2, Foxa3), a recent study demonstrates adenoviral-mediated knockdown of Foxa1, 2, and 3 results in adult liver results in global downregulation of liver genes and a resulting pre-hepatic cell fate after 7 days and liver failure by day 15 (Reizel, Morgan et al. 2020). These studies indicate that Foxa factors are required to initiate liver development and maintain nucleosome structure in adult liver cells. Foxa2 has profound effects on differentiation and metabolism in rodent systems, but the effects are less established in human cells.

The role for Foxa factors in regulating human GRN has been investigated in models of human development using human pluripotent stem cells (hPSC), which can be differentiated towards human endoderm, gut tube endoderm, and hepatic progenitor cells. A recent study performed a CRISPRi screen during endoderm induction, and demonstrated that Foxa2 controls expression of over 250 transcripts when targeted in endoderm (Genga, Kernfeld et al. 2019). GRN within the mesendoderm are affected, including Hex, Mixl1, Cer1, and Gsc, but Sox 17 is not affected by Foxa2 knockdown, and an altered chromatin landscape is evident. However, these studies to do not employ targeting of both Foxa1 and Foxa2, and despite the fact that there were

effects on foregut induction and hepatic differentiation, it is possible that Foxa1 compensation for Foxa2 knockdown occurred. Foxa1 priming of human liver genes within endoderm was studied in the context of developmental competence. It was shown that poised enhancer states at developmental enhancers, dictated by levels of H3K4 methylation and H3K27 acetylation, together with Foxa1binding, primed genes for liver and pancreatic gene expression (Wang, Yue et al. 2015). However, these studies also did not employ analysis of Foxa1 and Foxa2. Foxa1 and Foxa2 have also received increased interest, because they have been widely used to program fibroblasts to hepatocytes in mouse and human cells (Sekiya and Suzuki 2011); (Guo, Tang et al. 2017); (Nakamori, Akamine et al. 2017); (Song, Pacher et al. 2016); (Rezvani, Espanol-Suner et al. 2016). Although there is high homology in whole genome sequences between species, the GRN that govern embryonic and organ development surprisingly, are structurally and functionally distinct (Odom, Dowell et al. 2006). Taken together, there remain gaps in our understanding of the combined role of both Foxa1 and Foxa2 in regulating GRN in human hPSC-derived endoderm, gut tube endoderm, and hepatic endoderm.

To address this gap, we first assessed the Foxa1/2 (-/-) phenotype in hPSC-derived endoderm progenitor cells and a novel model of hPSC-liver induction under hypoxic conditions that relies on minimal growth factors. Employing RNAi, in endoderm progenitor cells, we observed widespread effects on mesendoderm/endoderm GRN, and during liver differentiation, we observe a significant downregulation of the albumin gene, consistent with a potential block in the activation of liver differentiation. The Foxa1/2 (-/-) phenotype was then analyzed in stable liver cells lines, and both siRNA transfection and stable shRNA transduction resulted in global downregulation of all major liver GRN, and unexpectedly, upregulation of germ layer genes in liver cells, including neural, cardiac, and endoderm genes. We performed transcriptomic and bioinformatics analysis of shRNA Foxa1/2 (-/-) phenotype. The data demonstrated downregulation of liver-specific genes, but also, upregulation in cardiac (mesoderm) and neural (neurectoderm) genes. This Foxa1/2 phenotype in a stable liver cell line shared great similarity with a phenotype presented after the analysis of transcriptomic data of seminal liver studies hPSC-derived hepatocyte-like cells, thereby validating the phenotype. These data suggest that Foxa1/2 plays a large role in human liver differentiation. These data suggest that Foxa1/2 phenotype has time-dependent, potentially reversible, effects on GRN and widespread effects on transcriptome, and that liver developmental GRN may function in a stable liver cell line.

METHODS:

Reagents/Materials

Dulbecco's modified Eagle's medium (DMEM) (Cat. #: 10566024), OptiMEM (Cat. #: 31985070), RPMI (Cat. #: 61870036), IMDM (Cat. #: 12440053), KO Serum (10828010), Penicillin-streptomycin (10000 U/mL) (Cat. #: 15140122), Fetal Bovine Serum (Cat. #: A3160602), 0.05% Trypsin-EDTA (Cat. #: 25300062), B27 (50x) (Cat. #: 17504044), Ham's F12 (Cat. #: 11765054), N2 Supplement (100x) (Cat. #:17502048), 50x B27 supplement (Cat. #: 17504044), Lipofectamine RNAiMAX Transfection Reagent (Cat. #: 13778030), Beta Actin Loading Control Antibody (Cat. #: MA5-15739), Puromycin: Liquid (20mL) (Cat. #: A1113802), DMEM (Invitrogen Cat. #: 10566024), DAPI (Cat. #: 62248), Precast gels-Bolt 4-12% Bis-Tris Plus 10-well gels (Cat. #: NW04120BOX), and nitrocellulose transfer stacks (Cat. #: IB23002) were purchased from Thermo Fisher (Waltham, MA). STEMdiff Definitive Endoderm induction medium (Cat. #: 05110), Activin A (Cat. #: 78001.1), mTESR1 (Cat. #: 85850), keratinocyte growth factor (KGF, FGF-7) (Cat. #: 78046), fibroblast growth factor 2 (FGF-2, bFGF) (Cat. #: 78003.1), epidermal growth factor (EGF) (Cat. #: 78006.1), and Gentle Cell Dissociation Reagent (GDR) (Cat: #: 07174), were purchased from StemCell Technologies (Vancouver, CA). G418 (Cat. #: G8168-10ML), Doxycycline, (Cat. #: D9891-1G), L-Ascorbic Acid (Cat. #: A4544), MTG (Monothioglycerol) (Cat. #: M6145), Polybrene (hexadimethrine bromine) (Cat. #: 107689), were purchased from Sigma Aldrich (St. Louis, MO). Aurum Total RNA Mini Kit (Cat. #: 7326820), DNase I (Cat. #: 7326828), iTaq Universal SYBR Green Supermix (Cat. #: 1725121), and iScript cDNA Synthesis Kit (Cat. #: 1708891) were purchased from Bio-Rad. Pooled siRNA for Foxa1 (siGENOME SMART POOL Human FOXA2 (#3170)) and for Foxa2 (siGENOME SMARTPOOL Human FOXA1 (#3169)) were purchased from Dharmacon (now Horizon Discovery Group (Waterbeach, UK). siGLO Cyclophilin B Control siRNA (D-001610-01-05) and non-targeting siGENOME nontargeting pool #1 (D-0012060-13-05) was purchased from Dharmacon GE life sciences (now Horizon Discovery Group (Waterbeach, UK). Matrigel, Growth factor-free (Cat. #: 40230), was purchased from Corning (Corning, NY). Bovine Serum Albumin (Cat. #: 10791-790) was purchased from VWR International (Radnor, PA). CHIR (CHIR99021) (Cat. #: 13122) was purchased from Cayman Chemical (Ann Arbor, MI). Rho-associated kinase

(ROCK) inhibitor (Y27632 2, Cat. #: MBS577605) was purchased from Mybiosource.com. Fugene HD Transfection Reagent (Cat. #: E2311) was purchased from Promega (Madison, WI). 96-well PCR plates (Cat. #: L223080), tissue culture treated 24-well plate (Cat. #: 702001), 75-cm^2 polystyrene tissue culture-treated flasks (Cat. #: 708003), PCR Plate Covers (Cat. #: HOTS-100) were purchased from Laboratory Product Sales, Inc (Rochester, NY). The pLKO.1 puro (Cat. #: 8453), pLKO-shFoxa1#1 (Cat. #: 70095), pLKO-shFoxa1#2 (Cat. #: 70096), and pLKO-shScramble. (Cat. #: 1864) plasmids were purchased from Addgene (Watertown, MA). Human hepatoma (HepG2) (Cat. #: HB-8065) were purchased from ATCC (Manassas, VA). All primers were purchased from either Integrated DNA technologies (IDT) (Newark, NJ), Sigma Aldrich (St. Louis, MO) or Thermofisher (Waltham, MA).

Antibodies

Mouse anti-human beta actin (Cat. #: MA5-15739), mouse anti-human Foxa2 monoclonal (Cat. #: MA5-15542), Goat anti-mouse secondary antibody (Alexa Fluor 488) (Cat. #: A-11034), and Goat anti-rabbit secondary antibody (Alexa Fluor 568) (Cat. #: A-11011) were purchased from Thermo Fisher (Waltham, MA). Mouse anti-human albumin monoclonal (Cat. #: ab10241) was purchased from Abcam (Cambridge, MA). Mouse anti-human albumin monoclonal antibody (Cat. #: sc-271604), mouse anti-human Alpha-fetoprotein (AFP) monoclonal antibody (Cat. #: sc-130302), Mouse anti-human Oct3/4 antibody (Cat. #: sc-5279), FITC- and TRITC-conjugated, anti-mouse, rabbit, and goat antibodies, and IgG control antibodies were purchased from Santa Cruz Biotechnology (Dallas, TX). Rabbit anti-human Sox17 antibody (Cat. #: NBP2-24568) was purchased from Novus Biologicals (St. Louis, MO). HRP-conjugated secondary antibodies were purchased from Jackson ImmunoResearch Laboratories (West Grove, PA).

Preparation of human pluripotent stem cell (hPSC) differentiation medium (SFD medium)

Basal differentiation medium for hPSC differentiation towards endoderm, gut tube, and liver contained SFD, a serum-free, defined medium based upon other mouse and human stem cell studies (Gadue, Huber et al. 2006). SFD medium contains 75% IMDM or RPMI supplemented with 25% Ham's F12, 0.5% N2 Supplement, 0.5% B27 supplement, 2.5 ng/ml FGF2, 1% Penicillin

+ Streptomycin, 0.05% Bovine Serum Albumin, 2mM Glutamine, 0.5mM Ascorbic Acid, and 0.4mM Monothioglycerol (MTG).

Feeder-free culture (maintenance) of human pluripotent stem cells

We performed experiments using the UCSF4 human embryonic stem cell (hESC) line (NIH Registry 0044, female), a kind gift from Susan Fisher, PhD, UCSF Institute for Regenerative Medicine and stem cells, a commercially available induced pluripotent stem cell (iPSC) line (BXS0114 ACS1028 female (ATCC), and an NIH registered iPSC cell line (HUES53, (NIH registry 0055, male)), a kind gift from Richard Gronostajski, PhD, University at Buffalo. Human pluripotent stem cells (hPSC) were cultivated at 90% N_2, 5% O_2, 5% CO_2, (Tri-gas HERAcell VIOS 160i CO2 incubators) using mTESR1 medium (warmed to room temperature), on 6-well tissue culture-treated plates, coated with 1:15 diluted (in DMEM) growth factor free-matrigel. Wells were coated with matrigel by adding 1 mL of diluted matrigel per well of a 6-well plate, and incubating for 1.5-2 hrs. at 37°C. Excess dilute matrigel was then removed and the wells were washed with PBS. Cell culture medium was changed every other day. For passaging, the mTESR1 serum-free maintenance medium was removed, 1 mL of gentle cell dissociation reagent (GDR) (Stem Cell Technologies) was added for 10-15 minutes, and single cells or clumps of cells were harvested from the dish. Cells were centrifuged 3 minutes at 800-1000 RPM (Eppendorf 5810 table top centrifuge) and resuspended. Cells were frozen in mTESR1 medium + 5% DMSO at -80°C overnight, followed by liquid nitrogen cryostorage. Passage number varied between 15-35 for all experiments.

Endoderm induction from human stem cells under low oxygen

Endoderm induction was carried out under low oxygen conditions (5% O_2 , 5% CO_2) using two different differentiation protocols. hPSCs were harvested by replating cells in mTESR with 10 μM ROCK inhibitor overnight at 50,000-200,000 cells in a growth factor-free matrigel-coated, tissue culture treated 24-well plate. Coating was accomplished by adding 300 μL of diluted matrigel (1:15 growth factor-free matrigel in DMEM), per well, for 1.5-2 hours of coating, with excess matrigel removed. The next day, endoderm was induced using one of two different protocols.

The first endoderm protocol was the STEMdiff definitive endoderm induction kit. 500 µL of Medium 1 (containing reagent A + B) was added from day 0 to day 1, and 500 µL of Medium 2 (containing reagent B) was added on days 2 to 4. A second endoderm protocol was developed specifically for improved survival at low oxygen and improved morphology. In RPMI medium with 1x B27 (no insulin) and 0.2% KO serum, definitive endoderm was induced with Activin A (100 ng/ml) and CHIR (3 µM) for 1 day, followed by Activin A (100 ng/ml) for 3 more days. Medium was changed daily and 500-750 µL medium was used per well. The 0.2% KO serum was added for improved viability at 5% O_2, and higher seeding densities improved culture morphology.

Gut tube induction from human pluripotent stem cells

Gut tube (GT) endoderm induction from definitive endoderm was performed in SFD medium (Gadue, Huber et al. 2006) (see: Preparation of SFD medium) containing keratinocyte growth factor (KGF or FGF7) (25 ng/ml) for an additional 2 days.

Hepatic differentiation from human pluripotent stem cells (hPSC)

Two protocols were used for liver differentiation. *Growth factor-free gut tube/ hepatic differentiation protocol:* To promote spontaneous differentiation of hPSC under 5% O_2 conditions, we added SFD medium with KGF (25 ng/ml) for a total of 10 days until day 14, with medium changes every day. Cells remained in 24-well plates, and SFD medium was changed daily. No additional growth factors were added.

Growth factor-positive hepatic differentiation protocol: To promote hepatic differentiation with growth factors, we adopted a protocol from the literature (Takebe, Sekine et al. 2013). Briefly, on day 6 of culture after gut tube induction, bone morphogenetic protein 4 (BMP4, 10 ng/ml), and fibroblast growth factor 2 (FGF2, 20 ng/ml) were added from days 7-10, and hepatocyte growth factor (HGF, 10 ng/ml), oncostatin (20 ng/ml), and dexamethasone (100nM) were added from days 10-14. Cells were assayed by qRT-PCR and immunostaining on day 14.

Stable liver cell line culture

Human hepatoma (HepG2) (Cat. #: HB-8065) cells were maintained with 10 mL of DMEM containing 1% Penicillin-Streptomycin (P/S), and 10% fetal bovine serum (FBS) in culture in 75-cm^2 polystyrene tissue culture-treated flasks (LPS), and incubated at 5% CO_2, 21% O_2, and at

37°C. Medium changes were performed every two days. HepG2 cells were passaged weekly by the addition of 5 ml of either 0.05% or 0.25% Trypsin-EDTA, and replated at 1:5-1:10 dilution, with passage numbers for all experiments ranging between 15-50.

RNA isolation, reverse transcription (RT) and quantitative polymerase chain reaction

Total RNA was purified with Aurum Total Mini Kit (Bio-Rad) using the spin column method with DNase 1 (Bio-Rad, Hercules, CA) (reconstituted in 10mM tris) treatment. RNA concentrations were determined by Nanodrop. RNA was converted to cDNA with an RT reaction using the iScript cDNA Synthesis Kit (Bio-Rad), and the mass of RNA was calculated such that 5 ng RNA per well to be run in the qPCR reaction. The RT reaction was performed using 5 minutes at 25º C, 20 minutes at 46º C, and 1 minute at 95ºC. Reactions were then held either at 4º C or on ice. We performed 10 µL qPCR (3 µL primers at a concentration of 0.3 µM, 1 µL nuclease free water, 1 µL cDNA, and 5 µL supermix) reactions with iTaq Universal SYBR Green Supermix (BioRad) in a 96-well PCR plate (LPS). The qPCR reaction was done in a CFX96 Touch Real-Time PCR Detection System (BioRad). The qPCR reaction consisted of polymerase activation and DNA denaturation at 98ºC for 30 seconds followed by, 40 to 45 cycles of 98ºC for 15 seconds for denaturation and 60º C for 60 seconds for annealing and extension. Melt curve analysis was performed at 65-95º C in 0.5ºC increments at 5 seconds/step. Relative, normalized, gene expression was analyzed using the delta-delta-Ct method (Livak and Schmittgen 2001), with three duplicates per gene tested. Primer sequences are as shown **(Supp. File 1).**

Short-interfering RNA (siRNA) transfection of human stem cell-derived cells.

Positive control siRNA (siRNA PPIB), negative control pooled siRNA, pooled siRNA targeting human Foxa1, and human Foxa2 siRNA, (Dharmacon) was used. In a tissue culture-treated 24-well plate (LPS), 100,000 hPSC (ESC or iPSC cell lines) were plated for endoderm differentiation using the STEMDiff protocol. Prior to transfection, cells were washed with 500 µL of fresh culture medium. Lipofectamine RNAiMAX reagent (1.5 µL) was mixed with Opti-MEM medium for a total of 50 µL, and 5 pmol siRNA (positive control, negative control, Foxa1, or Foxa2) was mixed with Opti-MEM medium for a total of 50 µL of siRNA mixture. The diluted RNAiMAX mixture and diluted siRNA mixture were mixed (1:1 ratio) at room temperature for an additional 5 minutes. Next, 50 µL of this mixture was added to a well of 24-well plate, mixed with 450 µL growth

medium, and medium was not changed for 48 hours. Cells were harvested and assessed by qRT-PCR.

Short-interfering RNA (siRNA) knockdown via reverse transfection in stable liver (HepG2) cells

Pooled siRNA targeting human Foxa1 and human Foxa2 was obtained from Dharmacon. 3 pmol of siRNA targeting Foxa1 and 3 pmol of siRNA Foxa2 were diluted into 100 µL of OptiMEM without serum per well in a 24-well tissue culture-treated plate and mixed gently. 1 µL of Lipofectamine RNAiMAX transfection reagent was added to each well and mixed gently. The mixtures were incubated for 10-20 minutes at room temperature. 500 µL of cells were resuspended in complete growth medium (DMEM and 10% FBS) without penicillin-streptomycin (Penn/Strep) at 100 cells/µL, seeded into each well, mixed gently by on a plate rocker, and placed in a 37°C incubator for 48 hours. Cells were collected and prepared for either qRT-PCR or western blot. Controls included cells maintained in transfection medium without siRNA, and cells treated with scrambled siRNA, and cells treated with a single siRNA pool but not both.

shRNA sequences

Four shRNA sequences were tested and evaluated for their ability to knockdown both Foxa1 and Foxa2. Each shRNA was cloned into the host pLKO.1-puro vector. PLKO-shFoxa1#1 (Cat. #: 70095) and PLKO-shFoxa1#2 (Cat. #: 70096) were obtained from Addgene. shFoxa2#1 and shFoxa2#2 were obtained from the literature (Tang, Shu et al. 2011). The sequences were:

pLKO-shFoxa1#1: 5' – TCTAGTTTGTGGAGGGTTATT – 3';

pLKO-shFoxa1#2: 5' – GCGTAC TACCAAGGTGTGTAT – 3;

pLKO-shFoxa2#1: 5' – GCAAGG GAGAAGAAATCCA – 3';

pLKO-shFoxa2#2: 5' – CTACTCGTACATCTCGCTC – 3'.

shRNA Lentiviral generation and transduction

Third generation lentivirus was produced with the help of the Lentivirus core facilities at Roswell Park Cancer Institute, using the cloned pLKO.1 puro vector. Transduction at multiplicity of infection (MOI) 5-10 (titer of 10^7 particles per ml) was performed serially with shFoxa1 transduced first followed by shFoxa2 to engineer shFoxa1/2 cell lines. Cells were seeded at 25% confluency in a 24-well tissue culture-treated plate one day before transduction. Cells were washed once with

PBS and 250µL of OptiMEM, 20 µL of virus, and polybrene (8 µg/mL) were added to each well for a 24-well plate. The cells were incubated for 6 hours and complete growth medium without Penn/Strep was added to each well for a total volume of 1 mL for 24-well plate. The cells were incubated for 48 hours and media was exchanged for an antibiotic selection medium composed of complete growth medium with Penn/Strep and 1.75 µg/mL of puromycin. The cells were incubated in the selection medium for 7 days with a media change after 3 days and transduction was repeated with the second lentivirus. The cells were then assayed for gene expression with qRT-PCR and for protein expression with Western Blot. The negative control was HepG2 cells maintained in the same conditions as experimental without the virus. The positive control was done using the same transduction protocol but with a pLKO-scrambled shRNA lentivirus. Positive control cells were also selected for using 1.75 µg/mL puromycin for 7 days prior to analysis.

Engineering of stable stem cell lines bearing conditional shRNA vectors

To engineer conditional cell lines which can downregulate or restore Foxa2 expression, we employed the pSingle-tTS-shRNA vector in which doxycycline activates shRNA expression. The vector was grown in Stbl3 *Escherichia coli* with ampicillin as a bacterial antibiotic resistance. For shRNA cloning, the Foxa2#2 shRNA was inserted downstream of the tTS promoter region using the XhoI and HindIII restriction sites. For HepG2 cell transduction, cells were plated at approximately 100,000 cells/well in DMEM+10% FBS in a 24-well plate. Cells were transfected using FugeneHD (Promega) transfection reagent with a Fugene: DNA ratio of 4.5:1. To transfect 3 wells, a total volume of 450 µL of 0.020 µg/µL plasmid solution was made up in OptiMEM (9.9 µg of plasmid in 450 µL total volume). 45 µL of Fugene reagent was then added to each well and incubated at room temperature for 15 minutes before 150 µL of the plasmid solution complex was added to each well and mixed thoroughly. After 48 hours, the medium was changed to DMEM +10 FBS +1% P/S for expansion, or for creation of a stable cell line, cells were selected with DMEM + FBS + P/S +1000 µg/ mL G418 (Sigma-Aldrich).

shRNA knockdown in stable liver cells and stem cells conditional shRNA vectors

For both transient transfection and stable cell line creation, knockdown was induced with the addition of doxycycline (Dox) medium. The Dox medium contained a final concentration of 1-2 µg/ml, but lower doses were tested as reported. For conditional experiments, the Dox containing

media was added on ((+) Dox condition) for 3 days, and was then removed for 3 days, and cells were harvested both after Dox treatment and after Dox removal, and analyzed by qRT-PCR.

Immunofluorescence

Cells were seeded at 50,000 cells/well in a 24-well tissue culture-treated plate one day before staining. Cells were rinsed three times with PBS at a volume of 500 µL per well. All wash steps were performed with 500 µL PBS/well unless otherwise specified. Cells were fixed with 500 µL of 4% paraformaldehyde at room temperature under cell culture hood. Cells were washed again three times for 5 minutes each wash before incubating with 1% PBS-Triton X for 30-60 minutes for permeabilization. Cells were rinsed thrice more before incubation in 1% BSA/PBS (Blocking Buffer) for 1 hour. Primary antibody incubation was done overnight at 4°C at a 1:500 dilution in either blocking buffer or PBS. After overnight incubation, cells were washed 4 times for 15 minutes before 1-hour incubation with a fluorescently labeled secondary antibody, diluted 1:500 in either blocking buffer or PBS. After secondary antibody incubation, cells were washed 4 times for 15 minutes each using 500 µL of 0.1% PBS-Triton X, cells were rinsed once more with Milli-Q water before adding 500 µL of a 1:1000 Dilution of DAPI in PBS before imaging. Cells were Imaged with Zeiss Axiovision SE64.

Protein isolation and Bradford assay

On the day of protein collection, cells were washed once with PBS and then incubated in 300 µL of 0.05% Trypsin for 6-10 minutes. 500 µL of DMEM containing 1% v/v of Pen-Strep (P/S) and 10% fetal bovine serum (FBS) was added and cells were then collected and transferred to sterile microfuge tubes. Cells were then pelleted at 5,000 RPM for 30 seconds and cell pellet was resuspended in 300 µL of 1x PBS, in which cells were washed twice. For the last wash, cell pellets were resuspended in 250 µL of RIPA buffer containing Protease inhibitor and incubated on ice for 5 minutes. Cells were then spun at 4ºC for 15 minutes at 14,000 RPM and the supernatant was collected for assaying via Bradford assay. Bradford assays were conducted by diluting protein samples by a factor of 10 with Milli water,, and 5µL of each sample or known standard was platted into each appropriate microplate well of a 96-well plate. 250 µL of Coomassie protein assay reagent (Thermo Scientific) was added to each well and was left to mix on a plate shaker for 30 seconds. Absorbances were measured at 595 nm with a plate reader and a standard cure was

constructed using the average of triplicated samples and known protein concentrations, a best fit line constructed using a 3rd degree polynomial was used to calculate the unknown sample concentrations.

Western blot

Equal concentrations of protein were loaded in all lanes of a 10% polyacrylamide gel with a 4% stacking gels were used and prepared per manufacturer's instructions (Invitrogen). Pre-made, Bolt 4-12% Bis-Tris Plus 10-well gels (Invitrogen) were also used. Protein sample preparation involved using equal concentrations of 10-50 μg/well for a final volume of 25 μL, for use in a 10-well polyacrylamide gel. Samples were heated, loaded, and run in 1X SDS-PAGE buffer for 45 minutes at 125V (Bio-Rad Mini Gel Tank). Pre-made gels were run at 200V for 22 minutes. Next, gels were washed and placed in a gel transfer assembly (Mini Blot Module, Invitrogen) for transfer to a nitrocellulose membrane (Bio-Rad) for approximately 60 minutes. Some transfers took place using the iBlot2 transfer module (Invitrogen) using nitrocellulose transfer stacks (Invitrogen). The membrane was washed in MilliQ water 2 times for 5 minutes each before being placed in 15 mL of blocking buffer consisting of 5% milk powder in PBS-Tween (PBS pH 7.4 + 0.1% Tween 20) for 1 hour at room temperature, washed, and incubated overnight with primary antibody overnight, washed three times, and incubated for 1 hour at room temperature in HRP-conjugated secondary antibodies, which were diluted in blocking buffer at a 1:100-1:5000 dilution. After incubation in secondary antibody, membranes were washed again in PBS-Tween before being placed in 0.1 mL working solution (equal parts stable peroxide solution and Luminol/Enhancer solution) per cm^2 membrane for 5 minutes before being placed for imaging in ChemiDoc Gel Imaging System (Bio-Rad).

Microscopy

Cell lines and stem cell cultures were imaged with a benchtop microscope (EVOS fluorescent, phase contrast microscope, #AMEFC4300R) at 4x, 10x, and 20x or with the Zeiss Axiovision SE64 microscope. Images were acquired and stored and used to visualize cells.

RNA-sequencing

Total RNA integrity was determined with the Agilent Technologies Fragment Analyzer RNA assay system. The TruSeq Total RNA Stranded Library Preparation kit (Illumina) with rRNA removal was used for preparation of RNA sequencing libraries. Multiplexed libraries were individually quality controlled with the Fragment Analyzer and quantified using the Quant-iT ds DNA Assay Kit (Invitrogen). The libraries were pooled to 10 nM and diluted for qPCR using the KAPA Library Quantification Kit (Kapa Biosystems). Subsequently the pooled libraries were normalized to 4 nM based on qPCR values. All samples were sequenced on the NextSeq500 (Illumina) in midoutput mode producing 160 million paired end 75 bp reads, ~25 million per sample.

RNA-seq (Transcriptome) analysis

The raw fastq files were first aligned using the hisat2-2.1.0 GRCh38 genome and converted into sam files. The sam files were converted into bam files using samtools-1.6. FeatureCounts was used to count reads for genes from the GRCh38 genome to create a text file with the two controls and three experimental samples. The text file was imported into R where it was analyzed with DESeq2. The gene expression data was normalized with Rlog normalization in DESeq2. The built in prcomp function was used on the normalized data to find the first and second principal component analysis (PCA1 and PCA2). PCA1 and PCA2 were plotted using R plotting tools. The volcano plot figure was created by importing the same count data into R. The DESeq2 function lfc shrink was used to shrink the data using the contrast between the condition and experimental groups. The res function was used to obtain the adjusted p-value and log2-fold change for each gene. The log2-fold change was then plotted on the x axis and the -log10 of the adjusted p-value was plotted on the y axis. The genes were colored if the adjusted p-value was under 0.05. The genes in red represent upregulation and the genes in blue represent down regulation based on the log2-fold change. The heatmap was created with the gplots function heatmap.2 in R. The normalized number of reads for each gene were compared between samples to determine the color. Green represents relatively higher gene expression, while red represents relatively lower gene expression. The 683 genes with an adjusted p-value under 0.05 were plotted and divided into upregulating and downregulating groups and listed by increasing adjusted p-value for each group. Every seventh gene in the list was labeled. Unbiased hierarchical clustering was used to group the samples with the built in hclust function within heatmap.2. The 683 significantly expressed genes were analyzed with Gene Ontology

Panther classification system using the "Statistical Overrepresentation Test" to find significant biological processes, cellular components, and molecular functions. Default options were used in this analysis except the reference list was replaced by a cumulative gene list including all genes found to have any expression from the RNA-seq analysis. Significance was determined by a (FDR)-qval <0.05.

RNA-seq normalized enrichment score analysis

Normalized enrichment scores (NES) for differential gene sets were analyzed using Gene Set Enrichment Analysis (GSEA). Gene expression data was normalized with the median of ratios method in DESeq2. The normalized read counts of the genes were compared between the experimental group (shFoxa1/2 -/-) and control group with a Student's t-test method and the genes were arranged from most upregulated to most downregulated. The GSEA 4.0.3 software from the Broad Institute analyzed the list using the GSEAPreranked tool. All default settings were used except the chip platform field was set to "no collapse" and the max set size was set to 2000. The version 7.0 gene sets from Gene Ontology (GO) and KEGG were downloaded from the Broad Institute and used as the gene sets database. The results were filtered for p-value<0.05 , NES -1.5< or > 1.5, and either (FDR)-qval or (FWER)-pval<0.1. The results were grouped based on enrichment and were sorted based on "SIZE". The signal-to-noise method was also used to sort the genes instead of the Student's t-test and obtained similar results.

Identification of liver GRN promoter/enhancer binding sites

To determine if genes downregulated in the shFoxa1/2 condition were likely to be regulated by Foxa, HNF4α, or HNF1ß, we analyzed a candidate list of hepatic genes that were downregulated. Information available in the "Genomics" section for each gene of interest through genecards.org was used to identify transcription factor binding sites in the promoter/enhancer sequences for the gene. Listed in Table 3 are the liver GRN (Foxa, HNF4α, HNF1ß) and the non-liver GRN. The non-liver GRN column does contain TF binding sites for some liver GRN, but we have separated it this way based on our current analytic approach.

Statistics

For statistical comparison between two groups, Student's t-test was performed, and p-values equal to or less than 0.05 was considered to have statistical significance and are delineated within the text and figures/figure legends. Statistics for the RNA-seq analysis is detailed in those sections.

RESULTS:

Directed differentiation of hPSC towards definitive endoderm, gut tube endoderm, and the hepatic lineage under low oxygen conditions.

We aimed to develop a reproducible hPSC-differentiation protocol to understand regulatory processes at the levels of GRN that influence liver differentiation. We aimed to differentiate hPSC towards definitive endoderm and then towards hepatic progenitors, all under hypoxic (5% O_2) incubator conditions, which mimics physiological liver development (Van Blerkom, Antczak et al. 1997); (Okazaki and Maltepe 2006) as we have previously done in mouse ESC (Parashurama, Nahmias et al. 2008);(Cho, Parashurama et al. 2008); (Park, Cho et al. 2007). We employed a commercially available endoderm induction protocol (STEMDiff for definitive endoderm differentiation) followed by gut tube induction (Wang, Yue et al. 2015), and hepatic differentiation (Takebe, Sekine et al. 2013) (**Fig. 1A**). Morphological analysis demonstrated that endoderm progenitor cells on day 4 did not have characteristic bright cell borders and cuboidal appearance, as observed previously (Parashurama, Nahmias et al. 2008), but expressed the combination of definitive endoderm specific proteins Foxa2 and Sox17 (**Fig. 1B, panels 1-3**), and formed either small clusters of cells or elongated cells by day 12 of differentiation. At day 4, immunostaining indicated that approximately 85-95% of cells were definitive endoderm. Kinetic analysis of differentiation by gene expression for multiple cell populations demonstrated a downregulation of the pluripotent marker Oct4, upregulation and downregulation of the primitive streak marker brachyury, and an upregulation of the primitive streak/endoderm marker goosecoid, all of which indicated formation of definitive endoderm **(Fig. 1C).** Consistent with this, we observed statistically significant differences in Oct4, Foxa2, and Sox17 **(Fig. 1D)** by day 4, including other endoderm markers like GATA4 and CXCR4 (data not shown). Next, we carried out the differentiation protocol as shown by the addition of KGF (25 ng/ml) from days 5-7, the addition of FGF2 and BMP4 on days 7-10, and the addition of HGF, oncostatin, and dexamethasone from days 10-12 **(Fig. 1A)** (Takebe, Sekine et al. 2013)**.** We observed the maintenance of Foxa2, the downregulation of Sox17, consistent with gut tube formation, and the large upregulation of alpha-

fetoprotein (Afp) consistent with liver differentiation, but minimal activation of albumin gene, with no significant differences and high variability. To ensure the capable activation of albumin gene in hPSC, we performed qRT-PCR after teratoma formation for 8 weeks, and an abundant activation of albumin occurred, indicating there was no lack of differentiation potential of the cells **(Fig. 1E).** These data suggested that in our differentiation protocol entirely under low oxygen, the lack of albumin gene activation indicated that hepatic differentiation was minimal (Cascio and Zaret 1991); (Calmont, Wandzioch et al. 2006).

Regulation of endoderm and hepatic nuclear factor/gene regulatory network (GRN) genes in definitive endoderm.

To understand GRN expression and hierarchy during endoderm induction, and its potential role in endoderm and liver differentiation, and we developed an RNAi transfection assay in hPSC-derived endoderm, followed by qRT-PCR for components of the endoderm and liver GRN. We are particularly interested in the role for Foxa2, due its widely established role in embryo, gut tube, and liver development (Zaret 2002). We transfected both pooled siFoxa1 and siFoxa2 and compared them to negative scramble controls. We performed single knockdown studies, of Foxa1 or of Foxa2, and demonstrated that the target factor achieved partial knockdown, but the corresponding factor (Foxa1 for a targeted Foxa2 knockdown) did not, demonstrating that Foxa1 can compensate for Foxa2 as reported in mouse in vivo studies (Lee, Friedman et al. 2005) (data not shown). We demonstrate that targeting both Foxa1 and Foxa2 during endoderm induction results in a statistically significant decrease in a wide range of gut tube and liver transcription factors, including Foxa1, Foxa2, GATA4, Hex, HNF6, with a upregulation of HNF4α and no significant changes in HNF1ß **(Fig. 2A)**. We identified a downregulation of master TFs, including Sox17 and Gsc, and an upregulation of both Cdx2 (gut, intestine) (Watts, Zhang et al. 2011) and Pax6, a master factor for neuroectoderm and eye development **(Fig. 2B)**. Finally, we analyzed endoderm and liver differentiation genes, and demonstrated a down regulation of Cxcr4, a cell surface receptor that correlates with endoderm induction (Fig. 2C). Further, even though the absolute levels of Afp and Alb were low in definitive endoderm (Fig 1), we observed a downregulation, supporting the idea that combined gene silencing of Foxa1 and Foxa2 can downregulate activation of liver differentiation markers in human stem cells (Fig. 2C).

Development of a simplified, cultivation system under hypoxia for endoderm induction from stem cells.

We sought to develop a simplified endoderm and liver induction protocol, under low oxygen conditions, for study of GRN that that would enable: 1) robust morphology of endoderm progenitors, as we have observed previously (Parashurama, Nahmias et al. 2008), 2) activation of liver differentiation programs and robust liver morphology and gene expression, as we have observed previously (Cho, Parashurama et al. 2008), and 3) differentiation under low oxygen for improved modeling of liver development. We first developed an endoderm protocol with the aim of improving endoderm morphology while maintaining the endoderm TF expression. We adopted a protocol (Rezania, Bruin et al. 2013) in which activin and Wnt stimulate primitive streak induction from hPSC, and a high concentration of activin stimulates endoderm induction from hPSC. Using this approach, we obtained extensive cell death at low oxygen. We reasoned that low oxygen resulted in cell death, so 0.2% KOSR was added to the medium to promote cell survival. We tested each pluripotent stem cell line **(Fig. 3A)** for endoderm induction, employing activin/CHIR (Wnt canonical pathway agonist) followed by activin under hypoxic conditions, with 0.2% KOSR throughout (Fig. 3A-B). Pluripotent stem cells **(Fig. 3C, top panels)** expressing Oct4, and not expressing Sox17, demonstrated downregulation of Oct4 and upregulation of Sox17, Foxa2, and low amounts of alpha-fetoprotein as judged by immunocytochemistry **(Fig. 3C)**. At day 4, immunostaining indicated that approximately 85-95% of cells were definitive endoderm. Morphological analysis demonstrated that cell survival was promoted in the presence of 0.2% KOSR, and by day 4 we typically observed robust cellular morphology, with cuboidal cell morphology, bright cell borders, and cell sheets consistent with robust definitive endoderm **(Fig. 3D, Day 1-4)**. This morphology compared extremely favorably with the morphology from the previous protocol under hypoxic conditions **(compare Fig. 1B, left panel and Fig 4D, right panel)**. As expected, gene expression analysis via qRT-PCR demonstrated statistically significant decreases in Oct4, and corresponding statistically significant increases in Foxa2 and Sox17, with clear trends in upregulation of Gata4 and Gsc **(Fig. 3E)**. Finally, we analyzed the expression kinetics and demonstrated Oct4 downregulation and brachyury upregulation followed by downregulation **(Fig. 3F)**, consistent with the formation of robust definitive endoderm by day 4, under hypoxic conditions.

Spontaneous differentiation of gut tube progenitors results in higher albumin activation

Foxa2 primes albumin gene activation (mRNA expression) by binding silent, inaccessible chromatin at the albumin regulatory regions in gut tube/hepatic endoderm progenitor cells, (Zaret 2002) and the abolishment of Foxa1/2 in hepatic endoderm during liver development completely inhibits liver organ development (Lee, Friedman et al. 2005). Further, albumin activation occurs prior to visible cell differentiation or organ development (Cascio and Zaret 1991). This supports the hypothesis that Foxa1/2 regulates albumin activation in murine systems, but this needs to be determined in human stem cell models of albumin gene activation. To test this, we first developed a novel differentiation protocol under low oxygen conditions **(Fig. 4A)**. The protocol involves induction of endoderm as in Fig. 3, differentiation of gut tube endoderm progenitor cells with KGF/Fgf7 from day 5 to day 7, and continued spontaneous differentiation towards hepatic progenitor cells between days 7 and 14, with no additional growth factors used during hepatic differentiation **(Fig. 4A and Fig. 1A)**. Spontaneous differentiation of mouse ESC to hepatic progenitors has been reported, with the absence of growth factors resulting in robust hepatic progenitors (Novik, Maguire et al. 2006), but this has not yet been reported in human stem cells. Notably, we switched the medium to SFD medium (Gadue, Huber et al. 2006) on day 4, which is a defined medium containing Ham's F12, N2, and B27 **(Fig. 4A)**. Morphological analysis demonstrated the induction of robust cells from day 4 through day 14 **(Fig. 4B)**. The cells on day 14 appear large and cuboidal, with bright cell borders **(Fig. 4B)**. We compared gene expression analysis of our data with the spontaneous, growth factor-free hepatic differentiation protocol to the growth factor containing protocol. We found that activation of albumin expression was higher in the spontaneous differentiation condition than in the growth factor containing condition **(Fig. 4C,** $P = 0.21$**)**, while Cdx2, a gut tube and intestinal marker, was expressed at varying amounts, and was not significantly different than the growth factor containing condition (Fig. 4D).

Foxa1/2 controls activation of albumin gene in a novel liver differentiation protocol under hypoxic conditions.

To understand how Foxa1/2 control albumin gene expression in gut tube progenitor cells, we performed cloning of scrambled shRNA and two shRNA sequences targeting Foxa1 and Foxa2, respectively. We engineered 4 separate lentiviruses bearing 2 separate shRNA sequences for Foxa1 and 2 separate shRNA sequences for Foxa2, for a total of 4 viruses (Foxa1 #1, Foxa1 #2, Foxa2

#1, Foxa2 #2) each of which could be selected with puromycin within the lentiviral PLKO vector. In a series of experiments (data not shown), we determined that shFoxa1 #2 and shFoxa2 #1 resulted in the most downregulation. We transduced scrambled virus alone at a MOI of 10 versus shFoxa1#2 on day 6, followed by shFoxa2#1 on day 7, in hPSC-derived gut tube endoderm. Cells were cultivated until day 14 and analyzed by qRT-PCR or western blot. We did not perform antibiotic selection. Nevertheless, we saw major differences between control and shRNA Foxa1/2 -/- conditions in terms of morphology. We performed gene expression analysis for Foxa1/2 and for albumin. Foxa2 levels did not show significant changes but trended downwards, while Foxa1 downregulation was significant **(Fig. 4E)**. Importantly, albumin gene was significantly downregulated when normalizing the data to pre-transfection across all experiments **(Fig. 4E)**. Finally, Western blot demonstrated that Foxa2 is indeed downregulated in gut tube endoderm progenitor cells **(Fig. 4F)**. We conclude that Foxa1/2 regulates albumin in hPSC-derived, gut tube progenitor cells that are differentiated towards liver.

Foxa 1/2 regulates the endoderm and liver GRN, and liver differentiation genes, in a stable liver cell line.

We hypothesized that if both siRNA and shRNA targeting both Foxa1/2 regulate albumin expression in the developing gut tube differentiated towards the liver, perhaps they can regulate albumin in established models of functional human liver cells. As human hepatocytes are difficult to isolate, culture, and have donor- dependent differences, we chose to apply this approach to human HepG2 cells, a well-established model of gene regulation, which expresses the immature marker alpha-fetoprotein, in addition the albumin. To understand how knockdown of both Foxa1 and Foxa2 can block differentiation, we employed siRNA knockdown, with pooled siRNA against both Foxa1 and Foxa2, in human stable liver cell lines (HepG2). In a series of optimization experiments, we found: 1) unaltered morphology after transfection at 48h **(Fig. 5A)**, 2) combined pools but not individual pools of siRNA lead to a ~50% knockdown **(Fig. 5B-C)**, and 3) scrambled siRNA had no effect on major regulatory transcription factors and markers of differentiation (Fig. 5D). Importantly, we found that combined Foxa1 and Foxa2 knockdown (siFoxa1/2 -/-) was possible at a total of 6 pmol (3 pmol each) and was associated with a ~50% albumin knockdown in liver progenitor cells **(Fig. 5D)**, indicating regulation of the albumin gene in HepG2 cells. We also observed an upregulation of E-Cad, with no major changes in Hex or other markers with

increasing dose **(Fig. 5D)**, and determined that 3 pmol for each siRNA pool was the ideal amount for transfection. We performed single knockdown studies, of Foxa2 or of Foxa1, and demonstrated that that target factor achieved partial knockdown, but the corresponding factor (Foxa1 for a targeted Foxa2 knockdown) did not, suggesting that Foxa1 can compensate for Foxa2 as reported in mouse *in vivo* studies (Lee, Friedman et al. 2005) (data not shown).

We then performed knockout studies in a series of experiments, with scrambled siRNA for control and double knockdown using pooled siRNA for Foxa1/2. These studies demonstrated that knockdown of Foxa1/2 resulted in a downregulation of albumin, but not alpha-fetoprotein **(Fig. 5E)**. Western blot studies at 48 hours confirmed knockdown of Foxa2 and albumin, suggesting that liver differentiation was partially inhibited **(Fig. 5F-G)**, and this was also supported by immunofluorescence analysis of albumin **(Fig. 5G)**. To further understand the effects of Foxa1/2 on differentiation in the stable liver cell line, we performed analysis of core regulatory transcription factors and other differentiation markers including pluripotency (Oct4, Nanog), neuroectoderm (Pax6), mesoderm (BMP4), endoderm (Sox7, Gata4, Cxcr4), pancreas (Pdx1), epithelial state (E-Cad, N-Cad), and gut tube/liver TF (Hex, HNF4α, HNF6). We identified trends in the data, including increased Pdx1, BMP4, E-Cad, and N-Cad **(Fig. 5H)**. Compared to scrambled controls, we found statistically significant upregulation of Nanog, Pax6, Sox7, and CXCR4, suggesting that these genes were activated in the Foxa1/2 -/- cells **(Fig. 5I)**. Further, we observed a statistically significant downregulation in Hex, a major regulator of liver differentiation (**Fig. 5I)**. Finally, we performed analysis of other differentiation markers and demonstrated trends in Pdx1 upregulation, BMP4 upregulation, and E-Cad upregulation **(Fig. 5J)**. These siRNA studies demonstrate the widespread changes in key master factors of germ layers, and differentiation genes, when both Foxa1/2 is knocked down in the liver cell line.

Lentiviral transduced shFoxa1 and shFoxa2 demonstrate a block in liver differentiation genes in a stable liver cell line

To determine if continuously expressing shRNA for Foxa1/2 would have a similar or a more extended phenotype, we cloned shRNA for both Foxa1 and Foxa2, and performed double transduction followed by antibiotic selection to engineer a stable liver cell line with Foxa1/2 -/- phenotype. Optimization experiments were performed for each virus **(Fig. 6A)**, confirming that shFoxa1 #2 and shFoxa2 #1 viruses resulted in downregulation of their target genes. We again

performed analysis of core regulatory transcription factors and other differentiation markers including neuroectoderm, (Pax6), endoderm (Gsc, Cxcr4), pancreas (Pdx1, Sox 9), biliary (CK19), and the endoderm/hepatic nuclear network (Foxa3, HNF4α, HNF1ß, Hex, Gata4, Tbx3). The data overwhelmingly demonstrated a statistically significantly decrease in albumin, a 0.87-fold decrease in signal the major gene that signifies liver differentiation **(Fig. 6B)**, and trending increases in alpha-fetoprotein (2-fold), Cxcr4 (4-fold), and CK19 (11-fold). Additionally, we observed statistically significant upregulation in Gsc expression (1.5-fold), Pdx1 downregulation a (0.84-fold), and trending increases in Pax6 and decrease in Sox 9 expression **(Fig. 6C)**. Other than the significant downregulation of both Foxa1 and Foxa2, we observed statistically significant changes in core members of the endoderm/gut tube/liver, including HNF4α (0.62-fold decrease), HNF1β (0.78-fold decrease), Hex (0.72-fold decrease), Tbx3 (0.36-fold decrease) **(Fig. 6D)**. We also observed a statistically significant upregulation in Foxa3 and GATA4 **(Fig. 6D)**. Western blot confirmed our gene expression analysis **(Fig. 6E)**. This analysis indicates a strong downregulation of liver differentiation genes and core gene regulatory network that establishes a liver cell fate, with a concomitant increase in early liver de-differentiation markers like Foxa3 (hepatic endoderm), alpha-fetoprotein (early hepatic progenitor), Cxcr4 (endoderm progenitors), an increase in Pax 6 (neuroectoderm), and an increase in biliary marker (CK19). These data demonstrate a coordinated decrease in liver differentiation, and evidence of immature states (hepatic endoderm) and alternate fates (biliary, neuroectoderm) consistent with a differentiate block. Importantly, our shFoxa1/2 -/- , double knockout cell line demonstrated no change in morphology as a function of scrambled virus **(Fig. 6F, top row)**. Further, we observed no overall change in morphology of colonies in the shFoxa1/2 -/- cell line, although at the edges of colonies in low density culture, at times we observed flat, elongated cells **(Fig. 6F, bottom row)**. Taken together, our data suggested that Foxa1/2 knockdown resulted in global response of repressed or blocked liver differentiation, upregulation of markers for alternate cell fates, and de-differentiation.

A conditional liver cell line demonstrates reversible control of master liver transcription factors

To determine whether the control of Foxa1/2 over downstream liver genes is reversible, we engineered a conditional Foxa1/2 cell line. We engineered the Dox-on, pSingle-tTS-shRNA

vector, in which shRNA for Foxa1/2 was controlled by the addition of Doxycyline (Dox). After cloning in the shFoxa2 #2 shRNA and scrambled shRNA into the pSingle vector, we engineered stable cell lines by DNA transfection followed by antibiotic selection in G418 **(Fig. 7A)**. We varied Dox concentration and determined that 1.5 ng/ml was optimal in terms of toxicity and knockdown response. We performed addition of Dox (1.5 ng/ml) for 3 days, followed by removal for an additional 3 days, and performed qRT-PCR on days 0, 3, 6. In control cell lines with scrambled vector, we observed no changes in gene expression (data not shown). After Dox addition, as expected, we observed downregulation of Foxa1, Foxa2, albumin, alpha-fetoprotein, Gata4, Hex, HNF1ß, and HNF4α **(Fig. 7B)**. After Dox removal for 3 days, we observed upregulation of Foxa1, Foxa2, albumin, Gata4, Hex, a more moderate upregulation of HNF4α, and a continued downregulation of Afp and HNF1ß **(Fig. 7B)**. Thus, in most cases, the global changes were reversible. The morphological changes observed in the stable shFoxa1/2 -/- cell line **(Fig. 6F, lower right)** were present during Dox addition, a convenient method to demonstrate knockdown occurred. These data suggest that much of the effect of Foxa1/2 knockdown on HepG2 cells could be reversed over 3 days, but there may be a hysteresis effect by day 6 of culture for HNF1ß and alpha-fetoprotein.

Whole transcriptome analysis of Foxa1/2 (-/-) Hep G2 cell phenotype demonstrates increased stem cell/ cell differentiation genes of alternate fates, and downregulation of metabolic genes

We hypothesized that a differentiation block affecting master TF's of hepatic differentiation in shFoxa1/2 -/- cells may have had global, genome-wide affects. To understand the differentiation block at the level of the genome, we performed whole transcriptome analysis on 2 control and 3 experimental conditions. Principal component analysis figure shows significant clustering of the experimental values and the control values along the PC1 axis **(Fig. 8A)**, shown as orange squares (experimental, shFoxa1/2 -/-) compared to controls, shown in blue triangle. A volcano plot **(Fig. 8B)**, demonstrates 448 upregulated genes (green) and 235 downregulated genes (red) between the two conditions (**Supp. File 2**). Each circle in the plot represents a color, or differentially expressed gene, if their adjusted p value was below 0.05, using a log2-fold change in expression levels as the threshold for significance. We then performed a heat map analysis, using unsupervised, hierarchical clustering, in which each gene was normalized and ranked by expression change,

demonstrating coordinated upregulation (green) and downregulation (red) in all genes. Labeled are 1 out of every 6 genes. Validating the RNA-seq data, we found Pax6 in the upregulated gene set, and albumin in the downregulated gene set, both of which we had identified by qRT-PCR **(Fig. 8C)**. A full list of the upregulated and downregulated genes is provided **(Supp. File 2)**. In comparing shFoxa1/2 -/- to HepG2 cells, we identified the top 30 downregulated genes (**Table 1**) and the top 30 upregulated genes (**Table 2**).

Among the top upregulated genes included genes involved in axon guidance and vascular development (NRP2, neuropillin 2), kidney epithelial function (Cdh16, cadherin 16), kidney cell interactions (SYNPO, synaptopodin), and cardiac development (Alpk2 (alpha kinase 2)). Interestingly, these data are consistent with a meta-analysis of hPSC-liver differentiation studies in terms of factors that are aberrantly upregulated during liver differentiation (Raju, Chau et al. 2018). Among the top 25 downregulated genes include those involved in energy metabolism (PPARGC1A (PPARG coactivator 1alpha)), liver development and liver function (albumin), and drug and steroid metabolism (AKRC12 (Aldo-Keto Reductase Family 1 Member C2)). We analyzed all the normalized RNA-seq data, not just differentially regulated genes, with a t-test to sort the genes from most upregulated to most downregulated. We analyzed this list and t-score values with gene sets from Gene Ontology (GO) and KEGG using the Gene Set Enrichment Analysis (GSEA) software to find significantly enriched gene sets based on stringent criteria: normalized enrichment score (NES) and false discovery rate (FDR) **(Fig. 8D, Supp. File. 4)**. Amazingly, amongst the pathways, we observed many processes associated with stem cell differentiation and cell differentiation. For processes with upregulated genes, we observed over 500 genes associated with epithelial cell differentiation, followed by processes associated with skin development, striated muscle differentiation, eye development and neuron migration **Fig. 8D, Supp. File. 4)**. Further, for processes with downregulated genes, including those associated with stem cell differentiation, Anterior-Posterior specification, **(Fig. 8D, Supp. File. 4)**. Overall, this suggests that the downregulation of shFoxa1/2 resulted in a downregulation of liver GRN, and resulted in a global block liver differentiation, while greatly altering differentiation state. The alteration of differentiation was characterized by clear activation of genes for alternate cell fates, as observed in a recent metanalysis of human stem-cell derived hepatic cells (Raju, Chau et al. 2018). We wanted to know if any key downregulated genes were controlled by the hepatic nuclear factor network. After reviewing the downregulated genes and their functions, we hand-picked a

series of genes of interest, based on hepatic expression and de-differentiated function, and examined whether their regulatory regions had a Foxa1/2, HNF4α, or HNF1 binding site. Nearly all of the list did, including APO-M, PPARGC1A, AKR1C1, IRS2, albumin, FGFR1, TGFBR3, TWNK, DLK1, NRARP, FASTKD5, UCA1, APOA4, APOC3, CYP24A1, CYP1A1, KRT7, CYP4F2, CYP4F2, CYP19A1 had a promoter/enhancer with a binding site for Foxa, HNF4α, or HNF1, again suggesting global downregulation by these factors (**Table 3**).

DISCUSSION

How developmental GRN function to control differentiation remains a central question in cell and developmental and stem cell biology, and the hPSC-derived liver differentiation is an excellent model to answer this question. GRN are known to control germ layer (neurectoderm, mesendoderm and endoderm induction), and cell specification from endoderm-derived gut tube, including the liver, pancreas, and lung. Further, GRN control differentiated liver functions. Foxa2 has a major role in controlling development, in part through controlling GRN. Studies of the role of Foxa2 in human cells are lacking compared to mouse, and many studies employing Foxa2 do not account for the compensatory functions of Foxa1. Here, we employed siRNA and shRNA to target Foxa2 and its compensatory molecule, Foxa1, in hPSC-derived endoderm, hPSC-derived gut tube endoderm progenitor cells differentiated along the liver lineage, and in a human stable liver cell line. In hPSC-derived endoderm, we demonstrated that siFoxa1/2 (-/-) phenotype is characterized by wide-spread changes in mesendoderm and endoderm GRN, suggesting repression of endoderm cell fate. To determine the role of Foxa1/2 in activation of the liver differentiation program, we performed shRNA-mediated knockdown in hepatic-fated cells. To do this, we first developed a novel model of endoderm induction and subsequent spontaneous liver differentiation model that exhibited robust morphology and lineage appropriate gene signatures under hypoxic conditions, which better mimic liver organ development. The shRNA Foxa1/2 (-/-) phenotype in gut tube endoderm, initiated on day 6, resulted in the downregulation of the albumin gene on day 14 culture, indicating potent control over albumin gene activation in gut tube progenitors differentiating towards the liver lineage. To further understand the effects of Foxa1/2 on liver differentiation, we employed the well-studied, stable liver cell line, HepG2. In HepG2 cells which exhibit liver-specific gene expression, we observed that shFoxa1/2 knockdown results in

widespread changes in liver gene expression, likely mediated by robust downregulation of liver-specific GRN components, including HNF4α, HNF1ß, Hex, and Tbx3. RNA-seq analysis demonstrates large changes to the transcriptome, and pathway analysis shows coordinated changes in stable liver cell lines, including upregulation of cell differentiation genes of other lineages. These findings provide novel insights into Foxa1/2 complex functions in various stages of differentiation, including endoderm, gut tube, and liver.

One of the main findings in this study was the identification of that Foxa1/2 downregulates liver GRN and liver-specific genes in differentiating hPSC-derived gut tube endoderm and in stable liver cell lines. The shRNA Foxa1/2 (-/-) phenotype in a stable liver cell line, to our knowledge, has not previously published. Interestingly, we observed statistically significant downregulation of all major liver GRN components including HNF4α, HNF1ß, Hex, and Tbx3, which is non-obvious result, since Foxa1 and Foxa2 knockout have been shown to have minimal effects on late stage fetal hepatocytes *in vivo* (Li, White et al. 2009). We believe this is effectively a differentiation collapse, as has been reported recently in adult hepatocytes (Reizel, Morgan et al. 2020). Foxa1/2 has been studied using human liver cell lines for analysis of gene regulation of alpha-fetoprotein (Kanaki and Kardassis 2017) and lipoprotein lipase (Lehner, Kulik et al. 2007). These studies both employed siRNA targeting Foxa2 in human cell lines, but not both Foxa1/2. Kanaki *et al*. demonstrated effects of Foxa2 downregulation on downregulation of albumin and transferrin, as well HNF4α and HNF1ß in HepG2 cells, but reports a decrease in HNF6, which we did not observe. Their investigation demonstrated that they also inadvertently also targeted the Foxa1 gene. However, these studies did not employ shRNA to induce sustained downregulation, nor did they investigate reversibility of the response using time-based assays siRNA or conditional shRNA vectors. Nonetheless, they demonstrated direct Foxa2 binding to both AFP regulatory regions, suggesting control over AFP expression, and they demonstrated downregulation of AFP, whereas our shRNA data demonstrated an upregulation of AFP (Kanaki and Kardassis 2017). Interestingly, when Foxa2 was knocked out of adult livers (by E18.5), there was no functional consequence (Sund, Ang et al. 2000), probably due to compensation by Foxa1, while Foxa1/2 knockout in embryonic liver at E16.5, exhibited phenotypic consequences in the biliary tract, did not demonstrate any changes in hepatocyte morphology, ultrastructure, or function, other than a 10-fold increase in alpha-fetoprotein expression. As mentioned above, adenovirus-mediated knockout of Foxa1 and Foxa2 in a Foxa3-deficient (Foxa1/2/3 -/-/-) adult liver demonstrated a

strong phenotype regarding strong collapse of adult liver and developmental liver genes (Reizel, Morgan et al. 2020). Our data is in the HepG2 shFoxa1/2 (-/-) cell line is partially consistent with the data in the Riezel *et al.* study, in that we observe downregulation of cytochrome P450 genes and apolipoproteins (**Table 3, Supp. File 4**), as well as albumin, and is consistent with a differentiation collapse or reversal occurs in the HepG2 cells. A new idea introduced in the Riezel et al. study, study is that histone modifications that predict gene activity (H3K27) are affected by the absence of Foxa factors, suggesting Foxa1/2 factors are required to maintain these epigenetic states in adult cells. These results are in contrast to studies of epigenetic marking during hPSC-derived gut tube differentiation towards pancreas, which demonstrate that H3K27 are not affected by Foxa1 knockdown (Wang, Yue et al. 2015). It is unclear how epigenetic mechanisms play a role in our study. In our shFoxa1/2 (-/-) cells, we observed upregulation of Gata4, which is not only and endoderm and mesoderm marker, but has a strong role in in pioneer factor activity at albumin (liver) genes (Bossard and Zaret 1998). Further, we observe an upregulation of Foxa3, which may play a role in maintaining chromatin state, as described in the Riezel et al. study. Future work would involve further gene silencing or CRISPR-based gene knockout strategies to fully understand the Foxa1/2 (-/-) phenotype, and CHIP-Seq analysis of Gata4, Foxa1/2/3, and histone modifications to further reveal clarify the underlying phenotype.

Another key finding of the Foxa1/2 -/- phenotype in our study is that it provides some relevance for our study for human stem cell research, which was also unexpected. Although there have been a series of key studies in the field that have advanced our knowledge of stem cell differentiation, a meta-analysis of 8 seminal studies hPSC-liver differentiation lists several major unsolved problems (Raju, Chau et al. 2018): 1) the resulting hPSC-HEPS are fetal-like, 2) they express a set of undesired genes, including cardiac mesoderm and neural/eye development 3) they are not fully functional metabolically, 4) they are missing signals from liver development, and 5) they express increased amount of intestinal master factors. Interestingly, the shFoxa1/2 (-/-) phenotype in stable liver cells also demonstrates upregulation of genes cardiac and neural/eye morphogenesis **(Fig. 8D).** Further, we also observed evidence of changes in genes associated with metabolism (i.e. PPARGC in Table 1) which were also identified in the metanalysis (Raju, Chau et al. 2018). A final connection between our study and the metanalysis is the fact that we observed increased Cdx2 expression in the shFoxa1/2 (-/-) liver cells **(Fig. 6D),** which was observed in the Raju *et al.* study. Further supporting this, we also observed increased Cdx2 expression when we

differentiated cells towards liver under low oxygen condition (**Fig. 4D**), and in the siFoxa1/2 (-/-) condition in human stem cell-derived endoderm **(Fig. 2D)**. Thus, engineering and analyzing shFoxa1/2 (-/-) cells, we are recapitulating several aspects of a phenotype seen in all major seminal studies of human stem cell differentiation (Raju, Chau et al. 2018). One way of interpreting this surprising connection is that the data suggests that Foxa1/2 mis-regulation is tied to problems in hPSC-liver differentiation by activation of inappropriate developmental factors in adult cells, loss of metabolic functions, and increased intestinal gene expression. This can be studied by further analysis of the Foxa1/2 (-/-) phenotype during liver differentiation.

We were particularly interested in how the Foxa1/2 -/- phenotype affects GRN in endoderm progenitors. It can be conceived of that there are at least 5 sets of GRN that are initiated and maintained, and partially downregulated during liver differentiation from hPSC. These include the mesendoderm GRN, endoderm GRN, gut tube GRN, foregut endoderm GRN, and hepatic GRN. In our experiments, siRNA was transfected on Day 1, which suggests that it is effective by Day 3-4 of culture. Our siRNA data of endoderm induction are consistent with the concept that in mesendoderm, Foxa1/2 controls initiation of the endoderm GRN, as demonstrated by our data demonstrating that Foxa1/2 downregulation results in decrease of Sox 17, Gsc, Hex, and HNF6 **(Fig. 2A).** Since endoderm is induced by Day 4, our observation of downregulation of mesendoderm and endoderm GRN suggests that Foxa1/2 plays a role in their activation and/or maintenance. Previous studies of GRN during mouse visceral endoderm differentiation of ESC in Foxa2 -/- cells demonstrated that Foxa2 is upstream of Foxa1, HNF1ß, and HNF4α, and thereby controls differentiation and metabolism (Duncan, Navas et al. 1998) in this population. Foxa2's role in activating the endoderm transcription network could explain why Foxa2 mutants in the embryonic germ layer lack the foregut (Cirillo, Lin et al. 2002);(Hallonet, Kaestner et al. 2002). Our data suggests that Foxa1/2 has time-dependent functions, and its inhibition may inhibit the initiation of different GRN at different time. Another factor that appear to be responsible for initiating GRN in human endoderm progenitors include GATA6 (Fisher, Pulakanti et al. 2017). Moreover, another factor that appears to initiate liver induction during hPSC-differentiation is HNF4α (DeLaForest, Nagaoka et al. 2011);(DeLaForest, Di Furio et al. 2018). Further studies investigating Foxa1/2 functions during hPSC-differentiation towards gut tube progenitors like liver, lung, and pancreas will continue to illuminate this subject.

To our knowledge, this is one of the first studies reporting downregulation of albumin in response to Foxa1/2 knockdown in during human stem cell differentiation. In early gut tube progenitors that have yet to be patterned, it is possible that Foxa1/2 knockdown would result in lack of formation of gut, and Cdx2 downregulation, while at later stages of gut tube formation, it may be possible to observe that Cdx2 is upregulated due to the loss of repressive role of Foxa2, as reported in murine studies (Watts, Zhang et al. 2011). Gut tube progenitors also exhibit complexity in their epigenetic state. It has been reported that endoderm progenitors exposed to KGF for two days begin modifying chromatin and undergo de novo activation of developmental enhancers, and that Foxa1 occupies these enhancers (Wang, Yue et al. 2015). As mentioned above, and pointed out in the literature (Kaestner 2015), knockout of Foxa1 but not both Foxa1 and Foxa2 (and/or Foxa3) in gut tube endoderm, was conducted in this study, and therefore it remains unknown whether Foxa2 binding precedes chromatin remodeling, and directs chromatin remodeling, at developmental enhancers. Our data in gut tube endoderm suggest that shRNA for Foxa1/2 downregulation on day 6, can lead to the lack of activation of albumin gene by day 14, suggesting that Foxa1/2 knockdown may prevent enhancer priming at later stages, perhaps due to its pioneer activity. Further studies and analysis of the human gut tube progenitors is needed, and our approach of spontaneous differentiation under low oxygen conditions, may enable a slower differentiation of gut tube progenitors than traditional studies, and thereby enabling further dissection of the expression and progression of GRN.

An important point here is that we have developed a growth factor-free liver differentiation system which activates the albumin gene. This has been previously done in murine systems (Novik, Maguire et al. 2006). The theoretical basis for this system is that GRN can spontaneously move forward in the presence of nutrient rich medium, and that default differentiation programs exist. To our knowledge, this is one of the first such systems that have been published for hPSC. It is possible that in existing liver cultivation systems, addition of growth factors when cells are not competent to receive them results in mixed cell signaling rather than liver induction. However, future work is needed to improve this system. First, the mechanism of spontaneous differentiation protocol of liver progenitors from stem cell-derived gut tube endoderm is not yet known. Our data suggests perhaps that mesoderm subsets are generated at low levels and provide signals to the cells to activate liver differentiation. Another possibility is that the liver can be chosen as a default pathway and can be established spontaneously in the absence of signals. A third possibility is that

the gut tube chromatin modeling continues and enables activation of albumin. Further characterization of this protocol, including maturation and transplantation, will be needed to establish how robust these cells in comparison to other protocols. It is also not yet clear why the addition of growth factors represses hepatic differentiation in our low oxygen protocols, and the mechanisms of spontaneous differentiation and repression of differentiation need to be further delineated.

There are limitations to our study worth mentioning. With our utilization of RNAi, we are not obtaining the complete Foxa1/2 (-/-) phenotype, which also is of obvious interest. However, our approach of siRNA and shRNA results in decreased Foxa1/2 levels, which have been shown to be of interest in the setting of nonalcoholic liver disease (Moya, Benet et al. 2012) and fibrosis (Wang, Yao et al. 2017). Future studies need to employ time-dependent CRISPR knockout of Foxa1 and Foxa2, at different times, to fully understand the Foxa2 phenotype. We employed antibiotic-based selection to generate our stable cell lines of interest, both with lentivirus (shRNA) and nonvirally (conditional cell line). While it is possible that antibiotic selection resulted in some changes in cell change phenotype, we did not observe any measurable changes. Nonetheless, future studies will employ reporter-based strategies can be performed with ease and less perturbation of cell state. While we focused on cells lines, future studies will focus on phenotypes. The extremely interesting and relevant phenotype we identified in stable liver cell line can be further dissected with single cell RNA-seq, as it is likely that the response is not homogeneous in all the cells. For example, we see upregulation of master factors of stem cell state like Nanog and as well as CK19.

Here, we have provided evidence that time-dependent GRN networks can be controlled by Foxa1/2 in human stem cells, at the level of endoderm induction and hepatic endoderm differentiation. Further, we find that Foxa1/2 phenotype in stable liver cells (HepG2), a key model for liver GRN studies, has potent, genome-wide effects in blocking liver differentiation and seemingly reversible effects on cells. It remains to be seen how our understanding of GRN can enable the improved differentiation and maturity of hepatic progenitors, and of alternate cell fates like instestine, pancreas, thyroid, and lung. Further understanding of GRN, like compensatory mechanism, concentration effects, and reversibility will likely alleviate challenges in engineering transcriptionally complex hepatocytes and pancreatic islets from hPSC and help coordinate appropriate epigenetic and genetic changes that occur during differentiation.

FIGURE LEGENDS: (Under Inserted Figures)

TABLES:

TABLE 1. Top 30 Upregulated genes in shFoxa1/2 (-/-) cells compared to HepG2

Gene Name	**Symbol**	**p-value**
neuropillin 2(NRP2)	NRP2	7.36E-22
cadherin 16(CDH16)	CDH16	1.47E-21
sodium channel epithelial 1 alpha subunit(SCNN1A)	SCNN1A	1.16E-16
versican(VCAN)	VCAN	2.37E-16
alpha kinase 2(ALPK2)	ALPK2	4.86E-16
sestrin 3(SESN3)	SESN3	5.82E-16
synaptopodin(SYNPO)	SYNPO	1.08E-13
glutaminyl-peptide cyclotransferase(QPCT)	QPCT	2.48E-13
L1 cell adhesion molecule(L1CAM)	L1CAM	2.87E-13
integrin subunit alpha 3(ITGA3)	ITGA3	3.24E-13
urothelial cancer associated 1 (non-protein coding)(UCA1)	UCA1	1.06E-12
serine peptidase inhibitor, Kazal type 1(SPINK1)	SPINK1	1.15E-12
extracellular leucine rich repeat and fibronectin type III domain containing 1(ELFN1)	ELFN1	2.73E-11
Kruppel like factor 7(KLF7)	KLF7	2.90E-11
dipeptidyl peptidase 4(DPP4)	DPP4	1.57E-10
troponin I2, fast skeletal type(TNNI2)	TNNI2	3.70E-10
bromodomain adjacent to zinc finger domain 2B(BAZ2B)	BAZ2B	3.93E-10
transgelin 2(TAGLN2)	TAGLN2	6.12E-10
ATP binding cassette subfamily A member 4(ABCA4)	ABCA4	1.39E-09
SET domain containing lysine methyltransferase 7(SETD7)	SETD7	3.64E-09
olfactomedin like 3(OLFML3)	OLFML3	6.00E-09
solute carrier family 1 member 2(SLC1A2)	SLC1A2	8.19E-09
Coiled-Coil Domain Containing 198	CCDC198	1.21E-08
filaggrin(FLG)	FLG	1.46E-08
shroom family member 4(SHROOM4)	SHROOM4	2.44E-08
HECW2 Antisense RNA 1	HECW2-AS1	4.18E-08
apolipoprotein A4(APOA4)	APOA4	4.61E-08
tripartite motif containing 29(TRIM29)	TRIM29	4.61E-08

TABLE 2. Top 30 Downregulated genes in shFoxa1/2 (-/-) cells compared to HepG2

Gene Name	*Symbol*	*p-value*
metallothionein 1G	MT1G	1.32E-12
aldo-keto reductase family 1 member B10	AKR1B10	5.16E-12
metallothionein 1E	MT1E	7.51E-11
aldo-keto reductase family 1 member C1	AKR1C1	2.13E-09
alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase	MGAT5	2.17E-09
cadherin like and PC-esterase domain containing 1	CPED1	1.24E-08
pannexin 2	PANX2	4.61E-08
aldo-keto reductase family 1 member C2	AKR1C2	3.41E-07
dual specificity phosphatase 7	DUSP7	9.34E-07
Kruppel like factor 15	KLF15	9.65E-07
acyl-CoA dehydrogenase short/branched chain	ACADSB	2.71E-06
lectin, mannose binding 1	LMAN1	3.11E-06
PPARG coactivator 1 alpha	PPARGC1A	5.70E-06
aldo-keto reductase family 1 member B15	AKR1B15	1.49E-05
solute carrier family 22 member 31	SLC22A31	2.67E-05
TBC1 domain family member 4	TBC1D4	3.26E-05
C2 calcium dependent domain containing 2	C2CD2	3.59E-05
cingulin like 1	CGNL1	5.32E-05
solute carrier family 6 member 14	SLC6A14	5.32E-05
RNA, variant U1 small nuclear 31	RNVU1-31	6.62E-05
lysozyme	LYZ	6.88E-05
GDP-mannose 4,6-dehydratase	GMDS	9.69E-05
albumin	ALB	1.07E-04
FAST kinase domains 5	FASTKD5	1.47E-04
ferritin light chain	FTL	1.48E-04
TUB bipartite transcription factor	TUB	1.57E-04
suppressor APC domain containing 2	SAPCD2	1.91E-04
collagen type XXVI alpha 1 chain	COL26A1	2.00E-04

TABLE 3. Select genes controlled by Liver GRN

Gene Symbol	*Gene Name*	*Promoter-Liver GRN*	*Promoter-Other GRN*
AKR1C1	aldo-keto reductase family 1 member C1	FoxA1, FoxA2	MRF-2, p53
ALB	albumin	FoxA1, FoxA2, HNF4α, HNF4α	C/EBPalpha, STAT1, STAT1alpha, STAT1beta, STAT2, STAT3, STAT4, STAT5A, STAT5B, TBP
APOA4	apolipoprotein A4	FoxA1, FoxA2, HNF4α	COUP, COUP-TF, COUP-TF1, PPAR-gamma1, PPAR-gamma2
APOC3	apolipoprotein C3	FoxA1, FoxA2, HNF1α, HNF4α	COUP, COUP-TF, COUP-TF1, PPAR-gamma1, PPAR-gamma2
APOM	apolipoprotein M	FoxA1, FoxA2, HNF1α, HNF4α	COUP, COUP-TF, COUP-TF1, PPAR-gamma1, PPAR-gamma2, STAT1
CD24	CD24 molecule		c-Myc, FOXC1, Max1, Nkx3-1, Nkx3-1 v1, Nkx3-1 v2, Nkx3-1 v3, POU3F2, SRY
CYP19A1	cytochrome P450 family 19 subfamily A member 1	FoxA1, FoxA2, HNF4α	PPAR-gamma1, PPAR-gamma2
CYP1A1	cytochrome P450 family 1 subfamily A member 1	FoxA1, FoxA2, HNF4α	Bach2, C/EBPbeta, GR, GR-alpha, MZF-1, p53, SREBP-1a, SREBP-1b, SREBP-1c
CYP24A1	cytochrome P450 family 24 subfamily A member 1	FoxA1, FoxA2, HNF1α, HNF4α	AML1a, AP-1, AREB6, ATF-2, c-Jun, C/EBPalpha, CREB, deltaCREB, FOXD3, YY1
CYP4F2	cytochrome P450 family 4 subfamily F member 2	HNF4α	FOXO4, GATA-2, HOXA5, HSF1 (long), p53, Pax-2, Pax-2a, RORalpha1, STAT3, YY1
DLK1	delta like non-canonical Notch ligand 1	FoxA1, FoxA2, HNF1α, HNF4α	Evi-1, GATA-2, PPAR-gamma1, PPAR-gamma2
FASTKD5	FAST kinase domains 5	FoxA1, FoxA2, HNF4α	c-Myb, CP2, GATA-2, Ik-2, MZF-1, NF-kappaB, NF-kappaB1, Pax-2, Pax-2a, Zic3

FGR1	fibroblast growth factor receptor 1	FoxA1, FoxA2, HNF1α, HNF4α	CREB, deltaCREB
IRS2	insulin receptor substrate 2	FoxA1, FoxA2, HNF4α	AP-1, ATF-2, c-Jun, C/EBPalpha, E2F, NRSF form 1, NRSF form 2
KRT7	keratin 7	FoxA1, FoxA2, HNF4α	AP-1, c-Fos, c-Jun, p53, PPAR-gamma1, PPAR-gamma2
NRARP	NOTCH-regulated ankyrin repeat protein	FoxA1, FoxA2, HNF4α	Oct-B1, oct-B2, oct-B3, POU2F1, POU2F1a, POU2F2, POU2F2 (Oct-2.1), POU2F2B, PPAR-gamma2
PPARGC1A	PPARG coactivator 1 alpha	FoxA1, FoxA2, HNF4α	aMEF-2, CREB, deltaCREB, MEF-2A
TGFBR3	transforming growth factor beta receptor 3	FoxA1, FoxA2, HNF1a, HNF4α	c-Myc, GCNF-2, Max1, MyoD, Pax-2, Pax-2a, Pax-2b, Sox9, Sp1
TWNK	twinkle mtDNA helicase	FoxA1, FoxA2, HNF4α	
UCA1	urothelial cancer associated 1	FoxA1	

FIGURE 1.

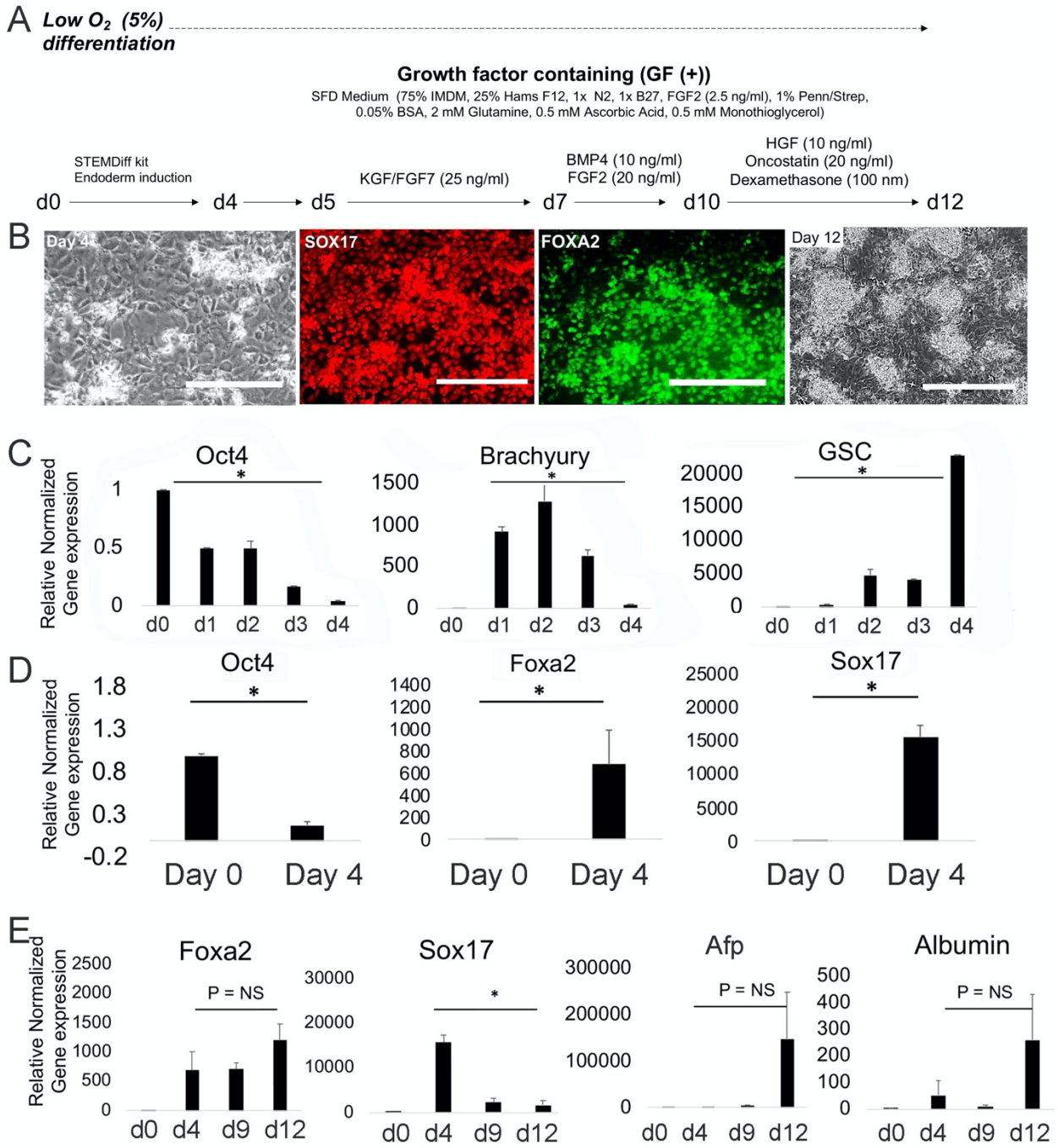
A
Low O2 (5%) differentiation
Growth factor containing (GF (+))
SFD Medium (75% IMDM, 25% Hams F12, 1x N2, 1x B27, FGF2 (2.5 ng/ml), 1% Penn/Strep, 0.05% BSA, 2 mM Glutamine, 0.5 mM Ascorbic Acid, 0.5 mM Monothioglycerol)
STEMDiff kit Endoderm induction
KGF/FGF7 (25 ng/ml)
BMP4 (10 ng/ml) FGF2 (20 ng/ml)
HGF (10 ng/ml) Oncostatin (20 ng/ml) Dexamethasone (100 nm)
d0
d4
d5
d7
d10
d12
B
Day 4
SOX17
FOXA2
Day 12
C
Oct4
Brachyury
GSC
Relative Normalized Gene expression
d0 d1 d2 d3 d4
D
Oct4
Foxa2
Sox17
Day 0
Day 4
E
Foxa2
Sox17
Afp
Albumin
P = NS
d0 d4 d9 d12

Figure 1. Endoderm induction and liver differentiation under low oxygen conditions using a conventional liver differentiation protocol.

A) Schematic of growth factor positive (GF (+)) protocol showing stages of endoderm induction from day 0-4 with a commercially available kit, differentiation towards gut tube with KGF from Days 5-7, and differentiation towards hepatic progenitors involves addition of BMP and FGF2 until Day 10, followed by addition of hepatocyte growth factor (HGF), oncostatin, and dexamethasone from Day 10 to Day 14. Protocol adopted from (Takebe, Sekine et al. 2013).

B) Morphology and immunochemistry of endoderm and hepatic progenitor cells. Left panel- Phase contrast image of Day 4 endoderm morphology. Bar = 200 µm. Middle panels- Sox 17 (red) and Foxa2 (green) staining of endoderm progenitor cells on Day 4. Right panel- phase contrast image of Day 12 morphology with elongated cells and prominent clusters of cells throughout.

C) Bar graph of gene expression kinetics (qRT-PCR) of Oct 4, brachyury, and goosecoid (gsc) during 4 days of endoderm induction from hPSC. N = 3 for each group compared. All comparisons between Day 0 and Day 4. Oct 4, P = 0.00093, brachyury, P = 0.011, goosecoid, P = 0.010. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

D) Same as C, except data demonstrates Day 0 and Day 4 gene expression. N = 3 for both control and experimental groups. All comparisons between Day 0 and Day 4. Oct 4, P = 0.000067, Foxa2, P = 0.037, Sox 17, P = 0.00029. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

E) Same as D, except data demonstrates endoderm TFs (Foxa2 and Sox 17), endoderm and liver differentiation genes (alpha-fetoprotein and albumin) during liver differentiation. N = 3 for both control and experimental groups. All comparisons between Day 4 and Day 12. Foxa2 , P = NS, Sox 17, P =0.0008, alpha-fetoprotein (Afp), P = NS (P = 0.10), albumin (Alb), P = NS (P = 0.18). Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

FIGURE 2.

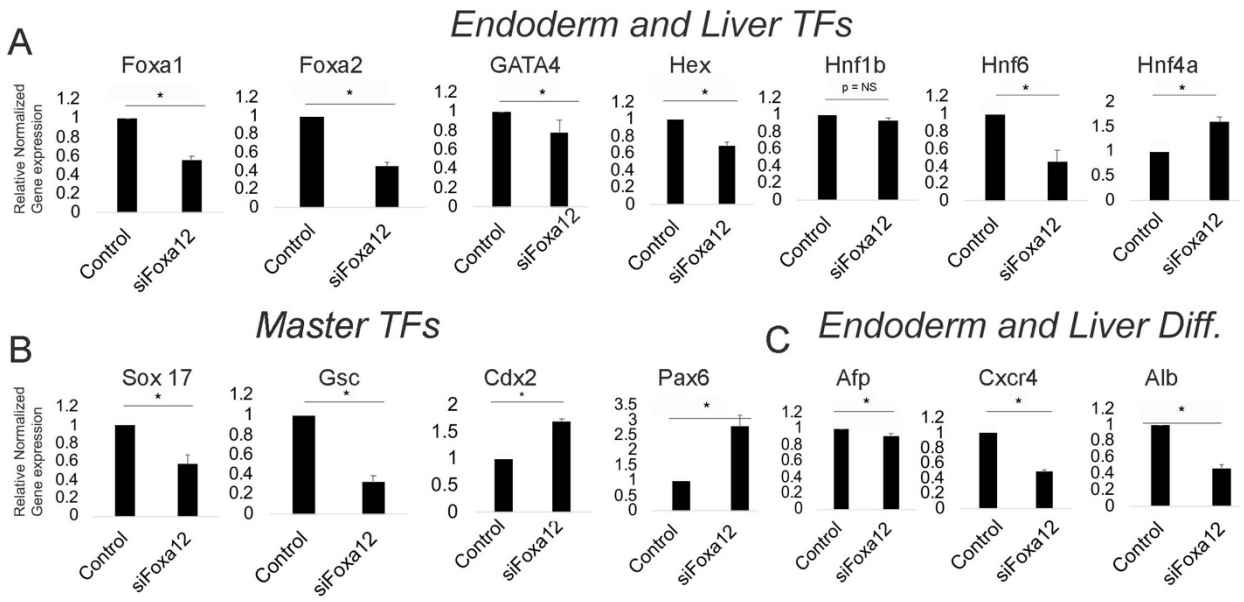

Figure 2. Gene expression of gene regulatory network in siRNA Foxa1/2 phenotype during endoderm induction from human stem cells.

A) Bar graph of gene expression kinetics (qRT-PCR) of endoderm and liver transcription factors (TFs) during endoderm induction from human stem cells, on Day 4 of culture. Cells transfected on Day 1 with siRNA. In control (scramble) and siFoxa1/2 conditions. N = 3 for control and siFoxa1/2 for all conditions. Foxa1 P = 0.000045, Foxa2, P = 0.00003, GATA4, P = 0.04, Hex P = 0.00002, HNF1ß, P = NS, Hnf6, P = 0.009, HNF4α, P = 0.00016. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

B) Same as A except master TFs during germ layer induction. Sox 17, P = 0.0064, Gsc, P = 0.0002, Cdx2, P = 0.0018, Pax6, P = 0.0010. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

C) Same as B except endoderm and liver differentiation markers. Afp, P = 0.004, Cxcr4, P = xxx, Alb, P = 0.000017. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

FIGURE 3.

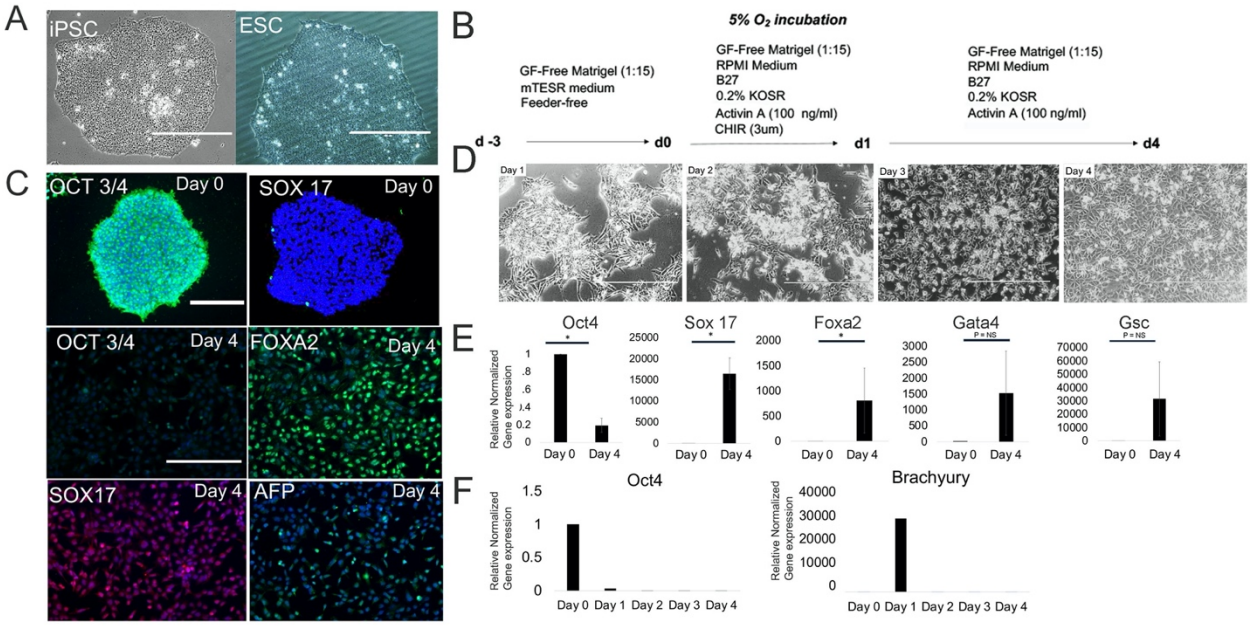
A
iPSC
ESC
B
5% O2 incubation
GF-Free Matrigel (1:15)
mTESR medium
Feeder-free
GF-Free Matrigel (1:15)
RPMI Medium
B27
0.2% KOSR
Activin A (100 ng/ml)
CHIR (3um)
GF-Free Matrigel (1:15)
RPMI Medium
B27
0.2% KOSR
Activin A (100 ng/ml)
d -3
d0
d1
d4
C
OCT 3/4
Day 0
SOX 17
Day 0
OCT 3/4
Day 4
FOXA2
Day 4
SOX17
Day 4
AFP
Day 4
D
Day 1
Day 2
Day 3
Day 4
E
Oct4
Sox 17
Foxa2
Gata4
Gsc
Relative Normalized Gene expression
Day 0
Day 4
F
Oct4
Brachyury
Relative Normalized Gene expression
Day 0
Day 1
Day 2
Day 3
Day 4

Figure 3. Endoderm induction from human stem cells in a hypoxic, growth factor-free, liver differentiation protocol.

A) Phase contrast images of human pluripotent stem cell (hPSC) colonies of iPSC and ESC were used in these studies.

B) Schematic of differentiation system for endoderm induction. iPSC were maintained under hypoxia on GF-free coated matrigel plates, then replated under hypoxic conditions, on freshly coated plates, in the presence of Activin (100 ng/ml)/ CHIR (3 µm) for 1 day followed by Activin (100 ng/ml) for an additional 3 days. Medium contains RPMI, 1x B27, and 0.2% KOSR for improved cell survival.

C) Immunofluorescent staining. Top row- Day 0, pluripotent colonies for Oct4/ DAPI (left) and Sox17/ DAPI (right), Bar = 400 µm. Middle- Day 4 cells stained for Oct4/ DAPI (left)and Foxa2/ DAPI (right), Bar = 200 µm. Bottom -Day 4 cells stained for Sox17/DAPI (left), and AFP/ DAPI (right), Bar = 200 µm.

D) Phase contrast images of cells during endoderm induction from Days 1-4. Day 4 cells (right panel) demonstrate endoderm morphology.

E) Bar graph of gene expression kinetics (qRT-PCR) of endoderm transcription factors (TFs) during endoderm induction from human stem cells, on Day 0 and Day 4 of culture. Cells transfected on Day 1 with siRNA. In control (scramble) and siFoxa1/2 conditions. N = 3 for control and siFoxa1/2 for all conditions. Oct4, P =0.000063, Sox17, P = 0.0015, Foxa2, P = 0.012, GATA4, P = NS (P = 0.011), GSC, P = NS. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

F) Bar graph of gene expression kinetics (qRT-PCR) of transcription factors (TFs) (Oct4, Brachyury) during endoderm induction from human stem cells, on Day 0-4. Plotted is mean for a n =1 experiments.

FIGURE 4.

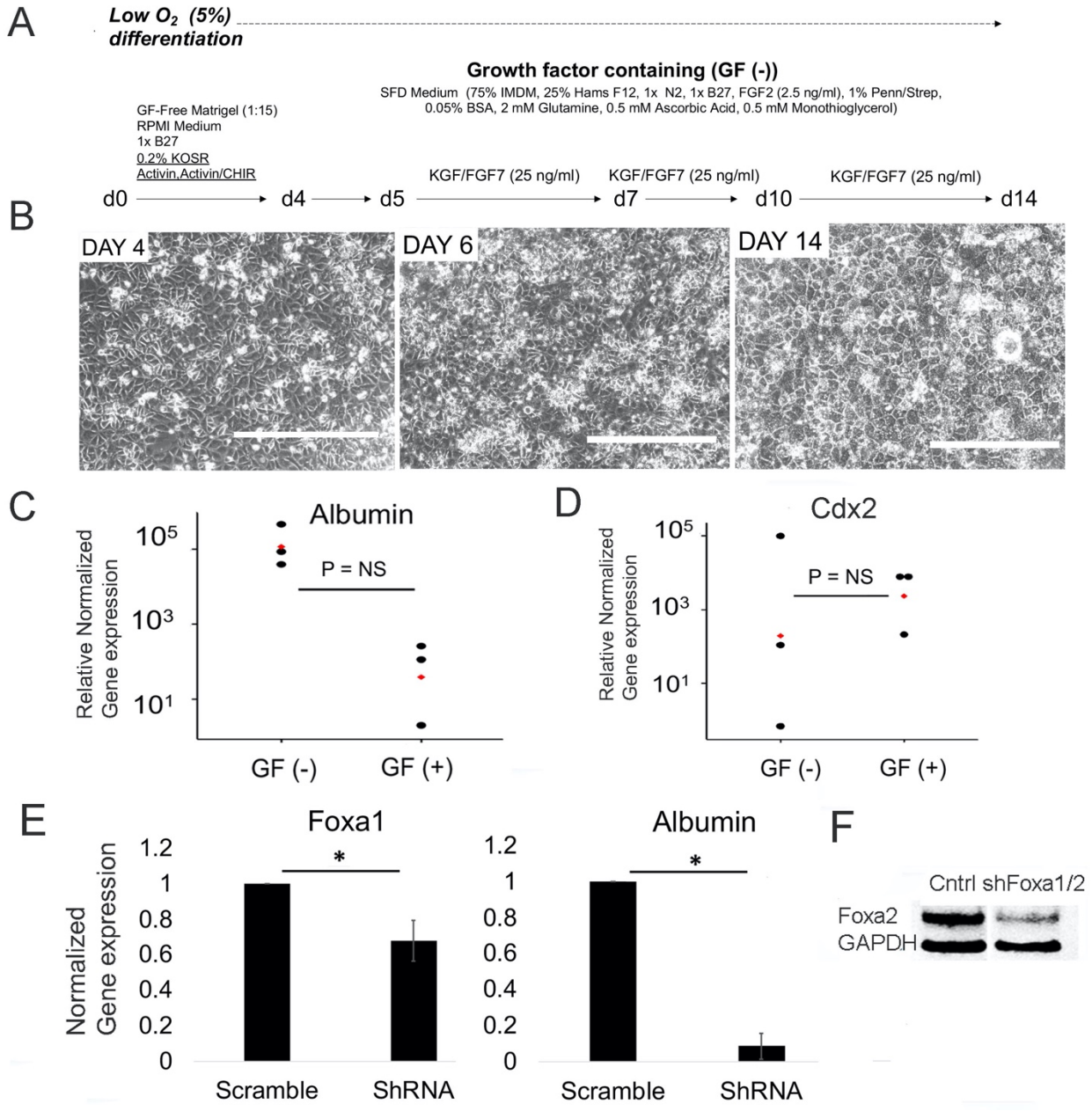
A
Low O2 (5%) differentiation
Growth factor containing (GF (-))
SFD Medium (75% IMDM, 25% Hams F12, 1x N2, 1x B27, FGF2 (2.5 ng/ml), 1% Penn/Strep, 0.05% BSA, 2 mM Glutamine, 0.5 mM Ascorbic Acid, 0.5 mM Monothioglycerol)
GF-Free Matrigel (1:15)
RPMI Medium
1x B27
0.2% KOSR
Activin,Activin/CHIR
KGF/FGF7 (25 ng/ml)
KGF/FGF7 (25 ng/ml)
KGF/FGF7 (25 ng/ml)
d0
d4
d5
d7
d10
d14
B
DAY 4
DAY 6
DAY 14
C
Albumin
P = NS
Relative Normalized Gene expression
GF (-)
GF (+)
D
Cdx2
P = NS
Relative Normalized Gene expression
GF (-)
GF (+)
E
Foxa1
Albumin
Normalized Gene expression
Scramble
ShRNA
F
Cntrl shFoxa1/2
Foxa2
GAPDH

Figure 4. Hepatic differentiation from human stem cells in a novel, simplified, growth factor-free, liver differentiation protocol under hypoxia with gene silencing for Foxa1/2.

A) Schematic of differentiation system. hPSC were differentiated towards endoderm, then gut tube endoderm, and then hepatic progenitor cells under hypoxic conditions (5% O_2). The protocol involves induction of endoderm by Day 5, and then switch of medium to SFD medium, followed by differentiation to gut tube by Day 7, and continues the addition of KGF with no additional growth factors from Day 7- Day 14, under hypoxic conditions.

B) Phase contrast microscopy during differentiation from Day 4 (endoderm), Bar = 400 µm, day 6 (gut tube endoderm), Bar = 200 µm, and Day 14 progenitors, Bar = 200 µm. Images demonstrate transition of endoderm morphology bright cell borders, and cuboidal morphology, consistent with hepatic progenitors.

C) Scatter plot of albumin gene expression (qRT-PCR) on Day 14 of differentiation, in the GF (-) compared to the GF (+) condition, on a log plot. N = 3 for both conditions. Albumin (Alb), P = NS (P = 0.21).

Plotted is individual values and mean. Significance (*) defined as $P \leq 0.05$.

D) Scatter plot of Cdx2 gene expression (qRT-PCR) on Day 14 of differentiation, in the GF (-) compared to the GF (+) condition, on a log plot. N = 3 for both conditions. Albumin (Alb), P = NS (P = 0.69). Plotted is individual values and mean. Significance (*) defined as $P \leq 0.05$.

E) Bar graph of gene expression (qRT-PCR) of Foxa1 and albumin (Alb) on Day 14, after shRNA scramble and shRNA Foxa1/2 knockdown on Day 6 of culture, for GF (-) condition. N = 3 for both conditions. Foxa 1, P = NS (P = 0.09), albumin (Alb), P = 0.000021. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

F) Western blot on Day 8 of GF (-) condition, after shRNA Foxa1/2 transduction on day 6 gut tube progenitor cells in the GF (-) condition, in Control (scrambled) and shFoxa1/2 condition. GAPDH is control.

FIGURE 5.

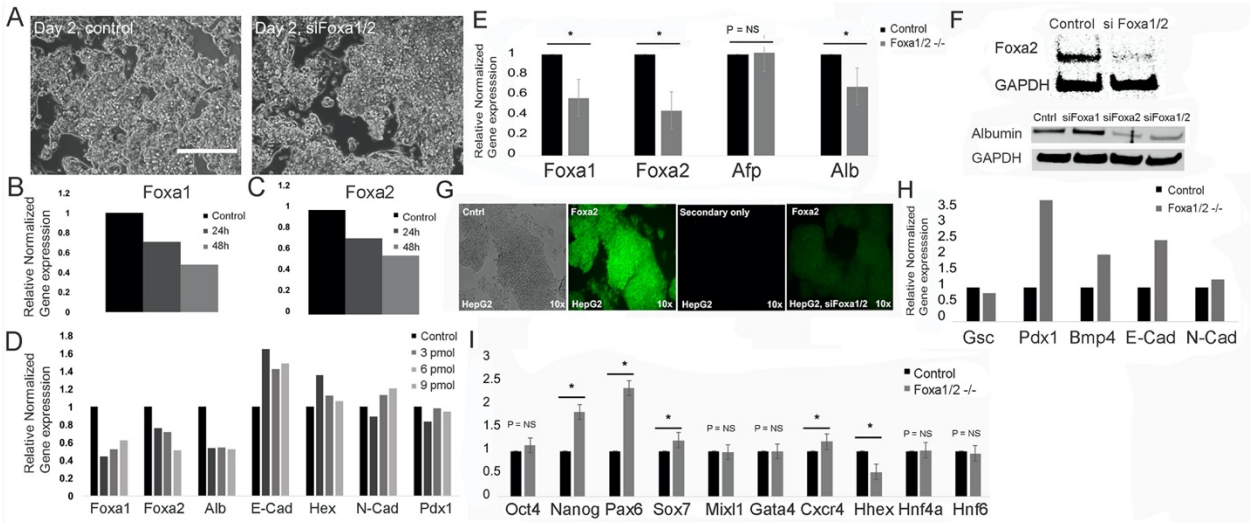
A
Day 2, control
Day 2, siFoxa1/2
B
Foxa1
Control
24h
48h
Relative Normalized Gene expresssion
C
Foxa2
Control
24h
48h
D
Control
3 pmol
6 pmol
9 pmol
Relative Normalized Gene expresssion
Foxa1
Foxa2
Alb
E-Cad
Hex
N-Cad
Pdx1
E
Control
Foxa1/2 -/-
Relative Normalized Gene expresssion
P = NS
Foxa1
Foxa2
Afp
Alb
F
Control
si Foxa1/2
Foxa2
GAPDH
Cntrl
siFoxa1
siFoxa2
siFoxa1/2
Albumin
GAPDH
G
Cntrl
Foxa2
Secondary only
Foxa2
HepG2
10x
HepG2, siFoxa1/2
H
Control
Foxa1/2 -/-
Relative Normalized Gene expresssion
Gsc
Pdx1
Bmp4
E-Cad
N-Cad
I
Control
Foxa1/2 -/-
P = NS
Oct4
Nanog
Pax6
Sox7
Mixl1
Gata4
Cxcr4
Hhex
Hnf4a
Hnf6

Figure 5. siRNA of Foxa1/2 regulates differentiation genes and GRN.

Human HepG2 cells were transfected for siRNA via reverse transfection.

A) Phase contrast images for morphological comparison of scrambled (left) and siRNA Foxa1/2 (right) 48 hours after transfection in HepG2 cells.

B) Bar graph of gene expression kinetics (qRT-PCR) for siRNA of HepG2 cells for Foxa1 (n = 1) after 0 (control), 24 and 48 hour time points.

C) Same as B except siRNA for Foxa2 (n = 1)

D) Bar graph of gene expression (qRT-PCR) data (n=1) of HepG2 cell for optimization of pmole siRNA added for pooled siFoxa1, and pooled siFoxa2, ranging from 3-9 pmoles. Control is scrambled siRNA. Markers tested are Foxa1, Foxa2, Alb, E-Cad, Hex, N-Cad, and Pdx1.

E) Bar graph of gene expression (qRT-PCR) data for siFoxa1/2 knockdown in HepG2 assayed after 48 hours, compared to scrambled controls. N = 5 for control (scrambled) and siRNA Foxa1/2 conditions. Foxa1, P = 0.000021, Foxa2 , P = 0.000014, alpha-fetoprotein (Afp), P = NS, albumin (Alb), P = 0.0003. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

F) Western blot analysis of control (scrambled) knockdown and siFoxa1/2 (-/-) in HepG2 48 hours after transfection. Above- Foxa2 and GAPDH (control) shown. Below-Albumin and GAPDH (control) shown.

G) Brightfield and immunohistochemistry of Foxa2 48 hours after siFoxa1/2 transfection. Left- Brightfield image of fixed HepG2 cells 48 h after transfection of scrambled control. Next is the same image with FITC (green) labeling of Foxa2 cells after scramble transfection. Next is secondary antibody staining only in the control condition. Right- FITC (green) labeling of Foxa2 cells in siFoxa1/2 (-/-) condition after 48 hours.

H) Bar graph of gene expression (qRT-PCR) for scramble and siFoxa1/2 (-/-) after 48 hours in HepG2 cells. All conditions are N = 1. Shown are Goosecoid (Gsc), Pdx1, Bmp4, E-Cadherin (E-Cad), and N-Cadherin (N-Cad) expression.

I) Same as H) except for analysis of several transcription factors. All conditions are N = 5. Oct 4, P = NS, Nanog, P = 0.00078, Pax 6, P = 0.00625, Sox7, P = NS (P = 0.102), Mixl1, P = NS, Gata4, P = NS, Cxcr4, P =NS, Hex, P = 0.0036, HNF4α, P = NS, Hnf6, P = NS. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

FIGURE 6.

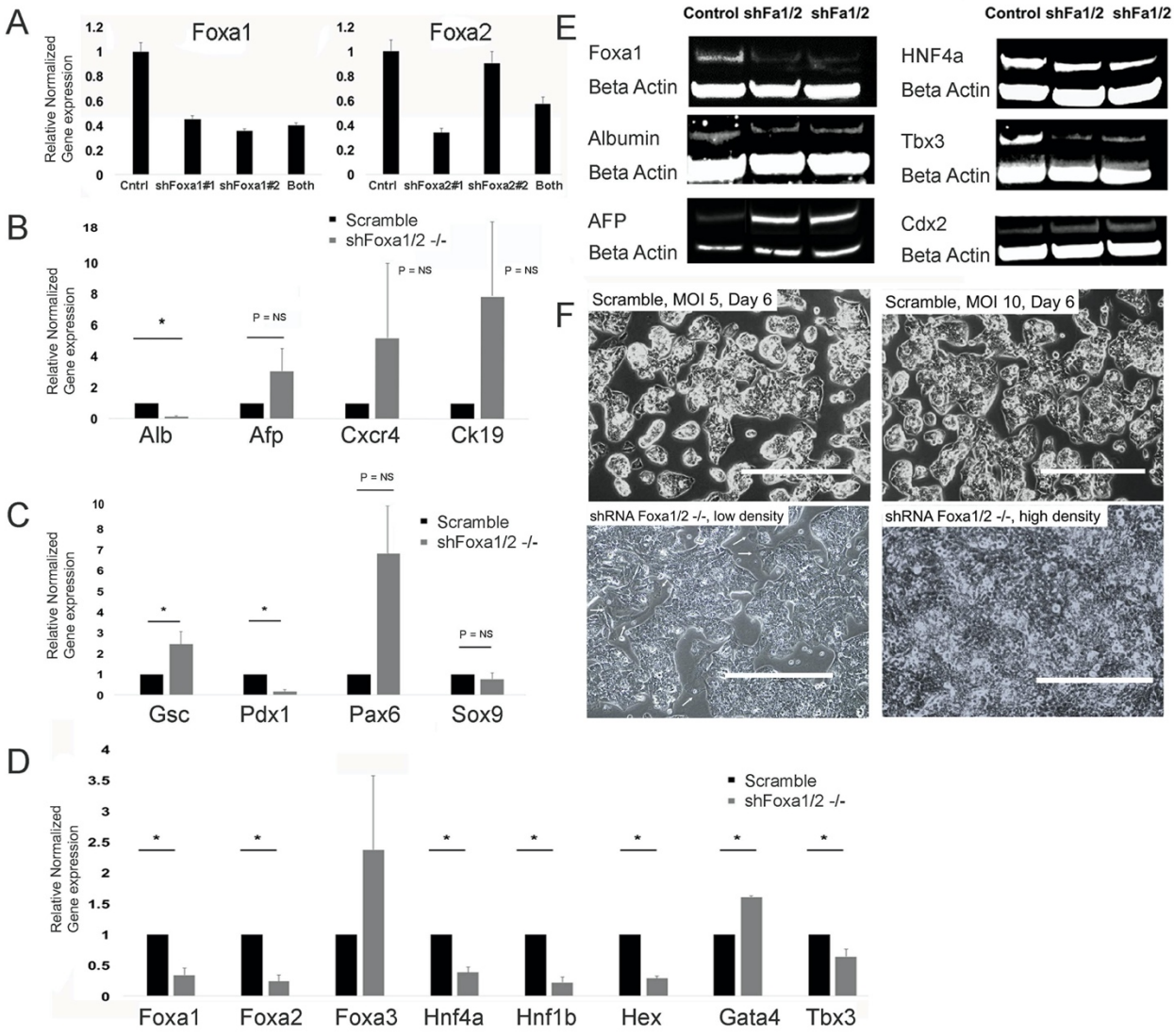
A
Foxa1
Foxa2
Relative Normalized Gene expression
Cntrl
shFoxa1#1
shFoxa1#2
Both
shFoxa2#1
shFoxa2#2
B
Scramble
shFoxa1/2 -/-
P = NS
Alb
Afp
Cxcr4
Ck19
C
Gsc
Pdx1
Pax6
Sox9
D
Foxa1
Foxa2
Foxa3
Hnf4a
Hnf1b
Hex
Gata4
Tbx3
E
Control shFa1/2 shFa1/2
Foxa1
Beta Actin
Albumin
AFP
HNF4a
Tbx3
Cdx2
F
Scramble, MOI 5, Day 6
Scramble, MOI 10, Day 6
shRNA Foxa1/2 -/-, low density
shRNA Foxa1/2 -/-, high density

Figure 6. shRNA of Foxa1/2 regulates differentiation gene and GRN.

Cells were transduced with 5-10 MOI of shFoxa1 and shFoxa2, or shRNA scrambled controls, continuously selected after several passages with puramycin selection, and analyzed.

A) Bar graph of gene expression (qRT-PCR) after shRNA transduction with shFoxa1, shFoxa2, shFoxa1/2, and scrambled shRNA control, followed by antibiotic selection. N = 1 shown. Left - Foxa1 expression in control, shFoxa1, shFoxa2, and both shFoxa1/2. Right- Foxa2 expression in control, shFoxa1, shFoxa2, and both shFoxa1/2. Plotted is mean for 3 replicates.

B) Bar graph of gene expression (qRT-PCR) after shRNA transduction with shFoxa1/2, and scrambled shRNA control, followed by antibiotic selection for two weeks. N = 3 for all conditions. Alb, P = 0.000019, Afp, P = NS, Cxcr4, P = NS, Ck19, P = NS. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

C) Bar graph of gene expression (qRT-PCR) after shRNA transduction of HepG2 cells with shFoxa1/2, and scrambled shRNA control, followed by antibiotic selection for two weeks. N = 3 for all conditions. Gsc, P = 0.027, Pdx1, P = 0.0002, Pax 6, P = NS, Sox 9, P = NS. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

D) Bar graph of gene expression (qRT-PCR) after shRNA transduction of HepG2 cells with shFoxa1/2, and scrambled shRNA control, followed by antibiotic selection for two weeks. N = 3 for all conditions. Foxa1, P = 0.001, Foxa2, P = 0.004, Foxa3, P = NS, HNF4α, P = 0.0006, HNF1ß, P = 0.0002, Hex, P = 0.000008, Gata4, P = 0.0000009, Tbx3, P = 0.013. Plotted is mean ± SD. Significance (*) defined as $P \leq 0.05$.

E) Western blotting for scrambled shRNA (control) and n = 2 cell lines analyzed. Beta Actin was control. Foxa1, albumin, Afp, HNF4α, Tbx3, and Cdx2 shown.

F) Morphology comparison between transfected and non-transfected HepG2 cells- Top row- shRNA scramble at MOI of 5 and MOI of 10, Bottom row- shRNA Foxa1/2 (-/-) condition. Bottom left panel-Small white arrows demonstrate elongated and flattened cells at the edges of colonies, that were seen routinely at low density. Bottom right panel- shFox2 -/- cell lines at high density.

FIGURE 7.

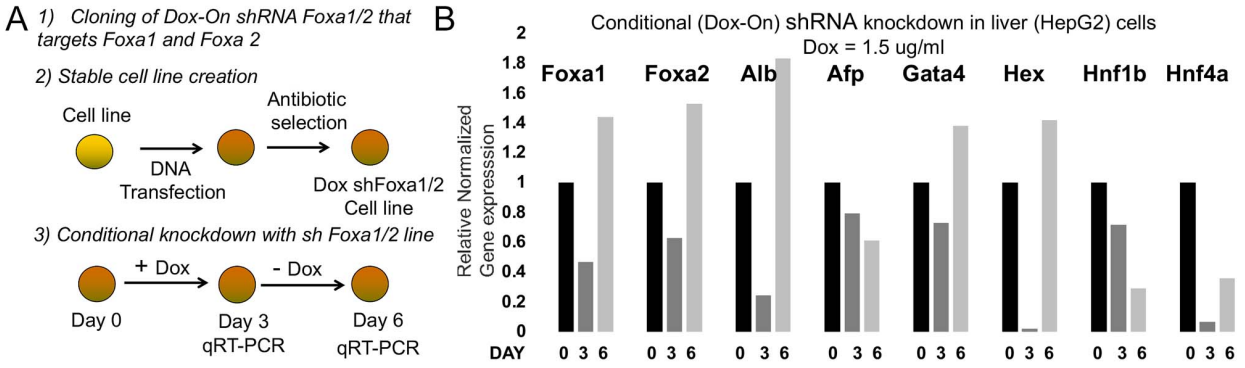

Figure 7. Conditional shRNA expression of Foxa1/2 -/- phenotype demonstrates gene reversibility.

A) Schematic of stable cell line creation for conditional Foxa1/2 (-/-) phenotype. Cells were transfected with Doxycycline (Dox)-On shRNA vector (shRNA Foxa2#1) targeting both shRNA for Foxa1 and Foxa2.

(1) shRNA was cloned into the Dox-on vector. (2) Next, DNA was transfected and selected for by antibiotic or shRNA scrambled controls, continuously selected after several passages with puramycin selection, and analyzed. (3) Next, cells are exposed to Dox (1.5 ug/ml) for 3 days and then Dox is removed for 3 days, with cell collection at Day 3 and Day 6 for qRT-PCR. B) Bar graph of gene expression (qRT-PCR) on Day 0, Day 3 (Dox addition, from Day 0- Day 3), Day 6 (Dox removed from Day 3-Day 6). Dox concentration was 1.5 µm.

FIGURE 8.

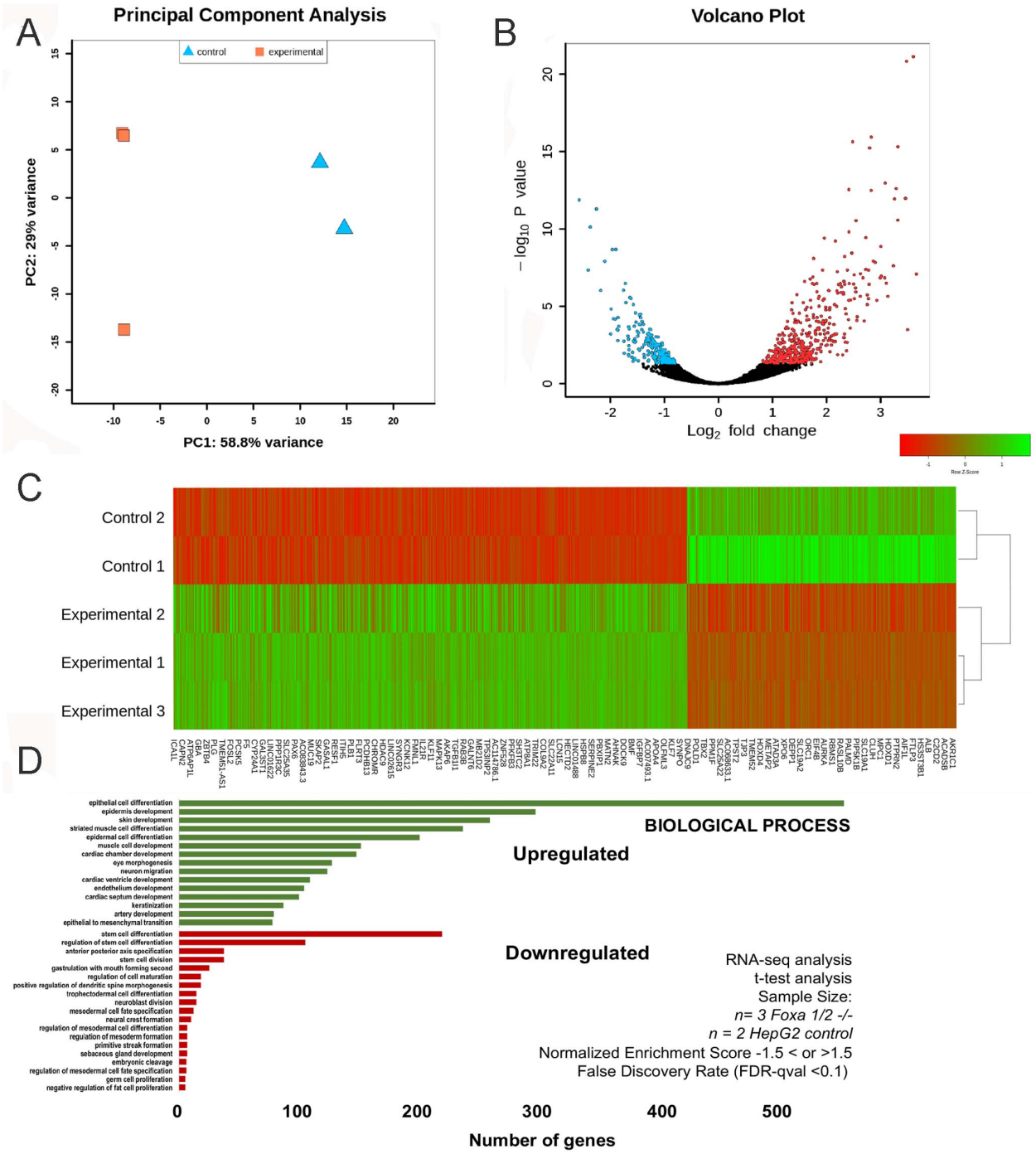
A
Principal Component Analysis
control
experimental
PC2: 29% variance
PC1: 58.8% variance
B
Volcano Plot
$-\log_{10}$ P value
Log_2 fold change
C
Control 2
Control 1
Experimental 2
Experimental 1
Experimental 3
D
BIOLOGICAL PROCESS
Upregulated
Downregulated
epithelial cell differentiation
epidermis development
skin development
striated muscle cell differentiation
epidermal cell differentiation
muscle cell development
cardiac chamber development
eye morphogenesis
neuron migration
cardiac ventricle development
endothelium development
cardiac septum development
keratinization
artery development
epithelial to mesenchymal transition
stem cell differentiation
regulation of stem cell differentiation
anterior posterior axis specification
stem cell division
gastrulation with mouth forming second
regulation of cell maturation
positive regulation of dendritic spine morphogenesis
trophectodermal cell differentiation
neuroblast division
mesodermal cell fate specification
neural crest formation
regulation of mesodermal cell differentiation
regulation of mesoderm formation
primitive streak formation
sebaceous gland development
embryonic cleavage
regulation of mesodermal cell fate specification
germ cell proliferation
negative regulation of fat cell proliferation
RNA-seq analysis
t-test analysis
Sample Size:
n= 3 Foxa 1/2 -/-
n = 2 HepG2 control
Normalized Enrichment Score -1.5 < or >1.5
False Discovery Rate (FDR-qval <0.1)
Number of genes

Figure 8. **RNA-seq analysis of control (HepG2) and experimental (shFoxa1/2 -/-) genes.**

A)Principal component analysis with experimental samples (n = 3), orange, and two control samples (n = 2), blue. Experimental group (shFoxa1/2 (-/-)) and control group are shown to cluster apart on the PC1 axis.

B) Volcano plot displaying differential expression analysis between the control and experimental group. The adjusted p-value from the "res" function in DESeq2 was used as the p-value for this plot. Each circle in the plot represents a gene. The genes are colored if their adjusted p-value is below 0.05, which was used to indicate significant differential expression. The genes in green represent upregulation and the genes in red represent down regulation based on the log2fold change.

C) Heat map of differentially expressed genes for the five samples. Green represents relatively higher gene expression, red represents relatively lower gene expression, compared to other samples. Genes with less significant expression level differences (adjusted p-value > 0.05) were categorized as not differentially expressed and therefore not included. 448 upregulating and 235 downregulating genes are grouped together and listed by increasing adjusted P value for each group. Every seventh gene in the list is labeled. Hierarchical clustering was performed using the "hclust" function in R.

D) Plot of biological process gene sets selected for relationship to stem cells and differentiation. Bioinformatics analysis of RNA seq data in which number of genes within each significant gene set is shown after removing genes not in the expression dataset. Red indicates a negative normalized enrichment score value or downregulation, green indicates a positive NES value or upregulation. Gene sets are from version 7.0 of the Gene Ontology (GO) gene sets from the Broad Institute. Number of genes within group shown after calculation, t-test comparison between groups ($P < 0.05$), for n = 2 control, n = 3 shFoxa1/2 (-/-) cells. Criteria was NES -1.5 < or > 1.5, and false discovery rate (FDR)-qval < 0.1).

Competing Interests: The authors hereby state no competing interest involved with the ideation, writing, or revision of this manuscript

ACKNOWLEDGEMENTS:

The authors acknowledge the UB Center for Excellence in Bioinformatics, and Roswell Park Lentiviral core facility. NP was supported by the UB CBE startup funds, NP and MM were supported by the New York State Stem Cell Science C024316 and the Stem cells in regenerative medicine (ScIRM) center. OO was supported by the Western NY Prosperity Fellowship.

AUTHOR CONTRIBUTIONS:

IY: Obtained data, analyzed data, built figures, wrote and approved manuscript

MM: Obtained data, analyzed data, built figures, edited and approved manuscript

DG: Obtained data, analyzed data, built figures, approved manuscript

OO: Obtained data, analyzed data, built figures, approved manuscript

TG: Obtained data, analyzed data, built figures, approved manuscript

TM: Obtained data, analyzed data, built figures, approved manuscript

SM: Obtained data, analyzed data, built figures, approved manuscript

XL: Obtained data, analyzed data, built figures, approved manuscript

AS: Obtained data, analyzed data, built figures, approved manuscript

RT: Obtained data, analyzed data, built figures, approved manuscript

PC: Obtained data, analyzed data, built figures, approved manuscript

RP: Obtained data, analyzed data, built figures, approved manuscript

NP: Conceptualized, acquired funding, investigated, supervised, wrote, edited manuscript, and approved manuscript.

REFERENCES

Ancey, P. B., S. Ecsedi, M. P. Lambert, F. R. Talukdar, M. P. Cros, D. Glaise, D. M. Narvaez, V. Chauvet, Z. Herceg, A. Corlu and H. Hernandez-Vargas (2017). "TET-Catalyzed 5-Hydroxymethylation Precedes HNF4A Promoter Choice during Differentiation of Bipotent Liver Progenitors." Stem Cell Reports **9**(1): 264-278.

Argemi, J., M. U. Latasa, S. R. Atkinson, I. O. Blokhin, V. Massey, J. P. Gue, J. Cabezas, J. J. Lozano, D. Van Booven, A. Bell, S. Cao, L. A. Vernetti, J. P. Arab, M. Ventura-Cots, L. R. Edmunds, C. Fondevilla, P. Starkel, L. Dubuquoy, A. Louvet, G. Odena, J. L. Gomez, T. Aragon, J. Altamirano, J. Caballeria, M. J. Jurczak, D. L. Taylor, C. Berasain, C. Wahlestedt, S. P. Monga, M. Y. Morgan, P. Sancho-Bru, P. Mathurin, S. Furuya, C. Lackner, I. Rusyn, V. H. Shah, M. R. Thursz, J. Mann, M. A. Avila and R. Bataller (2019). "Defective HNF4alpha-dependent gene expression as a driver of hepatocellular failure in alcoholic hepatitis." Nat Commun **10**(1): 3126.

Bochkis, I. M., J. Schug, D. Z. Ye, S. Kurinna, S. A. Stratton, M. C. Barton and K. H. Kaestner (2012). "Genome-wide location analysis reveals distinct transcriptional circuitry by paralogous regulators Foxa1 and Foxa2." PLoS Genet **8**(6): e1002770.

Bossard, P. and K. S. Zaret (1998). "GATA transcription factors as potentiators of gut endoderm differentiation." Development **125**(24): 4909-4917.

Calmont, A., E. Wandzioch, K. D. Tremblay, G. Minowada, K. H. Kaestner, G. R. Martin and K. S. Zaret (2006). "An FGF response pathway that mediates hepatic gene induction in embryonic endoderm cells." Dev Cell **11**(3): 339-348.

Cascio, S. and K. S. Zaret (1991). "Hepatocyte differentiation initiates during endodermal-mesenchymal interactions prior to liver formation." Development **113**(1): 217-225.

Cho, C. H., N. Parashurama, E. Y. Park, K. Suganuma, Y. Nahmias, J. Park, A. W. Tilles, F. Berthiaume and M. L. Yarmush (2008). "Homogeneous differentiation of hepatocyte-like cells from embryonic stem cells: applications for the treatment of liver failure." FASEB J **22**(3): 898-909.

Cirillo, L. A., F. R. Lin, I. Cuesta, D. Friedman, M. Jarnik and K. S. Zaret (2002). "Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4." Mol Cell **9**(2): 279-289.

DeLaForest, A., F. Di Furio, R. Jing, A. Ludwig-Kubinski, K. Twaroski, A. Urick, K. Pulakanti, S. Rao and S. A. Duncan (2018). "HNF4A Regulates the Formation of Hepatic Progenitor Cells from Human iPSC-Derived Endoderm by Facilitating Efficient Recruitment of RNA Pol II." Genes (Basel) **10**(1).

DeLaForest, A., M. Nagaoka, K. Si-Tayeb, F. K. Noto, G. Konopka, M. A. Battle and S. A. Duncan (2011). "HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells." Development **138**(19): 4143-4153.

Desai, S. S., J. C. Tung, V. X. Zhou, J. P. Grenert, Y. Malato, M. Rezvani, R. Espanol-Suner, H. Willenbring, V. M. Weaver and T. T. Chang (2016). "Physiological ranges of matrix rigidity modulate primary mouse hepatocyte function in part through hepatocyte nuclear factor 4 alpha." Hepatology **64**(1): 261-275.

Duncan, S. A., M. A. Navas, D. Dufort, J. Rossant and M. Stoffel (1998). "Regulation of a transcription factor network required for differentiation and metabolism." Science **281**(5377): 692-695.

Fisher, J. B., K. Pulakanti, S. Rao and S. A. Duncan (2017). "GATA6 is essential for endoderm formation from human pluripotent stem cells." Biol Open **6**(7): 1084-1095.
Gadue, P., V. Gouon-Evans, X. Cheng, E. Wandzioch, K. S. Zaret, M. Grompe, P. R. Streeter and G. M. Keller (2009). "Generation of monoclonal antibodies specific for cell surface molecules expressed on early mouse endoderm." Stem Cells **27**(9): 2103-2113.
Gadue, P., T. L. Huber, P. J. Paddison and G. M. Keller (2006). "Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells." Proc Natl Acad Sci U S A **103**(45): 16806-16811.
Gao, N., J. LeLay, M. Z. Vatamaniuk, S. Rieck, J. R. Friedman and K. H. Kaestner (2008). "Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development." Genes Dev **22**(24): 3435-3448.
Genga, R. M. J., E. M. Kernfeld, K. M. Parsi, T. J. Parsons, M. J. Ziller and R. Maehr (2019). "Single-Cell RNA-Sequencing-Based CRISPRi Screening Resolves Molecular Drivers of Early Human Endoderm Development." Cell Rep **27**(3): 708-718 e710.
Guo, R., W. Tang, Q. Yuan, L. Hui, X. Wang and X. Xie (2017). "Chemical Cocktails Enable Hepatic Reprogramming of Mouse Fibroblasts with a Single Transcription Factor." Stem Cell Reports **9**(2): 499-512.
Hallonet, M., K. H. Kaestner, L. Martin-Parras, H. Sasaki, U. A. Betz and S. L. Ang (2002). "Maintenance of the specification of the anterior definitive endoderm and forebrain depends on the axial mesendoderm: a study using HNF3beta/Foxa2 conditional mutants." Dev Biol **243**(1): 20-33.
Harries, L. W., J. E. Brown and A. L. Gloyn (2009). "Species-specific differences in the expression of the HNF1A, HNF1B and HNF4A genes." PLoS One **4**(11): e7855.
Huck, I., S. Gunewardena, R. Espanol-Suner, H. Willenbring and U. Apte (2019). "Hepatocyte Nuclear Factor 4 Alpha Activation Is Essential for Termination of Liver Regeneration in Mice." Hepatology **70**(2): 666-681.
Jung, J., M. Zheng, M. Goldfarb and K. S. Zaret (1999). "Initiation of mammalian liver development from endoderm by fibroblast growth factors." Science **284**(5422): 1998-2003.
Kaestner, K. H. (2015). "An epigenomic road map for endoderm development." Cell Stem Cell **16**(4): 343-344.
Kanaki, M. and D. Kardassis (2017). "Regulation of the human lipoprotein lipase gene by the forkhead box transcription factor FOXA2/HNF-3beta in hepatic cells." Biochim Biophys Acta Gene Regul Mech **1860**(3): 327-336.
Lau, H. H., N. H. J. Ng, L. S. W. Loo, J. B. Jasmen and A. K. K. Teo (2018). "The molecular functions of hepatocyte nuclear factors - In and beyond the liver." J Hepatol **68**(5): 1033-1048.
Lee, C. S., J. R. Friedman, J. T. Fulmer and K. H. Kaestner (2005). "The initiation of liver development is dependent on Foxa transcription factors." Nature **435**(7044): 944-947.
Lehner, F., U. Kulik, J. Klempnauer and J. Borlak (2007). "The hepatocyte nuclear factor 6 (HNF6) and FOXA2 are key regulators in colorectal liver metastases." FASEB J **21**(7): 1445-1462.
Levinson-Dushnik, M. and N. Benvenisty (1997). "Involvement of hepatocyte nuclear factor 3 in endoderm differentiation of embryonic stem cells." Mol Cell Biol **17**(7): 3817-3822.
Li, Z., P. White, G. Tuteja, N. Rubins, S. Sackett and K. H. Kaestner (2009). "Foxa1 and Foxa2 regulate bile duct development in mice." J Clin Invest **119**(6): 1537-1545.

Moya, M., M. Benet, C. Guzman, L. Tolosa, C. Garcia-Monzon, E. Pareja, J. V. Castell and R. Jover (2012). "Foxa1 reduces lipid accumulation in human hepatocytes and is down-regulated in nonalcoholic fatty liver." PLoS One **7**(1): e30014.
Nakamori, D., H. Akamine, K. Takayama, F. Sakurai and H. Mizuguchi (2017). "Direct conversion of human fibroblasts into hepatocyte-like cells by ATF5, PROX1, FOXA2, FOXA3, and HNF4A transduction." Sci Rep **7**(1): 16675.
Novik, E. I., T. J. Maguire, K. Orlova, R. S. Schloss and M. L. Yarmush (2006). "Embryoid body-mediated differentiation of mouse embryonic stem cells along a hepatocyte lineage: insights from gene expression profiles." Tissue Eng **12**(6): 1515-1525.
Odom, D. T., R. D. Dowell, E. S. Jacobsen, L. Nekludova, P. A. Rolfe, T. W. Danford, D. K. Gifford, E. Fraenkel, G. I. Bell and R. A. Young (2006). "Core transcriptional regulatory circuitry in human hepatocytes." Mol Syst Biol **2**: 2006 0017.
Okazaki, K. and E. Maltepe (2006). "Oxygen, epigenetics and stem cell fate." Regen Med **1**(1): 71-83.
Parashurama, N., Y. Nahmias, C. H. Cho, D. van Poll, A. W. Tilles, F. Berthiaume and M. L. Yarmush (2008). "Activin alters the kinetics of endoderm induction in embryonic stem cells cultured on collagen gels." Stem Cells **26**(2): 474-484.
Park, J., C. H. Cho, N. Parashurama, Y. Li, F. Berthiaume, M. Toner, A. W. Tilles and M. L. Yarmush (2007). "Microfabrication-based modulation of embryonic stem cell differentiation." Lab Chip **7**(8): 1018-1028.
Raju, R., D. Chau, T. Notelaers, C. L. Myers, C. M. Verfaillie and W. S. Hu (2018). "In Vitro Pluripotent Stem Cell Differentiation to Hepatocyte Ceases Further Maturation at an Equivalent Stage of E15 in Mouse Embryonic Liver Development." Stem Cells Dev **27**(13): 910-921.
Reizel, Y., A. Morgan, L. Gao, Y. Lan, E. Manduchi, E. L. Waite, A. W. Wang, A. Wells and K. H. Kaestner (2020). "Collapse of the hepatic gene regulatory network in the absence of FoxA factors." Genes Dev **34**(15-16): 1039-1050.
Rezania, A., J. E. Bruin, J. Xu, K. Narayan, J. K. Fox, J. J. O'Neil and T. J. Kieffer (2013). "Enrichment of human embryonic stem cell-derived NKX6.1-expressing pancreatic progenitor cells accelerates the maturation of insulin-secreting cells in vivo." Stem Cells **31**(11): 2432-2442.
Rezvani, M., R. Espanol-Suner, Y. Malato, L. Dumont, A. A. Grimm, E. Kienle, J. G. Bindman, E. Wiedtke, B. Y. Hsu, S. J. Naqvi, R. F. Schwabe, C. U. Corvera, D. Grimm and H. Willenbring (2016). "In Vivo Hepatic Reprogramming of Myofibroblasts with AAV Vectors as a Therapeutic Strategy for Liver Fibrosis." Cell Stem Cell **18**(6): 809-816.
Sekiya, S. and A. Suzuki (2011). "Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors." Nature **475**(7356): 390-393.
Shim, E. Y., C. Woodcock and K. S. Zaret (1998). "Nucleosome positioning by the winged helix transcription factor HNF3." Genes Dev **12**(1): 5-10.
Song, G., M. Pacher, A. Balakrishnan, Q. Yuan, H. C. Tsay, D. Yang, J. Reetz, S. Brandes, Z. Dai, B. M. Putzer, M. J. Arauzo-Bravo, D. Steinemann, T. Luedde, R. F. Schwabe, M. P. Manns, H. R. Scholer, A. Schambach, T. Cantz, M. Ott and A. D. Sharma (2016). "Direct Reprogramming of Hepatic Myofibroblasts into Hepatocytes In Vivo Attenuates Liver Fibrosis." Cell Stem Cell **18**(6): 797-808.
Sund, N. J., S. L. Ang, S. D. Sackett, W. Shen, N. Daigle, M. A. Magnuson and K. H. Kaestner (2000). "Hepatocyte nuclear factor 3beta (Foxa2) is dispensable for maintaining the differentiated state of the adult hepatocyte." Mol Cell Biol **20**(14): 5175-5183.

Takebe, T., K. Sekine, M. Enomura, H. Koike, M. Kimura, T. Ogaeri, R. R. Zhang, Y. Ueno, Y. W. Zheng, N. Koike, S. Aoyama, Y. Adachi and H. Taniguchi (2013). "Vascularized and functional human liver from an iPSC-derived organ bud transplant." Nature **499**(7459): 481-484.
Tauran, Y., S. Poulain, M. Lereau-Bernier, M. Danoy, M. Shinohara, B. D. Segard, S. Kato, T. Kido, A. Miyajima, Y. Sakai, C. Plessy and E. Leclerc (2019). "Analysis of the transcription factors and their regulatory roles during a step-by-step differentiation of induced pluripotent stem cells into hepatocyte-like cells." Mol Omics.
Van Blerkom, J., M. Antczak and R. Schrader (1997). "The developmental potential of the human oocyte is related to the dissolved oxygen content of follicular fluid: association with vascular endothelial growth factor levels and perifollicular blood flow characteristics." Hum Reprod **12**(5): 1047-1055.
Wan, H., S. Dingle, Y. Xu, V. Besnard, K. H. Kaestner, S. L. Ang, S. Wert, M. T. Stahlman and J. A. Whitsett (2005). "Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis." J Biol Chem **280**(14): 13809-13816.
Wang, A., F. Yue, Y. Li, R. Xie, T. Harper, N. A. Patel, K. Muth, J. Palmer, Y. Qiu, J. Wang, D. K. Lam, J. C. Raum, D. A. Stoffers, B. Ren and M. Sander (2015). "Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates." Cell Stem Cell **16**(4): 386-399.
Wang, A. W., Y. J. Wang, A. M. Zahm, A. R. Morgan, K. J. Wangensteen and K. H. Kaestner (2019). "The dynamic chromatin architecture of the regenerating liver." Cell Mol Gastroenterol Hepatol.
Wang, H. and C. B. Wollheim (2005). "Does chasing selected 'Fox' to the nucleus prevent diabetes?" Trends Mol Med **11**(6): 262-265.
Wang, J., C. P. Zhu, P. F. Hu, H. Qian, B. F. Ning, Q. Zhang, F. Chen, J. Liu, B. Shi, X. Zhang and W. F. Xie (2014). "FOXA2 suppresses the metastasis of hepatocellular carcinoma partially through matrix metalloproteinase-9 inhibition." Carcinogenesis **35**(11): 2576-2583.
Wang, W., L. J. Yao, W. Shen, K. Ding, P. M. Shi, F. Chen, J. He, J. Ding, X. Zhang and W. F. Xie (2017). "FOXA2 alleviates CCl4-induced liver fibrosis by protecting hepatocytes in mice." Sci Rep **7**(1): 15532.
Watts, J. A., C. Zhang, A. J. Klein-Szanto, J. D. Kormish, J. Fu, M. Q. Zhang and K. S. Zaret (2011). "Study of FoxA pioneer factor at silent genes reveals Rfx-repressed enhancer at Cdx2 and a potential indicator of esophageal adenocarcinoma development." PLoS Genet **7**(9): e1002277.
Ye, D. Z. and K. H. Kaestner (2009). "Foxa1 and Foxa2 control the differentiation of goblet and enteroendocrine L- and D-cells in mice." Gastroenterology **137**(6): 2052-2062.
Zaret, K. S. (2002). "Regulatory phases of early liver development: paradigms of organogenesis." Nat Rev Genet **3**(7): 499-512.