Foxa1/2 downregulates liver differentiation markers and the endoderm and liver gene regulatory network in human stem cells and in a human stable liver cell line

Iyan Warren[1], Mitchell Maloy[1], Daniel Guiggey[1], Ogechi Ogoke[1], Theodore Groth[1], Tala Mon[1], Saber Mearmadoost[1], Xiaojun Liu[1], Antoni Szeglowski[4], Ryan Thompson[1], Peter Chen[3], Ramasamy Paulmurugan [5], 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 Biological Sciences, University at Buffalo (State University of New York), Cooke Hall, Buffalo, NY 14260

5 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 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, hypoxia,

ABSTRACT:

Foxa2 has garnered considerable interest in its pioneer functions and its role endoderm gut, and liver differentiation, and development, and initiator of gene regulatory networks (GRN) in these tissues. Although Foxa2 has been investigated in these systems, Foxa1 also compensates for its function, and thus may also have compensatory effects in studies of Foxa2 regulation. In this study, we focus on the role for both Foxa1 and Foxa2 in controlling endoderm GRN, liver activation in gut tube, and liver GRN. We first compare endoderm induction protocols, and develop a novel model of liver induction under hypoxic conditions that relies on minimal growth factors and enhanced morphology. We employ an RNAi (siRNA and shRNA) approach to demonstrate the effects of Foxa1/2 on endoderm induction, gut tube activation of the albumin gene. Our data demonstrates widespread regulation of the endoderm and mesendoderm GRN, and albumin, which is significant for activation of the liver differentiation program. We then analyze the Foxa1/2 phenotype in stable liver cell lines, and engineer stable cell lines that demonstrate potential reversibility of cell state. Finally, we perform RNA-seq and bioinformatics analysis that demonstrates global GRN changes and transcription changes due to Foxa1/2 perturbation, including expansive changes in cellular differentiation and metabolism. These data suggest that Foxa1/2 phenotype has time-dependent effects on GRN and widespread effects on GRN, the liver transcriptome, and liver metabolism.

INTRODUCTION:

Hepatic gene expression is precisely controlled by regulatory TFs which form a core GRN, well-studied, evolutionarily conserved, TFs that together form a GRN termed the hepatic nuclear factor network established in rodent and humans. The hepatic GRN include Foxa1, Foxa2, Foxa3, HNF1a, HNF1b, HNF4a, 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 improved understanding 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 promoter of the albumin gene 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 the promoters of hundreds of silent, liver specific genes, and it's activity was 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, with a lack of both alpha-fetoprotein (Afp) and albumin (Alb) expression, key markers in the endoderm and hepatic endoderm (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 at earlier stages. In the adult liver, 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 location studies indicate that ~1/4-1/5 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 HNF1a, HNF4a, and that the Foxa1/2 co-occupied sequences also had increased recognition motifs for HNF1a and HNF4a, indicating an active developmental genetic regulatory network (GRN) in adult mouse liver, support by another study (Watts, Zhang et al. 2011). Taken together, Foxa2 has profound effects on differentiation and metabolism in rodent systems, but less so in human cells.

GRN are less understood in human endoderm, gut tube, and hepatic progenitor cells than in murine models. A recent study performed a CRISPRi screen during endoderm induction demonstrated that Foxa1/2 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 occurs. However, these studies to do not employ targeting of both Foxa1 and Foxa2, even though they demonstrated effects on foregut induction and hepatic differentiation. Foxa1 priming of 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 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, this suggests that we need to better understand the role of Foxa1 and Foxa2 and the GRN they regulate in endoderm, gut tube endoderm, and hepatic endoderm.

In this study, we compare endoderm induction protocols, and develop a novel model of liver induction under hypoxic conditions that relies on minimal growth factors and enhanced morphology. We demonstrate the effects of Foxa1/2 on endoderm induction, gut tube activation of the albumin gene. Our data demonstrates widespread regulation of the endoderm and mesendoderm GRN, and albumin, which is significant for activation of the liver differentiation program. We then analyze the Foxa1/2 phenotype in stable liver cell lines, and engineer stable cell lines that demonstrate reversibility of cell state via Foxa1/2. Finally, we perform RNA-seq and bioinformatics analysis that demonstrates global GRN changes and transcription changes due to Foxa1/2 perturbation, including expansive changes in liver cellular differentiation and metabolism. These data suggest that Foxa1/2 phenotype has time-dependent, potentially reversible, effects on GRN and widespread effects on transcriptome and cell metabolism in hepatic cell lines.

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), and 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. 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), and epidermal growth factor (EGF) (Cat. #: 78006.1), and Gentle Cell Dissociation Reagent (GDR) (Cat: #: 07174), were purchased from StemCell Technologies. G418 (Roche) (Sigma Aldrich 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. 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. 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. Matrigel, Growth factor-free (Cat. #: 40230), was purchased from Corning. Bovine Serum Albumin (Cat. #: 10791-790) was purchased from VWR. CHIR (CHIR99021) (Cat. #: 13122) was purchased from Cayman Chemical. Rho-associated kinase (ROCK) inhibitor (Y27632 2, Cat. #: MBS577605) was purchased from Mybiosource.com. Fugene HD Transfection Reagent (Cat. #: E2311) was purchased from Promega. 96-well PCR plates (Cat. #: L223080), Tissue Culture Treated 24-well plate (LPS Catalog #702001), 75-cm2 Polystyrene tissue Culture-Treated Flasks (LPS Catalog#708003), PCR Plate Covers (LPS Catalog #HOTS-100) were purchased from Laboratory Product Sales, Inc. The pLKO.1 puro (Cat. #: 8453), pLKO-shFoxa1#1 (Cat. #: 70095), pLKO-shFoxa1#2 (Cat. #: 70096), and pLKO-shScramble. #: 1864) plasmids were purchased from Addgene. Human hepatoma (HepG2) (Cat. #: HB-8065) were purchased from ATCC. All primers were purchased from either Integrated DNA technologies (IDT), Sigma Aldrich, or Thermofisher.

Antibodies

Mouse anti-human beta actin (Thermo Fisher, Cat. #: MA5-15739), mouse anti-human Foxa2 monoclonal (Thermo Fisher, Cat. #: MA5-15542) were purchased from Thermo Fisher. Mouse anti-human albumin monoclonal (Abcam, Cat. #: ab10241) was purchased from Abcam. Mouse anti-human albumin monoclonal antibody (sc-271604) and mouse anti-human AFP monoclonal antibody (sc-130302) were purchased from Santa Cruz Biotechnology. Rabbit anti-human Sox17 antibody (NBP2-24568) was purchased from Novus Biologicals. Mouse anti-human Oct3/4 antibody (sc-5279) was purchased from

Santa Cruz Biotechnology. HRP-conjugated secondary antibodies were purchased from Jackson ImmunoResearch Laboratories, Inc. antibodies. Goat anti-mouse Secondary antibody (Alexa Fluor 488) (cat no. A-11034) was purchased from Thermofisher. Goat anti-rabbit Secondary antibody (Alexa Fluor 568) (cat no. A-11011) was purchased from Thermofisher. FITC and TRITC conjugated, anti-mouse, rabbit, and goat antibodies, and IgG control antibodies were purchased (Santa Cruz Biotechnology).

Preparation of stem cell differentiation medium (SFD medium)

Basal differentiation medium for human pluripotent stem cell 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 RPMI supplemented with 75 % IMDM, 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 commercially available induced pluripotent stem cell (iPSC) line (BXS0114 ACS1028 female (ATCC), and a NIH registered iPSC cell line (HUES53, (NIH registry 0055, male), a kind gift from Richard Gronostajski, PhD, University at Buffalo. UCSF4 cell line was a kind gift from Susan Fisher, PhD, UCSF Institute for Regenerative Medicine and stem cells. Human pluripotent stem cells (hPSC) were cultivated at 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^{o}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 medium was removed, 1mL of gentle cell dissociation reagent (GDR) was added for 10-15 minutes, and cells were removed from the dish. For passaging, the mTESR1 medium was removed, 1mL of gentle cell dissociation reagent (GDR) was added for 10-15 minutes, and cells were harvested from the dish. Cells were centrifuged 3 minutes at 800-1000 RPM (Ependorf 5810 table top centrifuge) and resuspended. Cells were frozen in mTESR1 medium + 5% DMSO at -80^{o}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 10uM 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 with adding 300uL 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. 500ul of Medium 1 (containing reagent A + B) was added from day 0 to day 1, and 500ul 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. In RPMI medium with B27 (no insulin) and 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 in RPMI + 1x B27 (no insulin), and 0.2% KO serum. Medium was changed daily and 500-750ul 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

Two protocols were used for liver differentiation.

Growth factor-free gut tube/ hepatic differentiation protocol: To promote spontaneous differentiation of human stem cells 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.

Stable liver cell line culture

Human hepatoma (HepG2) (Cat. #: HB-8065) cells were maintained with 10mL 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) (reconstituted in 10mM tris) treatment. RNA concentrations were determined by Nanodrop. RNA was converted to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad) and the mass of RNA was calculated by adding 5ng RNA per well to be run in the qPCR reaction. cDNA Reaction protocol was performed using 5 minutes at 25^o C, 20 minutes at 46^o C, and 1 minute at 95^{o}C. Reactions were then held either at 4^o C or on ice. We performed 10µL qPCR (3uL primers at a concentration of 0.3uM, 1uL nuclease free water, 1uL cDNA, and 5uL supermix) reactions with iTaq Universal SYBR Green Supermix (BioRad) in a 96-well PCR plate (LPS). The RT-qPCR reaction was done in a CFX96 Touch Real-Time PCR Detection System (BioRad). The RT-qPCR reaction consisted of Polymerase Activation and DNA Denaturation at 98^o C for 30 seconds followed by, 40 to 45 cycles of 98^{o}C for 15 seconds for Denaturation and 60^o C for 60 seconds for Annealing and Extension. Melt curve analysis was performed at 65-95^o C in 0.5^{o}C increments at 5 seconds/step. Relative, normalized, gene expression was analyzed using the delta-delta-Ct method, with three replicates per gene tested.

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 human stem cells (UCSF4) were plated for endoderm differentiation using the STEMDiff protocol. Prior to transfection, cells were washed with 500ul of fresh culture medium. Lipofectamine RNAiMAX reagent (1.5 ul) was mixed with Opti-MEM medium for a total of 50 ul, and 5 pmol siRNA (positive control, negative control, Foxa1, or Foxa2) was mixed with Opti-MEM medium for a total of 50 ul 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 ul of this mixture was added to a well of 24-well plate, mixed with 450 ul 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 siRNA 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 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 (#70095) and PLKO-shFoxa1#2 (#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 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 P/S 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 P/S 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 1.75ug/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 ug/μL plasmid solution was made up in OptiMEM (9.9ug 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 +1000ug/mL G418 (Roche).

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

For both transient transfection and stable cell line creation, knockdown was induced addition of Dox medium. The Dox medium contained a final doxycycline concentration of 1-2 ug/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 RT-qPCR.

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 500μL PBS/well unless otherwise specified. Cells were fixed with 500μL of 4% paraformaldehyde at room temperature under cell culture hood. Cells were washed again three times for 5 minutes each wash before incubating with 0.1% PBS-Triton X for 30 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 blocking buffer. 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 blocking buffer. 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 500µL 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 300µL of 0.05% Trypsin for 6-10 minutes. 500uL 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 move into 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, cells were washed twice. On the last wash, cell pellets were resuspended in 250µL 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, and 5µL of each sample or known standard was platted into each appropriate microplate well of a 96-well plate. 250µL 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 595nm 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 25µL, 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% Tween20) 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 0.1mL 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.

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 rlog transformation in DESeq2. The read counts of the genes were compared between the Foxa1/2 group and HepG2 control group with a Student's t-test and the genes were arranged from most upregulated to most downregulated. The GSEA software analyzed this list with gene sets from GO and KEGG. The FDR q-val and normalized enrichment score values from the results were used to filter the list for FDR values under 0.15 and absolute NES values above 1.45. The results were grouped based on upregulation and downregulation and were sorted based on SIZE, which is the number of genes within the gene set after filtering genes not in the expression dataset.

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 human stem cells towards definitive endoderm, gut tube endoderm, and along the hepatic lineage under low oxygen conditions.

We aimed to develop a reproducible hPSC-differentiation protocol to understand regulatory processes at the levels of gene regulatory networks that influence liver differentiation. We aimed differentiate hPSC towards definitive endoderm and then towards hepatic progenitors all under hypoxic ($5\%O_2$) incubator conditions, as occurs during 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 down regulation 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). We observed the maintenance of Foxa2, the downregulation of Sox 17, 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 gene regulatory network expression and hierarchy during endoderm induction, and its potential role in endoderm and liver differentiation, and we developed a RNAi transfection assay in hPSC-derived endoderm, followed by qRT-PCR for components of the endoderm and liver GRN. It's important to note that RNA-seq does not typically identify differences in GRN, unless increasing sequence depth is pursued, because of their low overall abundance in the transcriptome. 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 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, 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 Foxa1/2, GATA4, Hex, Hnf1b, and a trend in decrease in HNF6 (Fig 2A). Further, the data demonstrated a trending increase in HNF4a. The master TFs Sox 17 and Gsc demonstrated a trend for downregulation, while the gut tube and intestine patterning transcription factor Cdx2 demonstrates a trend towards upregulation as supported by mouse studies (Watts, Zhang et al. 2011), as does the master neuroectodermal transcription factor Pax6 (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 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), 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 Want stimulate primitive streak induction from hPSC, and 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 all three pluripotent stem cell lines (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 Sox 17 demonstrated downregulation of Oct4 and upregulation of Sox 17, 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 vi qRT-PCR demonstrated statistically significant decreases in Oct4, and corresponding statistically significant increases in Foxa2 and Sox 17, with clear trends in upregulation of Gata 4 and goosecoid (Fig. 3E). Finally, we analyzed the expression kinetics and demonstrated Oct4 downregulation and brachyury up and 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

Strong evidence exists for the regulation of liver development by Foxa2. 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, and 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 defined medium that has several medium supplements including Hams 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, 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 (Fig. 4C-D). We found that activation of albumin expression was significantly higher in the spontaneous differentiation condition than in the growth factor containing condition (Fig 4C), 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 as well as 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 puramycin 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, human pluripotent stem-cell derived gut tube endoderm, using scrambled sequences and shRNA to both Foxa1/2. 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. Importantly, albumin gene was significantly downregulated, when normalizing the data to pre-transfection across all experiments. Finally, Western blot demonstrated that Foxa2 is indeed downregulated in gut tube endoderm progenitor cells. We conclude that Foxa1/2 regulates albumin in 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 and Foxa2 regulate albumin expression in the developing gut tube differentiated towards liver, perhaps they can regulate albumin in established models of human liver regulation. 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 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 unaltered morphology after

transfection at 48h (Fig. 5A), that combined pools but not individual pools of siRNA lead to a ~50% knockdown (Fig. 5B-C), and that scrambled siRNA had no effect on major regulatory transcription factors and markers of differentiation (Fig. 5D). Importantly, we 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 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 of stable cell lines, we performed analysis of core regulatory transcription factors and other differentiation markers including pluripotency (Oct4, Nanog), neuroectoderm (Pax6), mesoderm (Bmp4), endoderm (Sox7, Mixl1, Gata4), pancreas (Pdx1), endoderm differentiation (CxcR4), epithelial state (E-Cad, N-Cad, and gut tube/liver (Hex, HNF4, HNF6). Compared to scrambled controls, we found statistically significant upregulation of Nanog, Pax6, Sox7, 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 validated that silencing of both Foxa1/2 is needed to observe the effect we identified.

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

Lentiviral shRNA was tested as a way to produce a stable line determine if sustained and long-term knockdown of shRNA in Foxa1/2 cells results in de-activation of the albumin gene and other members of the developmental, endoderm, and liver transcription factor (gene regulatory) networks. Optimization experiments were performed for each virus at MOI of 5-10 (Fig. 6A), again confirming that shFoxa1 #2

and shFoxa2 #1 viruses resulted in most downregulation of their target genes. We again performed analysis of core regulatory transcription factors and other differentiation markers including neuroectoderm, (Pax6), endoderm (Gsc), pancreas (Pdx1, Sox 9), biliary (CK19), endoderm differentiation (Cxcr4), and the endoderm/hepatic nuclear network (Foxa3, HNF4a, Hnf1b, Hex, Gata4, Tbx3). The data overwhelmingly demonstrated 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, statistically significant upregulation in Gsc expression (1.5-fold) and Pdx1 downregulation a (0.84-fold), and trending increases in Pax6 and decrease in Sox 9 expression (Fig 6B). 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 HNF4a (0.62-fold decrease), HNF1β (0.78-fold decrease), Hex (0.72-fold decrease), Tbx3 (0.36-fold decrease) (Fig. 5D). We also observed a statistically significant upregulation in Foxa3 and GATA4 (Fig. 5D). 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 markers like Foxa3 (hepatic endoderm), Afp (early hepatic progenitor), Cxcr4 (endoderm progenitors), an increase in Pax 6 (neuroectoderm), and increased biliary marker (CK19). These data demonstrate a coordinated decrease in liver differentiation, and evidence of immature states (hepatic endoderm), alternate fates (biliary, neuroectoderm) consistent with a differentiate block and subsequent response. Cells with scrambled virus demonstrated no changes in morphology (Fig 6F), but in shFoxa1/2 -/- cells, we routinely observed flat, elongated cells at low densities and at the edges of colonies that resolved at higher densities (Fig 6F, bottom left vs bottom right). 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 better understand 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 in these cells and determined that Dox concentration of 1.5 ng/ml was optimal in terms of toxicity and knockdown response. In traditional monolayer culture, we performed addition of Dox for 3 days, followed by removal for an additional 3 days, and performed qRT-PCR on days 0, 3, 6. In control cell

lines, we observed no changes in gene expression (data not shown). After Dox addition, we observed downregulation of Foxa1/, Foxa2, albumin, alpha-fetoprotein, Gata4, Hex, Hnf1b, and Hnf4a (Fig. 7B). After Dox removal for 3 days, we observed rapid over upregulation of Foxa1, Foxa2, albumin, Gata 4, Hex, a more moderate upregulation of Hnf4a, and a continued downregulation of Afp and Hnf1b (Fig. 7B). Thus, in most cases, the global changes were reversible. The morphological changes observed during knockdown, including the large, flat cells at the edges of colonies (Fig. 7F, lower right) were present during Dox addition, further demonstrating that global changes occurred in the cells. These data suggest that much of the effect of Foxa1/2 knockdown on HepG2 cells could be reversed over 3 days.

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 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. Each circle in the plot represents a color, or differentially expressed gene, if their adjusted p value was below 0.05, which shows the genes had significant differential expression. The log2-fold change determined whether the genes were up or down regulated. 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. Validating the RNA-seq data, we found Pax6, and in the downregulated genes, we identified albumin, both of which we had identified by qRT-PCR (Fig. 8C). A full list of the upregulated and downregulated genes is provided (Supplemental Fig 2). Among the top 16 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)). 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 RNA-seq data, not just differentially regulated genes, using a calculated normalized enrichment score (NES) with stringent criteria, the false discovery rate (FDR), and t-test between control and knockdown conditions. We obtained new gene lists then

performed KEGG pathways- Biological Process analysis (Fig 8D, Supp. Fig. 3). Amazingly, amongst the pathways, we observed many processes associated with stem cell 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. Further, for processes with downregulated genes, including those associated with stem cell differentiation, AP-specification, and gastrulation. Overall, this suggests that global block of the liver machinery greatly altered the differentiation trajectory, perhaps resulting in the activation of genes for alternate cell fates. Surprisingly, we identified many metabolic-related pathways during KEGG pathway-cellular component analysis (Fig 9A, Supp. Fig. 3). For example, when plotting just this class of cell component analysis, we identified approximately 1500 mitochondria-related genes that were altered, including 1000 genes associated with mitochondrial envelope, over 450 genes associated with mitochondrial matrix, and over 200 genes associated with mitochondrial protein complex. The rest of this list strongly suggested that the cells had not only been altered from the differentiation point of view, but also from the metabolic point of view. To further validate this, we again analyzed biological process lists of genes, associated with metabolism (Fig 9B, Supp. Fig. 3). Surprisingly, we observed wide ranges of metabolic processes that were downregulated, including over 1600 genes associated with nitrogen biosynthesis and 1200 genes associated with nitrogen catabolism, 900 genes associated with oxygen reduction processes, and 600 genes associated with generation of precursor metabolites. These data strongly suggest that the genes associated with metabolism are altered in the experimental condition.
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 they had a Foxa1/2, HNF4a, HNF1 binding site, and 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, again suggesting global downregulation by these factors.

DISCUSSION

We were interested in studying how gene regulatory networks (GRN) function in human cells to control cell fate. GRN are known to control early germ layer formation, mesendoderm and endoderm induction, cell specification from gut tube, including liver and pancreas. Here, we employed siRNA and shRNA to target Foxa2 and its compensatory molecule, Foxa1, in human stem cell derived endoderm, human gut tube endoderm progenitor cells, and in human liver stable cells lines. We tested Foxa1/2

combined knockdown in human pluripotent stem cell-derived endoderm, demonstrating wide-spread control of gene activation of mesendoderm, endoderm GRN. Suspecting that liver differentiation would be regulated, we targeted the activation of albumin gene, widely used to model the activation of liver differentiation, but traditional addition of growth factors did not result in liver differentiation under hypoxic conditions which emulate liver organ development. To do this, we 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. In this model, no specific liver differentiation growth factors are used. shRNA targeting Foxa1 and Foxa2 resulted in the downregulation of the albumin gene on day 14 culture, indicating potent control over albumin gene activation in differentiating gut tube progenitors. To further understand the effects of Foxa1/2 on liver differentiation, we employed the well-studied, stable liver cells 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 modulation of liver-specific GRN components including HNF4a, HNF1b, amongst others. RNA-seq analysis demonstrates that large changes to the transcriptome, and KEGG pathway analysis shows coordinated changes in stable liver cell lines, including upregulation of cell differentiation genes of other lineages and a subsequent vast downregulation of metabolism-related genes. These findings provide novel insights into Foxa1/2 complex functions in various stages of differentiation, including endoderm, gut tube, and liver.

We were particularly interested in the effects of Foxa1/2 on GRN in endoderm progenitors, differentiating gut tube endoderm, and in the stable liver cell line HepG2. Initial studies of GRN during mouse visceral endoderm differentiation of ESC in a Foxa2 -/- cells, demonstrated that is upstream of Foxa1, HNF1a, and HNF4a, and thereby controls differentiation and metabolism (Duncan, Navas et al. 1998). Foxa2's role in activating 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 siRNA data of endoderm induction are consistent with the concept that in mesendoderm, Foxa1/2 may control 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 2). Further, downregulation of Foxa2 increased Cdx2 expression (Fig 2), which we have observed at later stages in our cultivation of gut tube progenitor cells at low oxygen (Fig 4D). Therefore, this data could suggest that low expression of Foxa2 in gut tube progenitor cells in our cultivation system, which we have observed, leads to the unwanted upregulation of the gut patterning factor Cdx2. As we did not perform detailed CHIP analysis, we must speculate whether this regulation is via Foxa2's pioneer TF function, or conventional transactivation/repression, or through more indirect means. Future studies of the endoderm TF network can help determine this. Interestingly, we also observed small amounts of albumin that present in our day

4 endoderm culture, that was regulated by Foxa1/2. It is unclear whether this represents small numbers of cells that are farther ahead in developmental stage, or whether albumin gene is active at low levels within endoderm cells, and whether convention or pioneer activity is responsible for its control.

The role of Foxa1/2 in differentiating gut tube progenitors also remains to be determined. 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). 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). These studies performed knockout of Foxa1 but not both Foxa1/2 in gut tube endoderm, and therefore it remains unknown whether Foxa2 first binds prior to enhancers is precedes H3K4me1 levels. Our data in gut tube endoderm suggest that shRNA for Foxa1/2 down regulation on day 6, can lead to 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 previously reported (Wang, Yue et al. 2015), thereby enabling further dissection of its molecular regulation.

One of the main findings in this study was the identification of that Foxa1/2 downregulates liver-specific genes in differentiating gut tube endoderm and in stable liver cell lines. 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 HNF4a and HNF1b in HepG2 cells, but reports a decrease in HNF6, which we did not observe. Their investigation demonstrated that they 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 or conditional vectors. Nonetheless, they demonstrated direct 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. On the other hand, Foxa1 and Foxa2 have been shown to have separate functions in the adult liver, with Foxa1 binding sites are associated with p53 binding sites, while Foxa2 is associated with hepatic metabolism (Bochkis, Schug et al. 2012). Based on these studies,

it is hard to predict how Foxa1/2 would affect gene expression. While we did expect changes in metabolic genes, we were surprised how the downregulation of Foxa1/2 resulted in global changes in metabolic genes in these cells, and the upregulation of alternate cell differentiation genes. This effect remains to be studies by employing metabolic analyses of these cells.

There are limitations to our study worth mentioning here. 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.

Nonetheless, we have provided evidence that time dependent GRN networks can be controlled by Foxa1/2, and that it can be have potent, genome-wide effects on liver differentiation and metabolism. 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 pancreas and lung. It remains to be seen how understanding of GRN will enable improved use of stem cells for therapy, and also our understanding of cancer differentiation and metabolism. Our studies indicated time-dependent Foxa1/2 effects, in human cells, need to be better understood to improve differentiation protocols, to balance epigenetic and genetic changes that occur during differentiation, and to avoid aberrant differentiation from pluripotent stem cells.

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, the New York State Stem Cell Science C024316 and the Stem cells in regenerative medicine (ScIRM) center. MM was supported by ScIRM program. OO was supported by the Western NY Prosperity Fellowship.

Figure Legends:

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. B) Morphology and immunochemistry of endoderm and hepatic progenitor cells. Left panel- Phase contrast image of Day 4 endoderm morphology. Bar = 200 um. 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 demonstrating elongated cells and prominent clusters of cells throughout. C) Gene expression (qRT-PCR) of expression kinetics of Oct4, brachyury, and goosecoid (gsc) during 4 days of endoderm induction. D) Same as C, except data demonstrates day and day 4 gene expression. $p<0.05$, * denotes significance. E) Same as D, except data demonstrates endoderm TF (Sox 17), endoderm and liver TF (Foxa2), and liver differentiation genes (Alpha-fetoprotein and albumin).

Figure 2. Gene expression of gene regulatory network in siRNA Foxa1/2 phenotype during endoderm induction from human stem cells. A) Gene expression (qRT-PCR) of endoderm and liver transcription factors (TFs) during endoderm induction from human stem cells. p values as shown. B) Same as A except analysis of master TFs in endoderm, gut tube, and neuroectoderm, p values as shown C) Same as B except endoderm and liver differentiation genes examined, p values as shown.

Figure 3. Endoderm induction from human stem cells in a hypoxic, growth factor-free, liver differentiation protocol. A) Human pluripotent stem cell colonies of iPSC and ESC were used in these studies. The data shown below is for the iPSC cell line. B) Schematic of differentiation system. iPSC were cultured under hypoxia on GF-Free coated matrigel plates, then replated on freshly coated plates in the presence of Activin/CHIR followed by Activin over 4 days. Medium contains RPMI, 1x B27, and 0.2% KOSR for improved cell survival. C) Immunofluorescent staining. Top two panels- Day 0, pluripotent colonies for Oct4/DAPI and Sox17/DAPI, bar = 400 um. Middle two panels- Day 4 cells stained for Oct4/DAPI and Foxa2/DAPI, bar = 200 um. Bottom two panels-Day 4 cells stained for Sox17/DAPI, and AFP/DAPI, bar = 200 um. D) Morphological analysis of cells during endoderm induction. Day 4 cells (right panel) demonstrate bright cell borders. E) Gene expression (qRT-PCR) analysis of gene expression comparing day 0 and day 4 for several endoderm genes. $P< 0.05$, * is significance. F) Gene expression (qRT-PCR) analysis of gene expression comparing expression kinetics of Oct4 (left) and Brachyury (right) from day 0 to day 4.

Figure 4. Hepatic differentiation from human stem cells in a novel, simplified, growth factor-free, liver differentiation protocol under hypoxia. A) Schematic of differentiation system. hPSC were differentiated

towards endoderm, then gut tube endoderm, and then hepatic progenitor cells under 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) Morphology of day 4 (endoderm) bar = 400 um, day 6 (gut tube endoderm) bar = 200 um, and day 14 progenitors bar = 200um, demonstrates bright cell borders, cuboidal morphology, consistent with hepatic progenitors. C) Gene (qRT-PCR) expression demonstrates kinetics of endoderm day 1- day 4 for Oct 4 and brachyury. D) High albumin gene expression (qRT-PCR) in the GF (-) compared to the GF (+) condition, on a log plot. $p < 0.05$. Cdx2 expression demonstrates variability between GF (-) and GF (+) conditions, p = NS. E) Gene (qRT-PCR) expression of Foxa1 and albumin (Alb) on day 14 after shRNA Foxa1/2 knockdown on day 6 of culture (n=3), $p < 0.05$. F) Western blot after two days after shRNA Foxa1/2 transduction into day 6 gut tube progenitor cells demonstrating Foxa2 knockdown.

Figure 5. siRNA of Foxa1/2 regulates differentiation genes and GRN. Human HepG2 cells were transfected for siRNA via reverse transfection. A) siRNA knockdown of Foxa1/2 in Hepg2 cells showed no morphology changes when compared to Hepg2 control cell cultures. B) qRT-PCR expression kinetics for siRNA for Foxa1(n=1) after 24 and 48-hour time points. C) Same as B except siRNA for Foxa2. D) qRT-PCR expression data (n=1) for optimization of pmole siRNA added for pooled siFoxa1, and pooled siFoxa2, ranging from 3-9 pmoles. Several genes tested in addition to Foxa1 and Foxa2. E) qRT-PCR expression data (n = 5) for siFoxa1/2 knockdown in HepG2 assayed after 48 hours, compared to scrambled controls. Significance values were calculated using t-test (2 Tails, Type 2) for Foxa1, Foxa2, alpha-fetoprotein (Afp) and albumin (Alb). F) To confirm the results of qRT-PCR, Western blot was done demonstrating protein level knockdown of Foxa2 in siFoxa1/2 (-/-) cell lines compared to scrambled siRNA controls. G) Immunohistochemistry of albumin in siFoxa2 -/- cell lines. From left to right- Epi-illumination image of fixed HepG2 cells. FITC labeling of Foxa2 cells after scramble transfection. Secondary antibody staining only. siFoxa1/2 (-/-) knockdown cells after staining. Data demonstrates minimal FITC signaling in siFoxa/2 (-/-). H) qRT-PCR expression for scramble and siFoxa1/2 (-/-) endoderm (Gsc), pancreas (Pdx1), mesoderm (Bmp4), and epithelial cell state (N-cadherin, E-cadherin) after 48 hours. I) Same as above (H) except for members of the pluripotency , endoderm, and hepatic nuclear network. Criteria was $p < 0.05$ for n =3.

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) Gene expression (qRT-PCR) after shRNA

transduction with shFoxa1, shFoxa2, shFoxa1/2, and scrambled shRNA control, followed by antibiotic selection. n=1 experiments. B) Gene expression (qRT-PCR) after shRNA transduction with shFoxa1/2, and scrambled shRNA control, followed by antibiotic selection. n=3 experiments, significance denoted by "*" for $p<0.05$. C) Same as B, except different gene set. D) Same as C, except different gene set. E) Western blotting for scrambled shRNA (control) and n = 2 cell lines. Gene of interest and housekeeping gene shown for each plot. F) Morphology comparison between top two panels (shRNA scramble at MOI of 5 and MOI of10) and bottom two panels (shRNA Foxa1/2). 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. Conditional shRNA expression of Foxa1/2 -/- phenotype demonstrates gene reversibility. Cells were transfected with Doxycycline (Dox)-On shRNA vector (shRNA Foxa2#1) targeting both shRNA for Foxa1 and Foxa2. A) First (1) shRNA was cloned into the Dox-on vector. Next, (2), DNA was transfected and selected for by antibiotic or shRNA scrambled controls, continuously selected after several passages with puramycin selection, and analyzed. Next (3) 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) Gene expression (qRT-PCR) after Dox addition (day 3) and Dox removal (day 6).

Figure 8. RNA-seq analysis of control (HepG2) and experimental (shFoxa1/2 -/-) genes. A) Principal component analysis with the three experimental samples and two control samples. 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) 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. Plot of biological process gene sets selected for relationship to stem cells and differentiation. Number of genes within group shown after calculation, t-test comparison between groups ($P<0.05$), for n = 2 control, n = 3 Foxa1/2 -/- cells. Criteria was NES -1.5< or > 1.5, and false discovery rate (FDR)-qval <0.1).

Figure 9. RNA-seq and bioinformatics analysis of Foxa1/2 -/- cell lines demonstrates global regulation of metabolic genes. A) Bioinformatics analysis of RNA seq data in which number of genes within each significant gene set is shown related to metabolism. 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. Plot of cellular component gene sets selected for relationship to metabolism. Number of genes within group shown after calculation, t-test comparison between groups ($P<0.05$), for n = 2 control, n = 3 Foxa1/2 -/- cells. Criteria was NES -1.48 < or > 1.5, and false discovery rate (FDR)-qval <0.09). B) Same as above, except Biological process gene sets, selected for relationship to metabolism. Criteria was NES -1.52 < or > 1.75, and false discovery rate (FDR)-qval <0.09).

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.

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.

Du, Y., J. Wang, J. Jia, N. Song, C. Xiang, J. Xu, Z. Hou, X. Su, B. Liu, T. Jiang, D. Zhao, Y. Sun, J. Shu, Q. Guo, M. Yin, D. Sun, S. Lu, Y. Shi and H. Deng (2014). "Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming." Cell Stem Cell **14**(3): 394-403.

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.

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.

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.

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.

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.

Tang, Y., G. Shu, X. Yuan, N. Jing and J. Song (2011). "FOXA2 functions as a suppressor of tumor metastasis by inhibition of epithelial-to-mesenchymal transition in human lung cancers." Cell Res **21**(2): 316-326.

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.

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.

Yu, B., Z. Y. He, P. You, Q. W. Han, D. Xiang, F. Chen, M. J. Wang, C. C. Liu, X. W. Lin, U. Borjigin, X. Y. Zi, J. X. Li, H. Y. Zhu, W. L. Li, C. S. Han, K. J. Wangensteen, Y. Shi, L. J. Hui, X. Wang and Y. P. Hu (2013). "Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors." Cell Stem Cell **13**(3): 328-340.

Zaret, K. S. (2002). "Regulatory phases of early liver development: paradigms of organogenesis." Nat Rev Genet **3**(7): 499-512.

A

Low O_2 (5%) differentiation ----------►

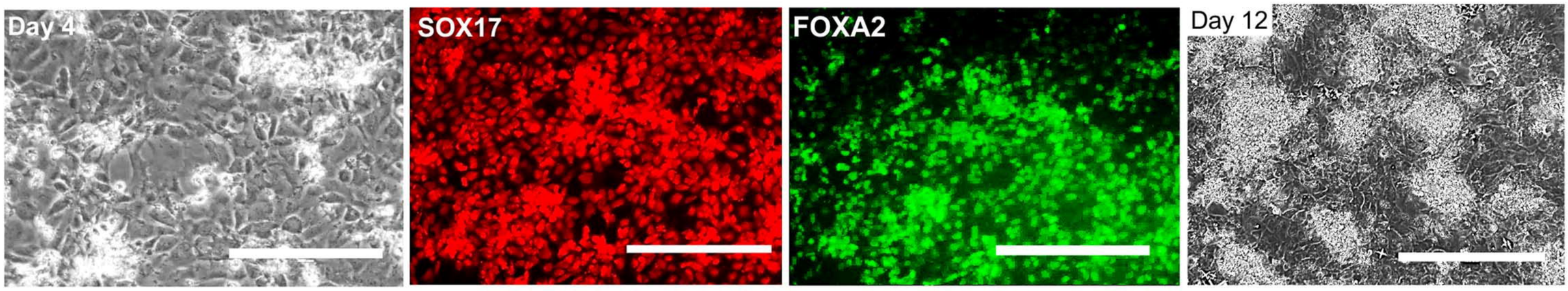
Growth factor containing (GF (+))

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)

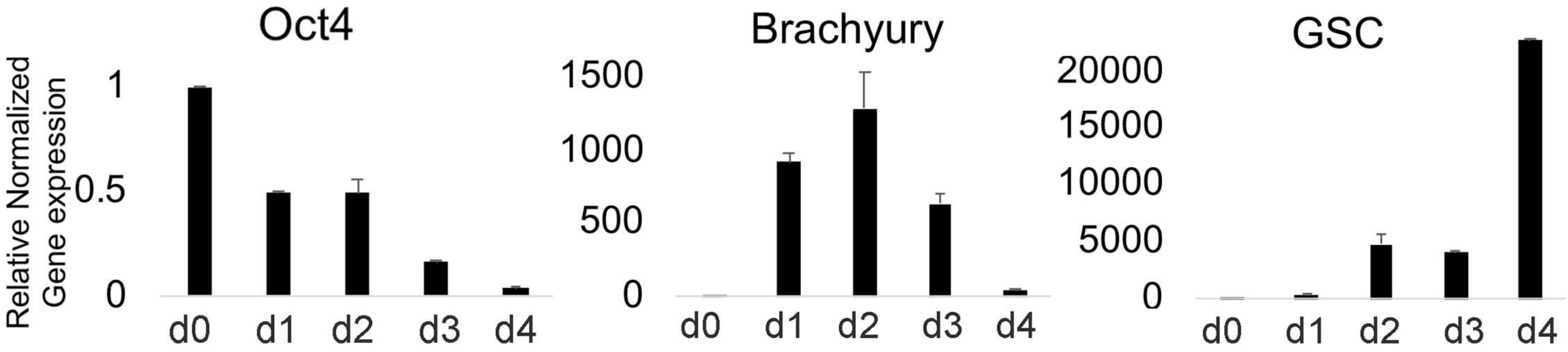
STEMDiff kit Endoderm induction → KGF/FGF7 (25 ng/ml) → BMP4 (10 ng/ml) FGF2 (20 ng/ml) → HGF (10 ng/ml) Oncostatin (20 ng/ml) Dexamethasone (100 nm)

d0 → d4 → d5 → d7 → d10 → d12

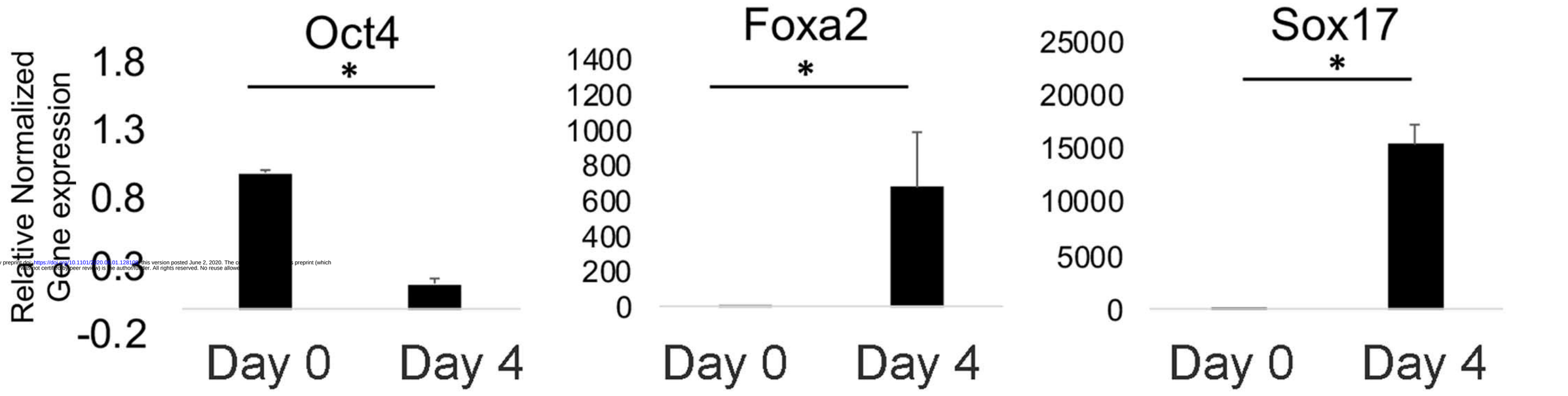
B



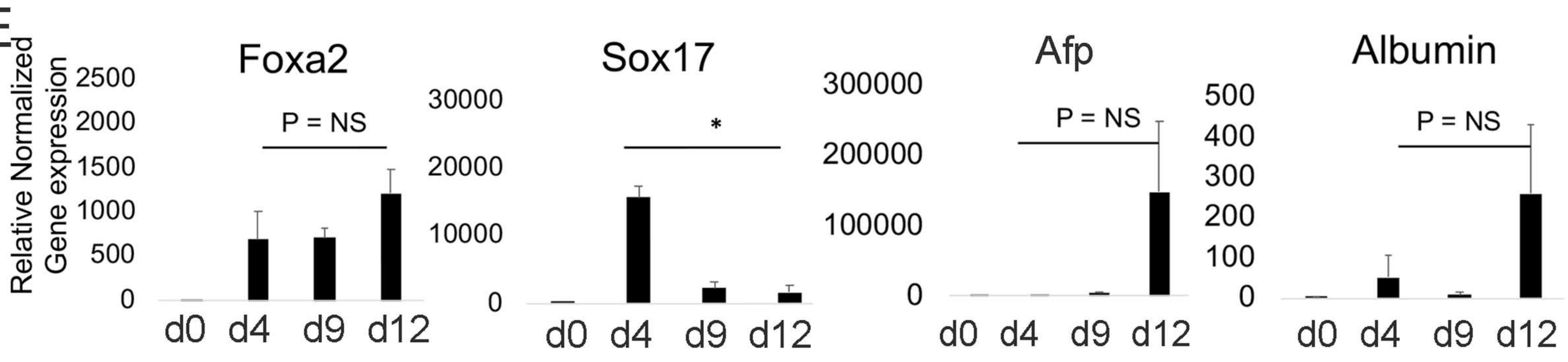
C

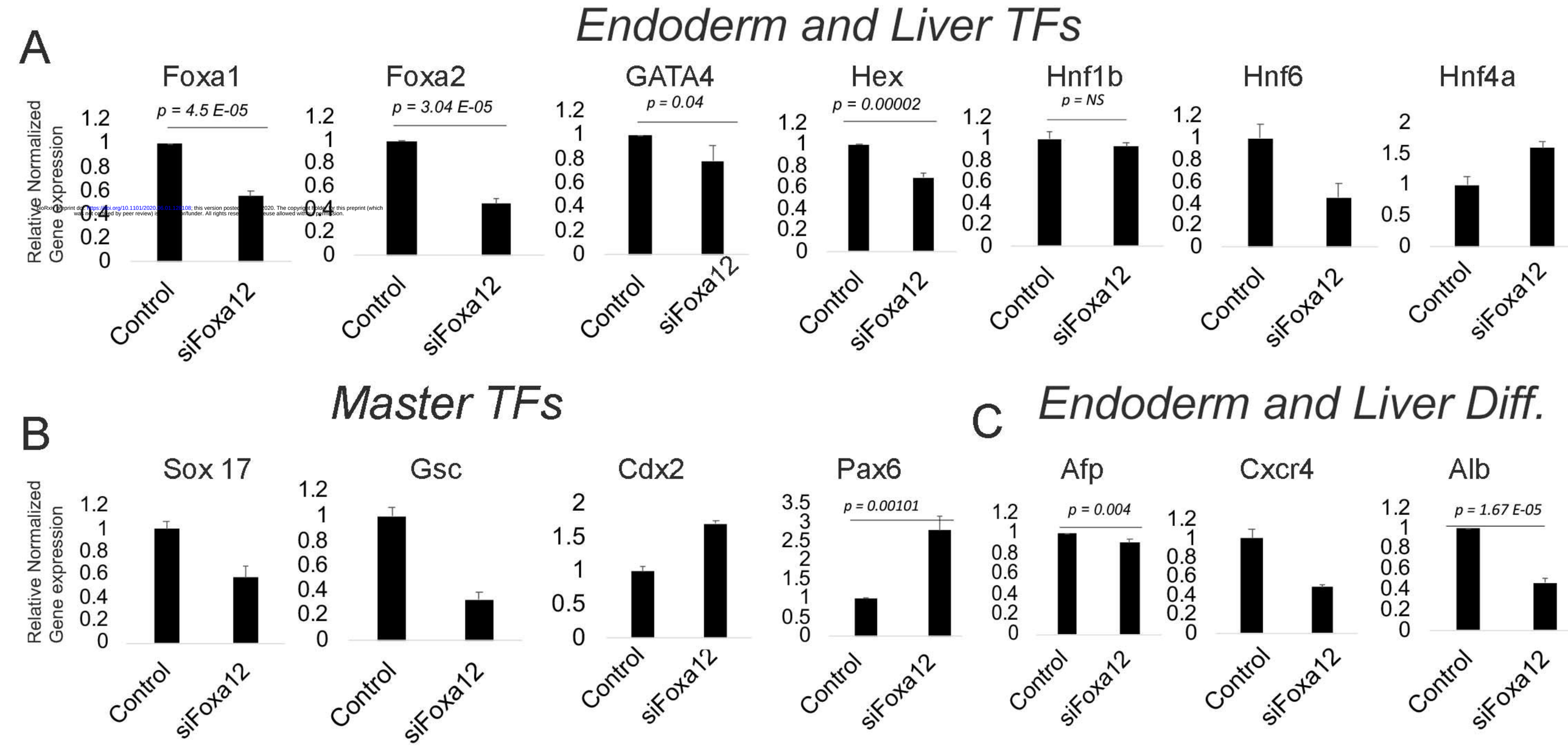


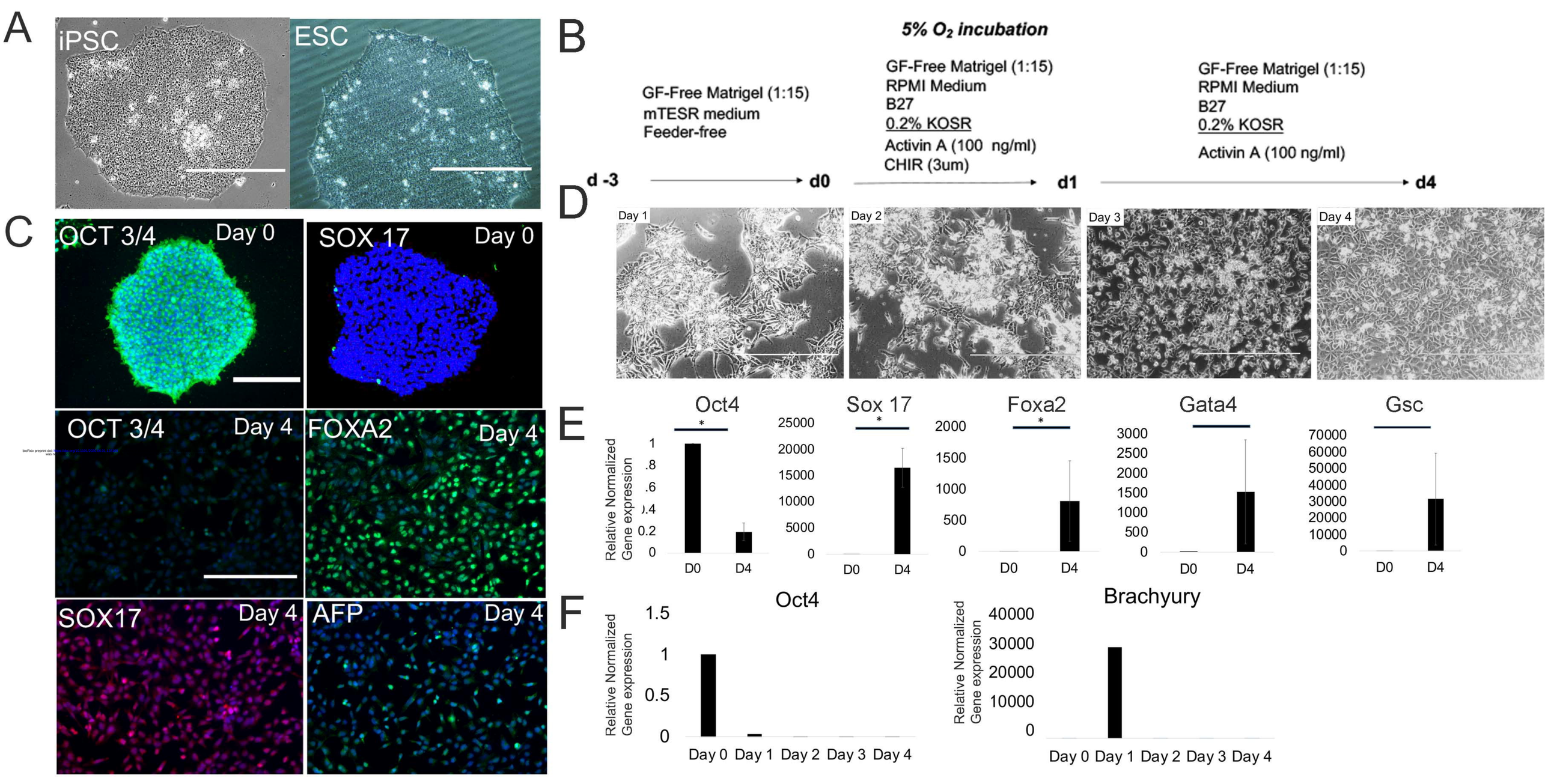
D



E






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
D0
D4
F
Oct4
Brachyury
Relative Normalized Gene expression
Day 0
Day 1
Day 2
Day 3
Day 4

A

Low O_2 (5%) differentiation

Growth factor containing (GF (-))

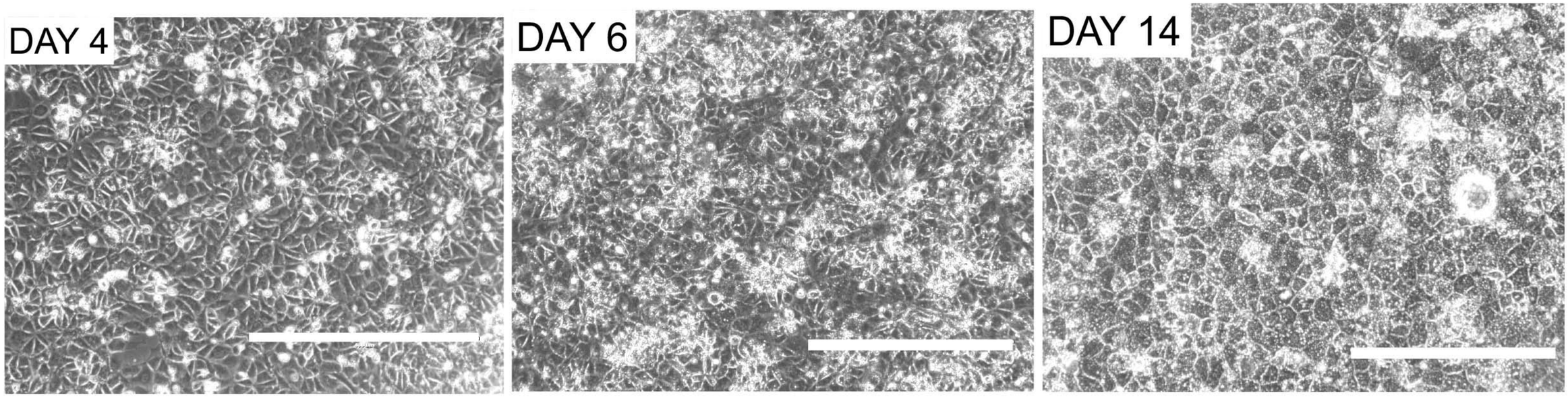
SFD Medium (75% IMDM, 25% Hams F12, 1x N2, 1x B27, FGF2 (2.5 ng/ml), 1% Penn/Strep, 0.05% BSA, 2 mM Glutamine, 0.5 mM Ascorbic Acid, 0.5 mM Monothioglycerol)

GF-Free Matrigel (1:15)
RPMI Medium
1x B27
0.2% KOSR
Activin,Activin/CHIR

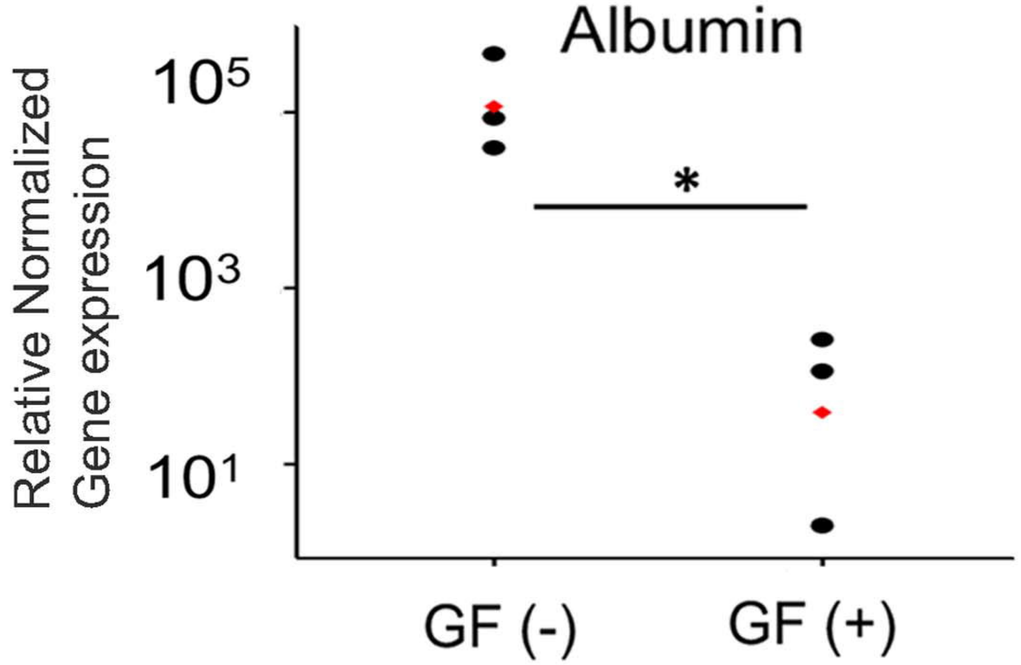
KGF/FGF7 (25 ng/ml) KGF/FGF7 (25 ng/ml) KGF/FGF7 (25 ng/ml)

d0 → d4 → d5 → d7 → d10 → d14

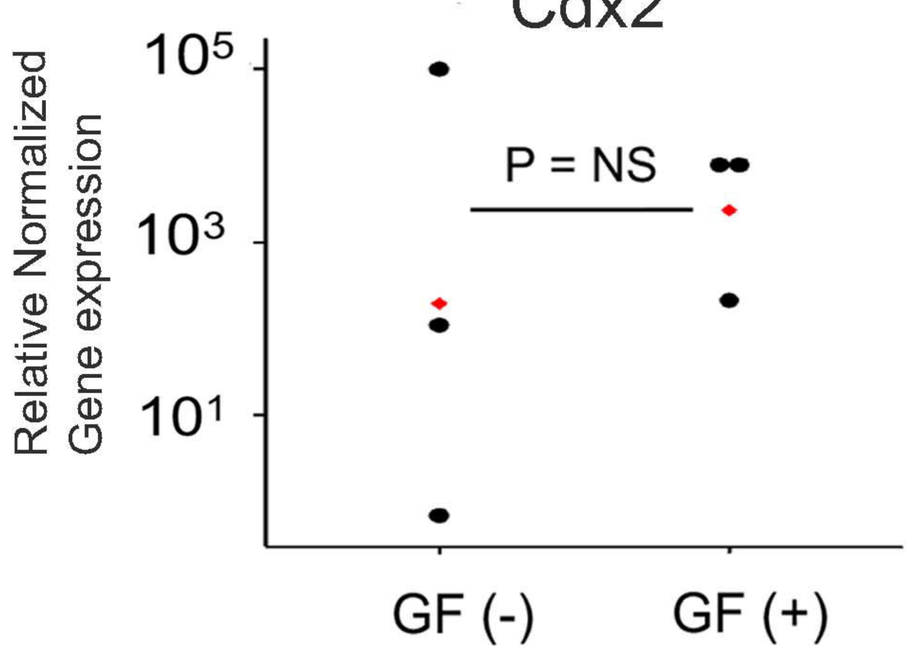
B



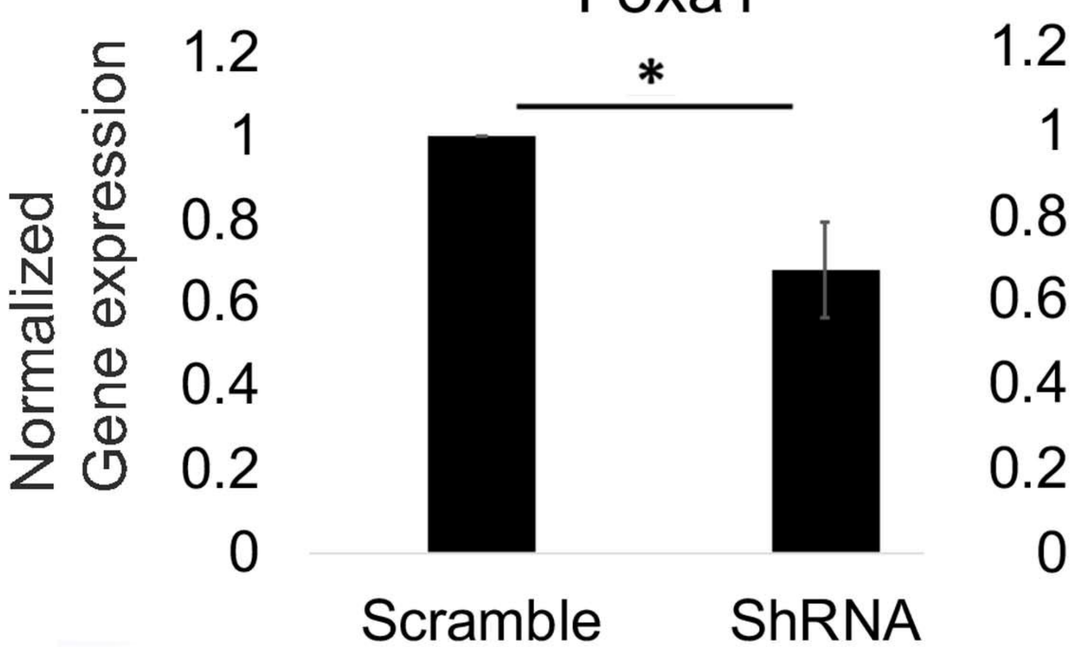
C

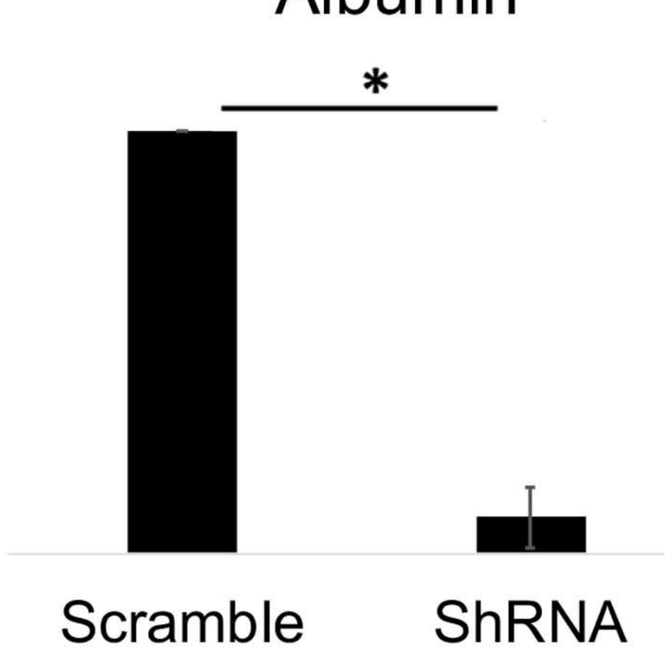


D

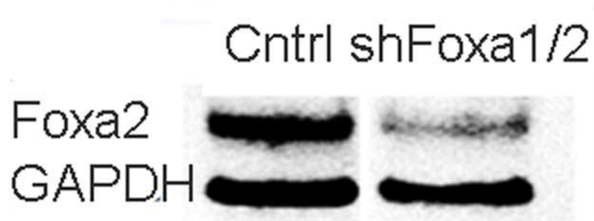


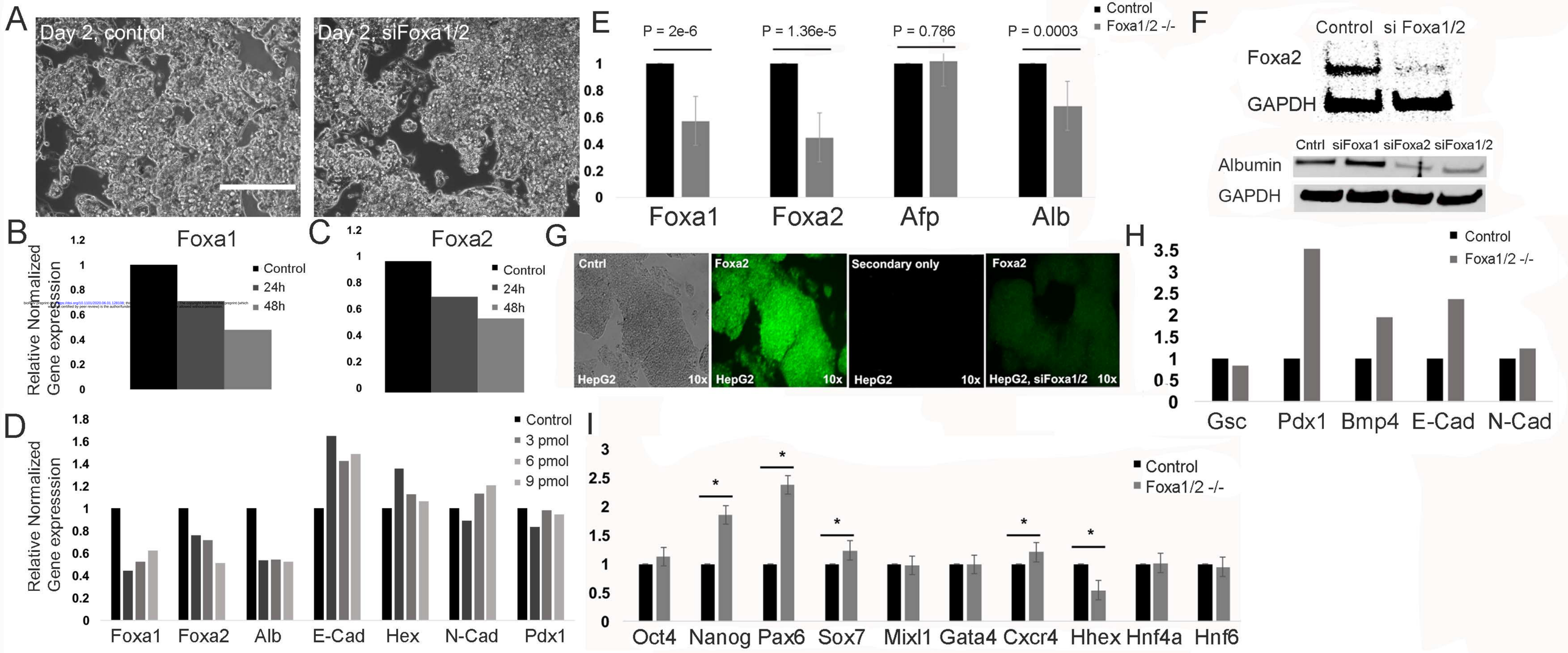
E



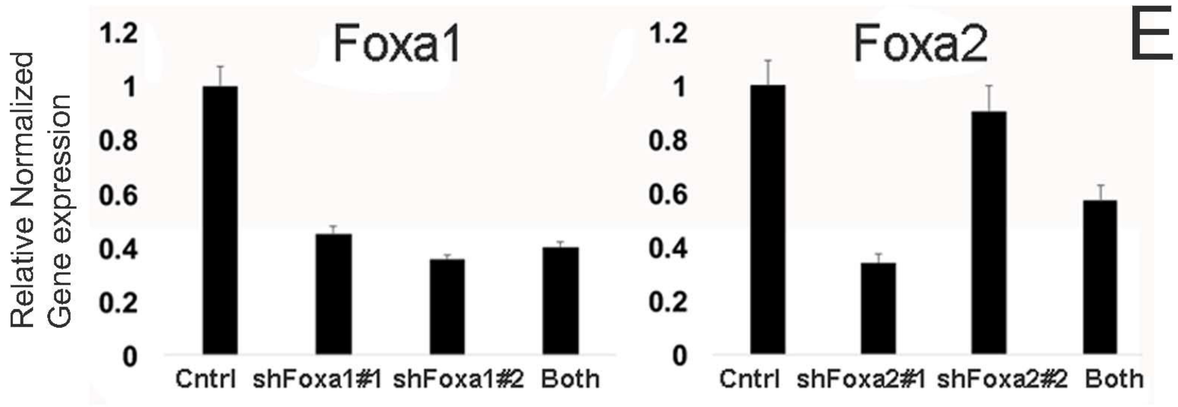


F

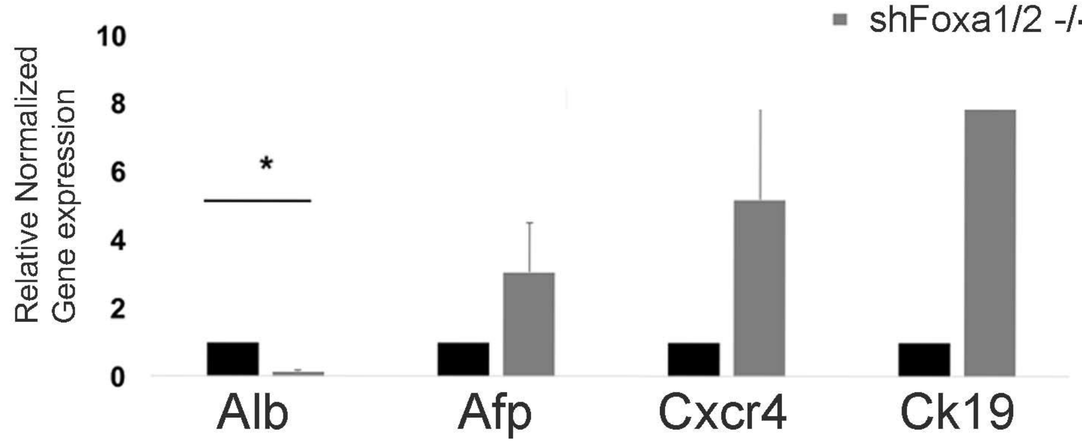



A
Day 2, control
Day 2, siFoxa1/2
B
Foxa1
Relative Normalized Gene expresssion
Control
24h
48h
C
Foxa2
Control
24h
48h
D
Relative Normalized Gene expresssion
Control
3 pmol
6 pmol
9 pmol
Foxa1
Foxa2
Alb
E-Cad
Hex
N-Cad
Pdx1
E
Control
Foxa1/2 -/-
P = 2e-6
P = 1.36e-5
P = 0.786
P = 0.0003
Foxa1
Foxa2
Afp
Alb
F
Control
si Foxa1/2
Foxa2
GAPDH
Cntrl
siFoxa1
siFoxa2
siFoxa1/2
Albumin
GAPDH
G
Cntrl
HepG2
10x
Foxa2
HepG2
10x
Secondary only
HepG2
10x
Foxa2
HepG2, siFoxa1/2
10x
H
Control
Foxa1/2 -/-
Gsc
Pdx1
Bmp4
E-Cad
N-Cad
I
Control
Foxa1/2 -/-
Oct4
Nanog
Pax6
Sox7
Mixl1
Gata4
Cxcr4
Hhex
Hnf4a
Hnf6

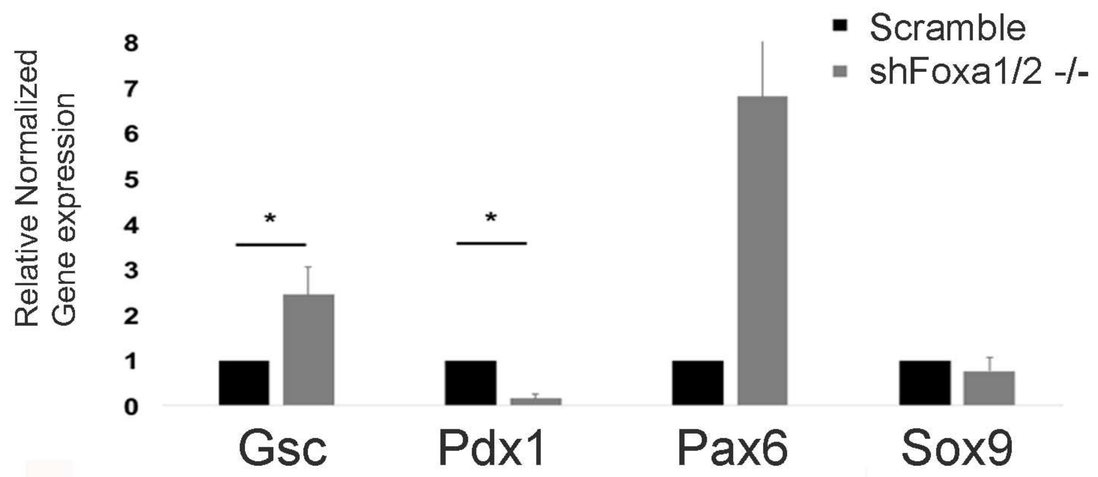
A



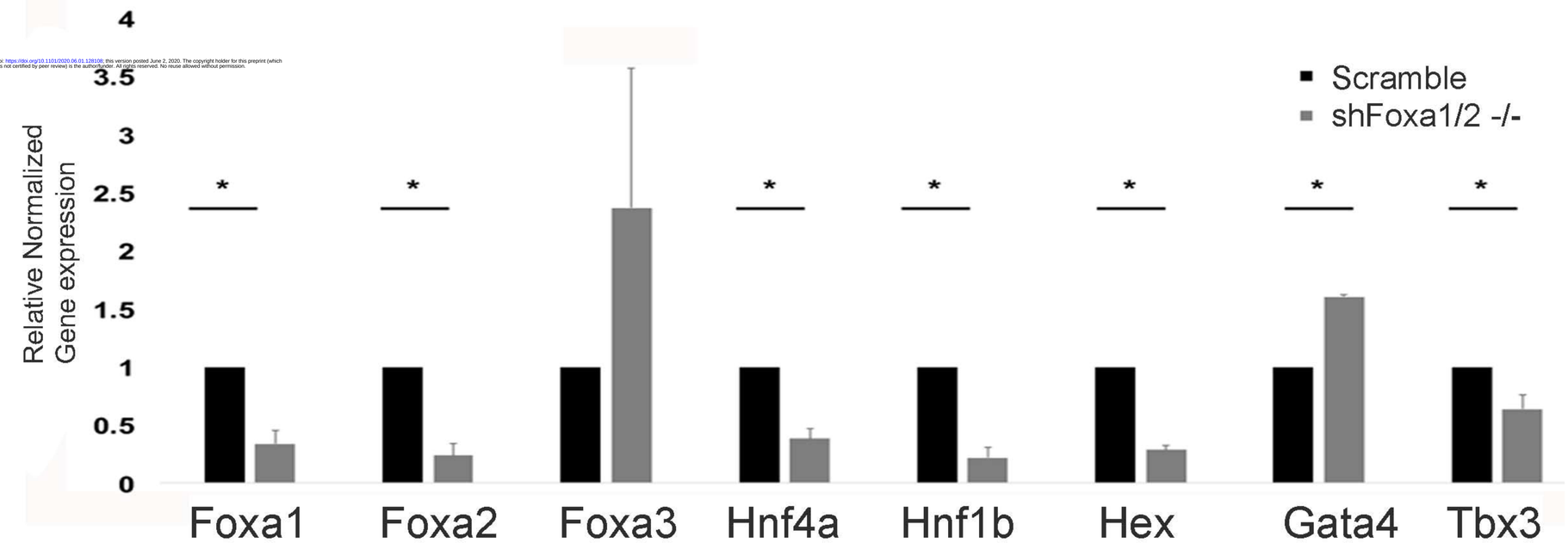
B



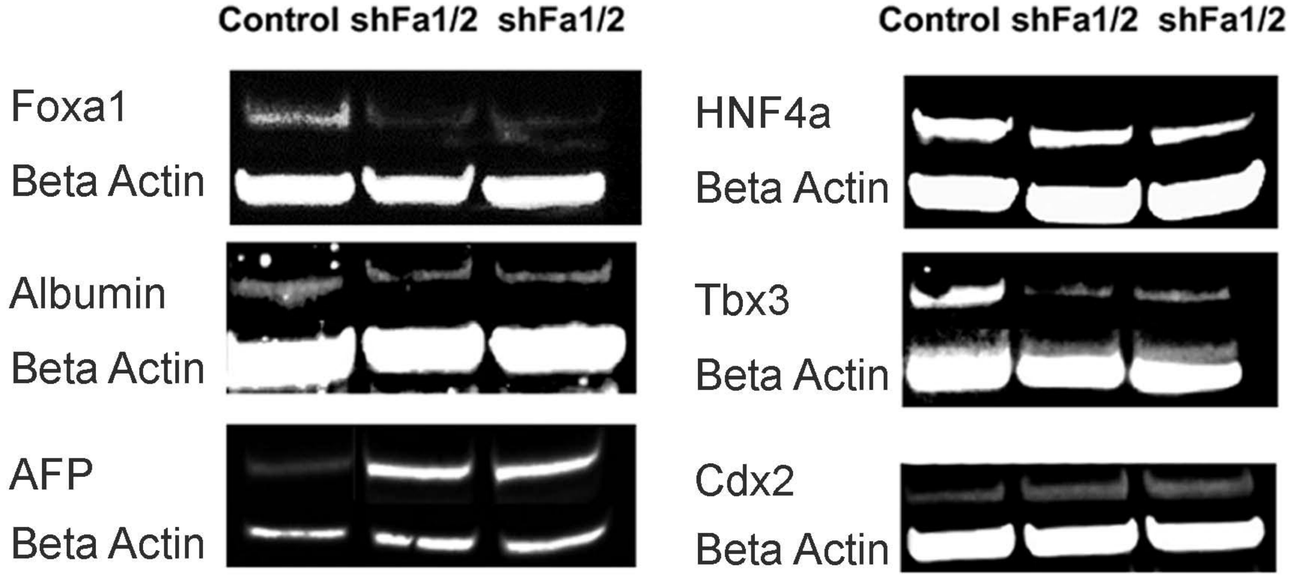
C



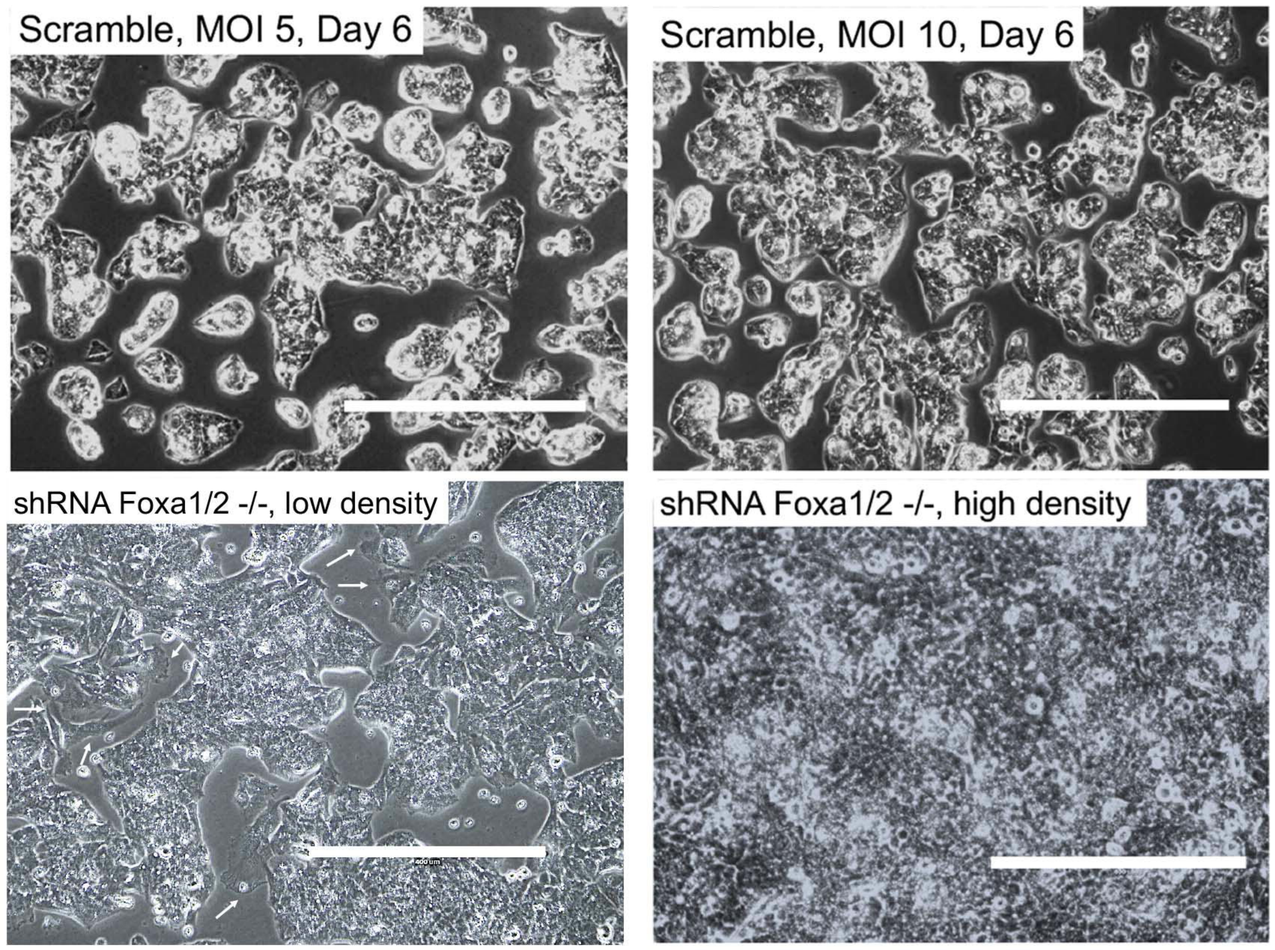
D

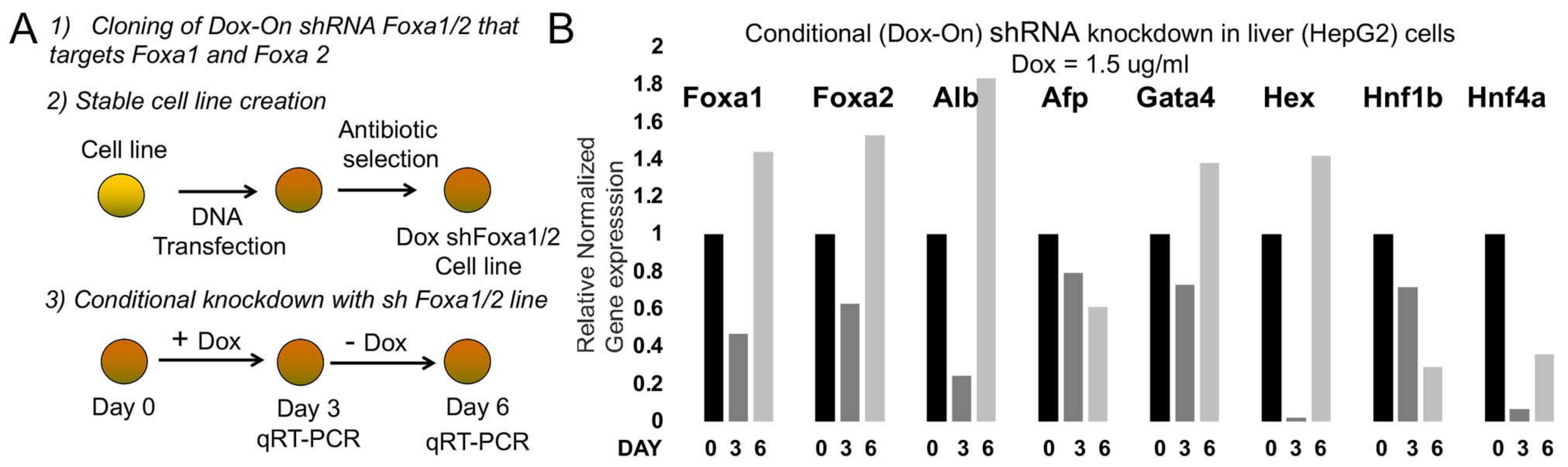


E

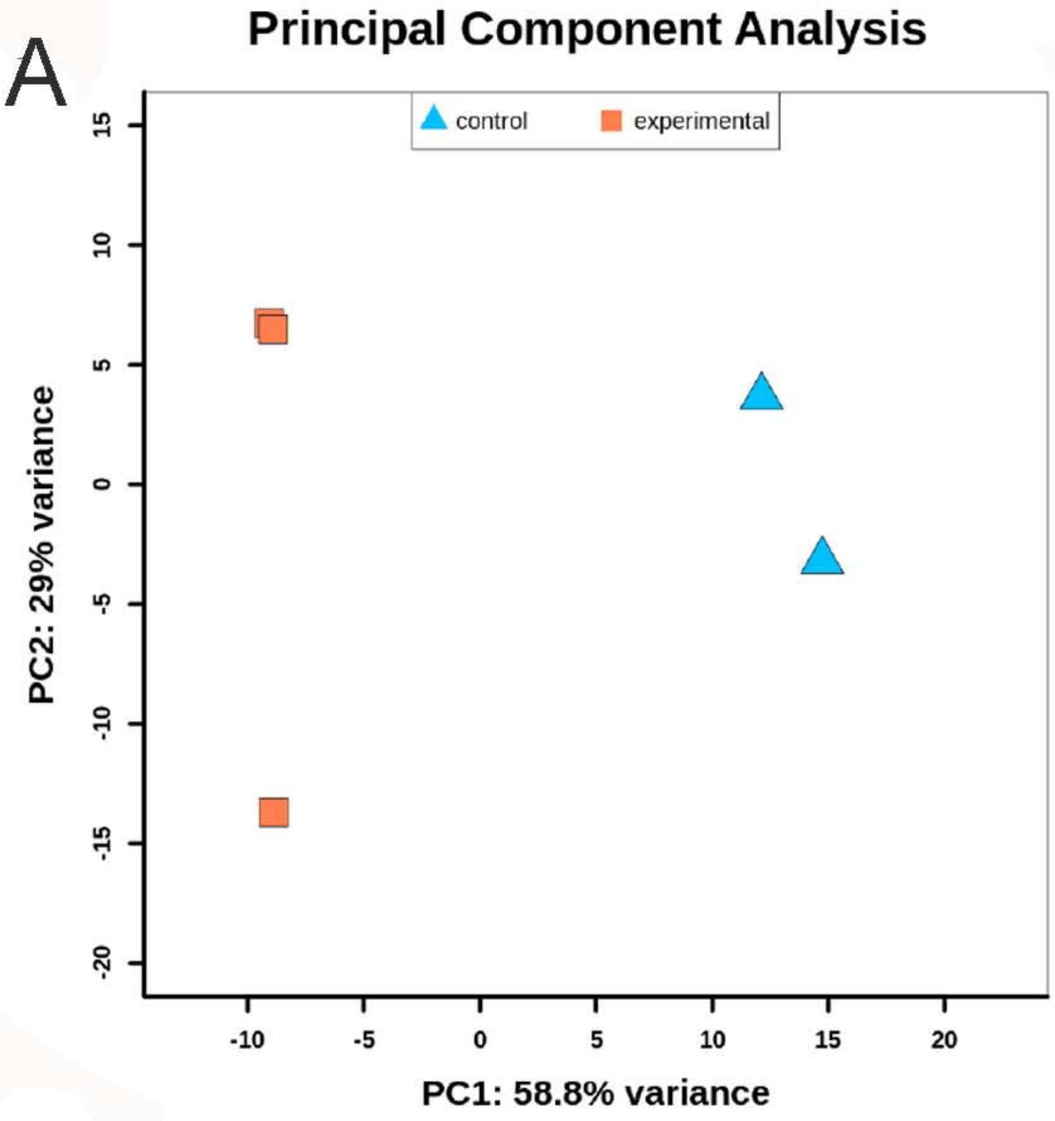


F

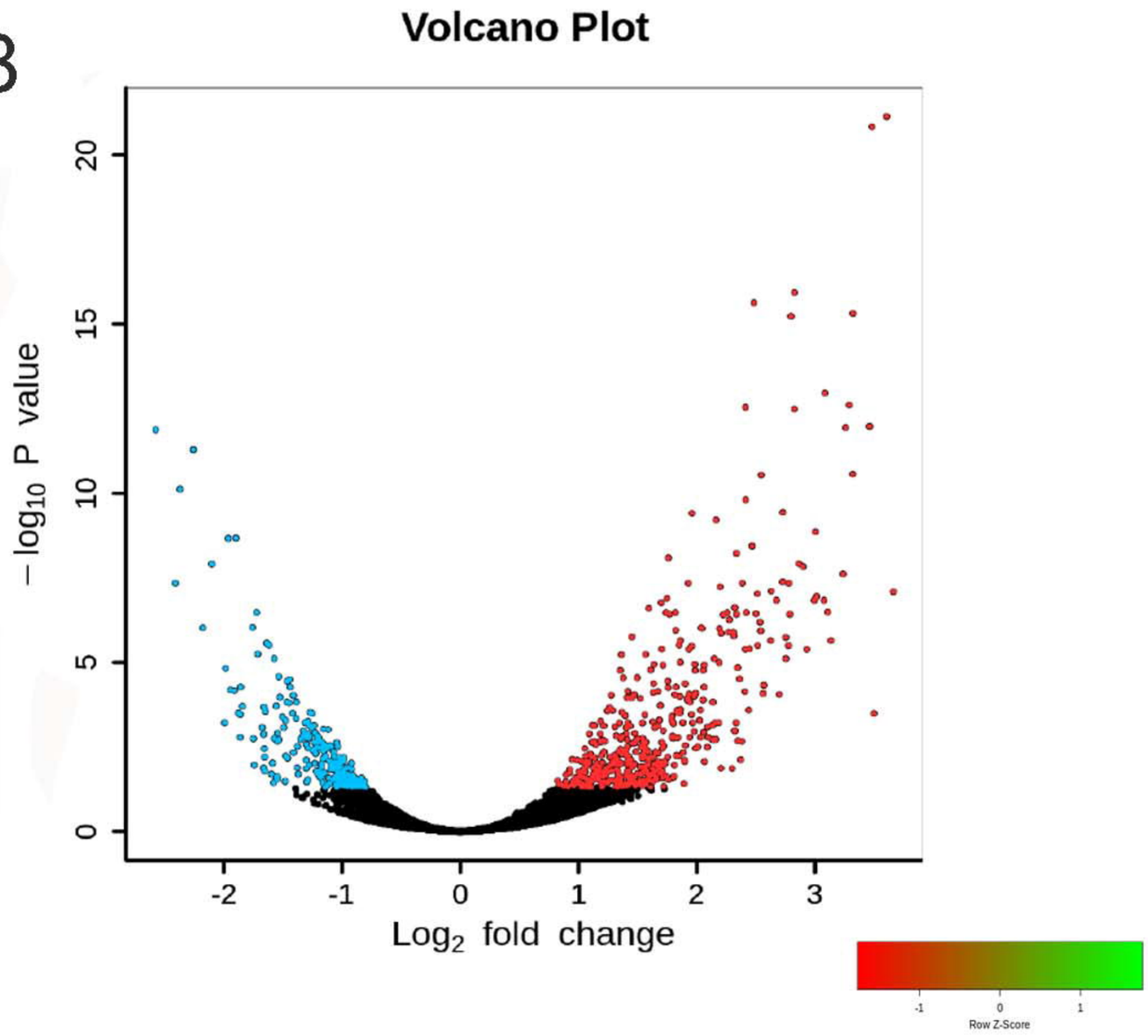



A
1) Cloning of Dox-On shRNA Foxa1/2 that targets Foxa1 and Foxa 2
2) Stable cell line creation
Cell line
DNA Transfection
Antibiotic selection
Dox shFoxa1/2 Cell line
3) Conditional knockdown with sh Foxa1/2 line
+ Dox
- Dox
Day 0
Day 3 qRT-PCR
Day 6 qRT-PCR
B
Conditional (Dox-On) shRNA knockdown in liver (HepG2) cells
Dox = 1.5 ug/ml
Relative Normalized Gene expresssion
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
DAY
Foxa1
Foxa2
Alb
Afp
Gata4
Hex
Hnf1b
Hnf4a
0 3 6

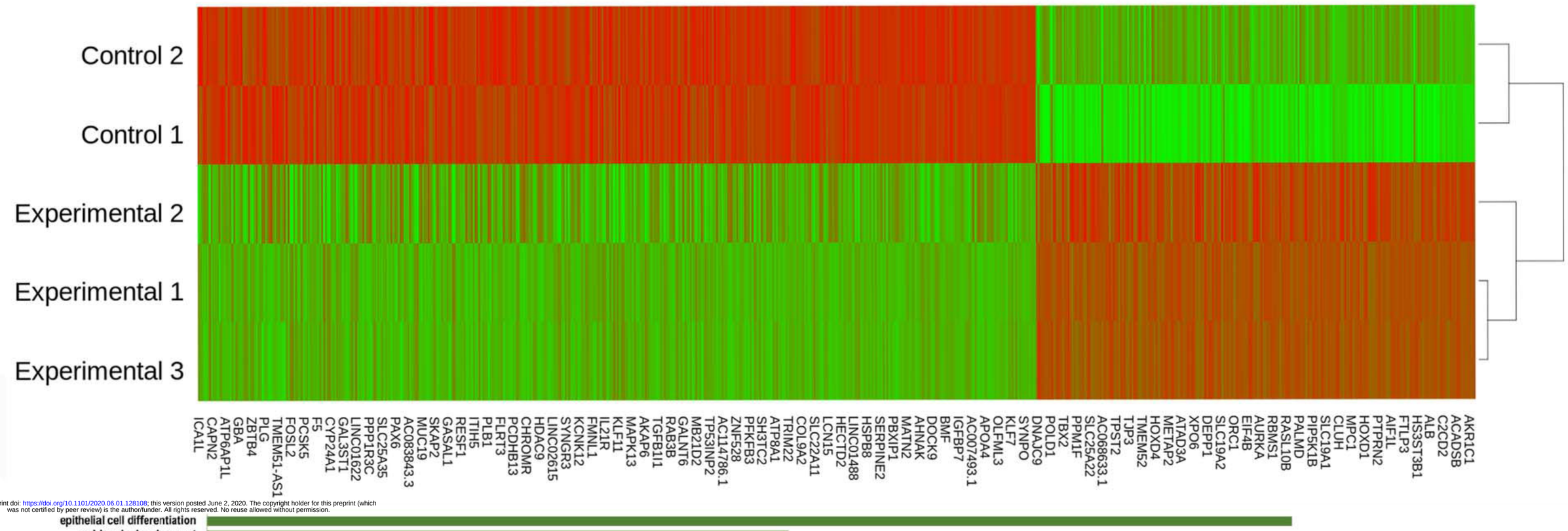
A



B

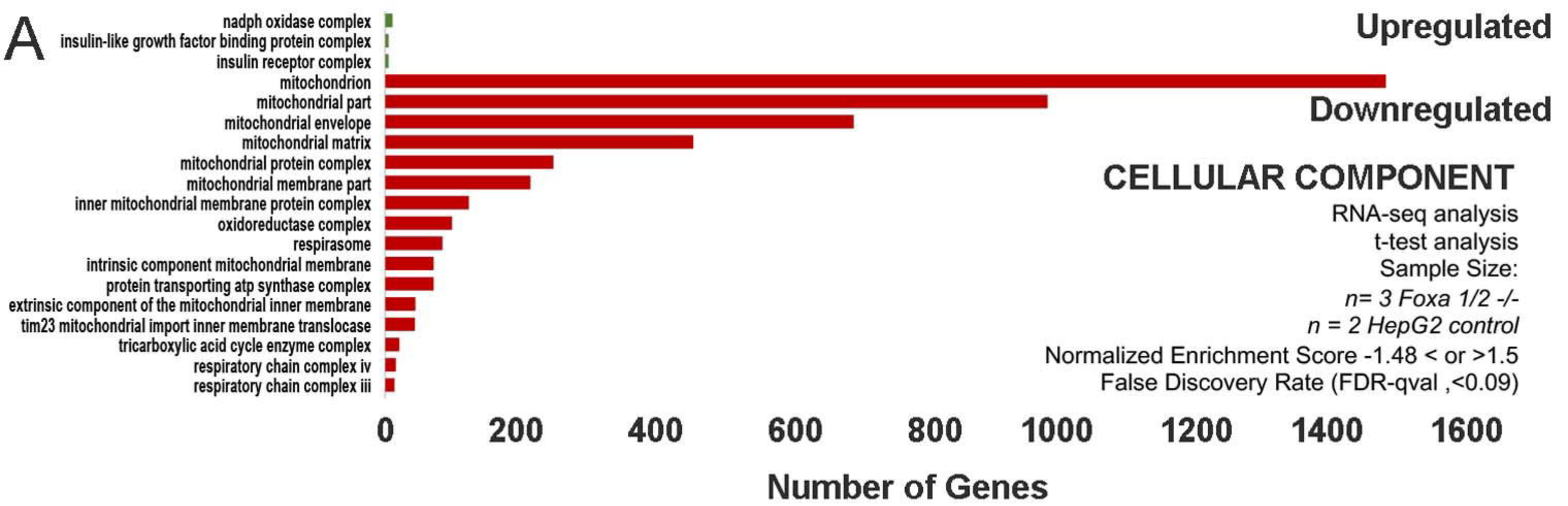


C



D

A



B