UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Role of the P1 and P2 Promoter-Driven HNF4α Isoforms in Cellular Proliferation and Differentiation in Human Colon Cancer and Mouse Embryonic Stem Cells

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Cell, Molecular, and Developmental Biology

by

Linh My Vuong

December 2014

Dissertation Committee:
Dr. Frances M Sladek, Chairperson
Dr. Nicole zur Nieden
Dr. Ernest Martinez

The Dissertation of Linh My Vuong is approved:

Committee Chairperson

University of California, Riverside

Acknowledgement

I thank Dr. Frances M. Sladek for being a great professor, research trainer, and mentor the last several years. Coming out from this research training, I have learned the necessary skills to become a critical thinker, a science presenter, and an autonomous researcher in cell and molecular biology. I will continue to improve on these skills as I move forward to the next career stage in my endeavors. I thank the people in the Sladek's lab for their help and answering all the questions I had, as well as the colloquial conversions about science and graduate student's life at UCR. I am grateful to the people at the Stem Cell Core facility who generated the iPS cells for us. I am also grateful to Dr. Joseph M. Dhahbi for taking his time and effort to analyze two RNA-Seq data and one ChIP-Seq data for us. I am in debt to my family for their persistent support and love.

Dedication to Mother and Professor

ABSTRACT OF THE DISSERTATION

The Role of the P1 and P2 Promoter-Driven HNF4α Isoforms in Cellular Proliferation and Differentiation in Human Colon Cancer and Mouse Embryonic Stem Cells

by

Linh My Vuong

Doctor of Philosophy, Graduate Program in Cell, Molecular, and Developmental Biology
University of California, Riverside, December 2014
Dr. Frances M. Sladek, Chairperson

Cellular proliferation and differentiation are critical events in normal development and cancer. Despite decades of research, much remains to be learned about how cells transition between the two states. To decipher one aspect of this switch, we focused on the transcription factor, Hepatocyte Nuclear Factor (HNF) 4α. HNF4α is a nuclear receptor that is important in development and in maintaining the homoeostasis of the adult liver and colon. There are multiple isoforms of the HNF4α that are generated by alterative promoter (P1 and P2) usage and 3' splicing events in different tissues. Both P1 and P2-HNF4α isoforms are expressed in the adult colon, while the P1-HNF4α is expressed in the adult liver. Others have shown that P1-HNF4α is downregulated in cancer, whereas P2-HNF4α is upregulated in hepatocellular carcinoma and colorectal cancer. This would suggest that P1-HNF4α is tumor suppressive, while P2-HNF4α may act as an oncogene although a mechanism has not been elucidated. One potential mechanism could be through a differential interplay with the Wnt/β-catenin/TCF pathway, which is activated in many cancers including liver and colon. To determine whether HNF4α and TCF4 cross-talk to control the switch between proliferation and differentiation and to elucidate the role of the HNF4α isoforms in early development and cancer, we generated Tet-On inducible systems that express either HNF4α2 (P1) or HNF4α8 (P2) under the control of doxycycline in human colon cancer and mouse

embryonic stem cells. We characterized the lines to look for morphological and functional differences between the isoforms. We performed RNA-Seq and ChIP-Seq on the HCT116 inducible lines to identify any changes in gene expression and regulation and compared the HNF4α2 and HNF4α8 lines to determine functional differences. Although we found some functional redundancy, there are discrete differences. HNF4α2 suppresses tumor growth, inhibits cellular proliferation, and competes with TCF4 for regulation of target genes more effectively than HNF4α8. Similarly, we also found that HNF4α2 decreases cell numbers more than HNF4α8 in the mES inducible lines. Our findings provide insight into the distinct role of the P1- and P2-HNF4α in cancer and normal development.

Table of Contents

List of Figures

List of Tables

Abbreviation

ADA2 adenosine deaminase

AF-1 Activation Function-1

AF-2 Activation Function-2

AFP alpha-fetoprotein

AOM azoxymethane

APC adenomatous polyposis coli

ApoB apolipoprotein

ApoCIII apolipoprotein C3

AR androgen receptor

ARP-1/COUP-TFII COUP transcription factor II

BCF Bovine Calf Serum

BSA bovine serum albumin

C/EBPβ CEBPB CCAAT/enhancer binding protein, beta

CAC colitis-associated colon cancer

CAG CAGG, PyF101, IRES

CAGG chicken β-actin and the rabbit β-globin promoter

CBP cAMP response element-binding protein

Cdx2 homeobox caudal

ChIP chromatin immunoprecipitation

CMV cytomegalovirus promoter

Co-IP co-immunoprecipitation

COS-7 African green monkey fibroblast like cells

CV coefficient variants

DBD DNA binding domain

DHM dimethydrazine

dn dominant negative

DOX doxycycline

d.p.c day post coitum

Dsh Dishevelled

DSS dextran sulfate sodium

DTT dithiothreitol

E embryonic day

EAR-2 V-erbA-related protein 2

EAR-3/COUP-TFI COUP transcription factor I

ES embryonic stem cells

FAP familial adenomatous polyposis

FoxA1/2 forkhead box A1/2

FBS fetal bovine serum

FC fold change

FDR false discovery rate

FPKM fragments per kilobase of exon per million fragments mapped

GATA4/6 GATA binding protein 4/6

GO Gene Ontology

GRIP-1 glucocorticoid receptor-interacting protein 1

GSK3 glycogen synthase kinase 3

GST glutathione S-transferase

HCMV human cytomegalovirus

HCT116 human colon cancer cells

HDAC histone deacetylase

HEK293 human embryonic kidney cells

HEK293T human embryonic kidney T antigen cells

hESCs human embryonic stem cells

HNF4α (NR2A1) hepatocyte nuclear factor 4 alpha

HNPCC hereditary non-poly

HRE HNF4 or a PPAR motif

HRP horseradish peroxidase

HstrgTstrg HNF4α strong TCF strong

HstrgTwk HNF4α strong TCF weak

HwkTstrg HNF4α weak TCF strong

IACUC Institutional Animal Care and Use Committee

IB immunoblot

ICM inner cell mass

IF immunoflorescence

iHep induced hepatocyte-like cells

IR (IRES) internal ribosomal entry site

iPS inducible pluripotent stem cells

kb kilobase

LA linoleic acid

LBD ligand binding domain

LEF lymphoid enhancer factor

LIF leukemia inhibitory factor

LRH-1 liver receptor homolog 1

LTFs liver-specific transcription factors

MEF mouse embryonic fibroblast

mESCs (D3) mouse embryonic stem cells D3

MIEP major immediate early promoter

MMR mismatch repair

MODY1 maturity-onset diabetes of the young one

NCOA1 nuclear receptor coactivator 1

NE nuclear extracts

NR nuclear receptor

PAX4/6 paired box 4/6

PANC-1 human pancreatic tumor cell line

PCR polymerase chain reaction

PBM protein binding microarrays

PBS phosphate buffered saline

PC4 positive coactivator 4

PL parental line

PMSF phenylmethylsulfonyl fluoride

PR progesterone receptor

PP periportal

P/S penicillin-streptomycin

PURO puromycin

PV perivenular

PWM Position weight matrices

PyF101 polyoma virus mutant enhancer

RAR retinoic acid receptor

RFP red fluoresce protein

RLSC-dH resident liver stem cell derived hepatocytes

rtTA reverse tetracycline-controlled transcriptional activator

RT-PCR real time PCR

SD standard deviation

SDS Sodium Dodecyl Sulfate

SEM standard error of the mean

Seq sequencing

SMRT silencing mediator for retinoid and thyroid receptors

SNP single nucleotide polymorphism

Tc tetracycline

TCF T cell factor proteins

TCGA The Cancer Genome Atlas

Tet tetracycline-controlled

TetO Tet operator

TetR Tet Repressor

TFs transcription factors

TRE tetracycline response element

TSS transcription start site

tTA Tet-controlled transcriptional activator

tTS tetracycline-controlledtranscriptional silencer

TTR transthyretin

VP16 activation domain of Herpes Simplex virus

WCE whole cell extracts

WRE Wnt responsive element

α2 HNF4α2 line

α8 HNF4α8 line

Chapter 1

Introduction

HNF4α and its function

Hepatocyte Nuclear Factor (HNF) 4α (NR2A1) is a liver-enriched transcription factor that regulates genes involved in embryogenesis and the maintenance of fetal and adult liver (Chen et al. 1994; Hayhurst et al. 2001). Sladek et al (1990) was the first to isolate the 54 KD HNF4α protein from rat liver nuclear extracts and found that the protein factor was able to bind to liver-specific genes such as transthyretin (TTR) and apolipoprotein CIII (ApoCIII). Classified as a member of the steroid hormone receptor superfamily of ligand-dependent transcription factor (Sladek FM et al. 1990), the amino acid structure of HNF4α consisting of a variable A/B domain that contains an activation function (AF-1) in the N-terminus; two conserved domains – the DNA binding domain (DBD) that has two zinc fingers and the hydrophobic ligand binding domain (LBD) that contains a second activation function (AF-2); and an F domain at the C-terminus that can inhibit the activity of the AF-2 function (Hadzopoulou-Cladaras et al. 1997; Mangelsdorf et al. 1995; Mangelsdorf and Evans 1995; Sladek et al. 1999). Defined as a new subclass of nuclear receptors (NRs), HNF4α is found only in the nucleus and binds exclusively as a homodimer to DNA (Jiang et al. 1995). HNF4α was categorized as an orphan nuclear receptor until the identification of its endogenous ligand - linoleic acid (LA, C18:2ω6) (Yuan et al. 2009). HNF4α binds LA in a reversible fashion making it a potential drug targets although the binding of LA does not appear to significantly influence its transcriptional function: HNF4α binds DNA and activates transcription in the absence of LA (Yuan et al. 2009), and hence remains a constitutive transactivator. The only discernable function of LA binding was an apparent decrease in protein stability (Yuan et al. 2009).

Many of the genes HNF4α regulates are involved in the transport and metabolism of nutrients, hematopoiesis, xenobiotic and drug detoxification, immune response, liver differentiation, and signal transduction, to name a few (Sladek and Seidel 2001; Odom et al. 2004; Bolotin et al. 2010). Although HNF4α was originally identified and characterized in the adult liver, other tissues also express HNF4α including the kidney, small intestine, colon, pancreas and stomach, as well as visceral endoderm of the developing embryo (Sladek et al. 1990; Sladek and Seidel 2001). Dysfunction of HNF4α, due to viral infections or inherited mutations in the HNF4α-binding sites of target genes or in the HNF4α gene itself, is linked to various human diseases such as hemophilia, atherosclerosis, diabetes, cancer, hepatitis, hypoxia anemia, and others (Sladek and Seidel 2001; Yusuf et al. 2012). This ancient and highly conserved nuclear receptor that is found in sponge to human (Sladek 2011) is part of a regulatory transcriptional network that is involved in the development and maintenance of the liver, including the gastrointestinal tract (Kyrmizi et al. 2006). Other factors involved in this circuitry complex include HNF1α/β, HNF6, GATA4/6, and HNF3 (Odom et al. 2004; Kyrmizi et al. 2006).

HNF4α promoter-driven isoforms

There are two HNF4 genes found in human and mouse: *HNF4A,* which encodes the protein originally identified as HNF4, and *HNF4G* (HNF4γ) (Sladek 2011). A third gene, *HNF4B* (HNF4β) has been found only in Xenopus (Holewa et al. 1997). The first is HNF4α, as mentioned above and will be emphasized throughout the discussion; it is evolutionarily conserved across the animal organisms (Sladek and Seidel 2001). The other is HNF4γ that is similar, but not identical, to HNF4α (Drewes et al. 1996;

Plengvidhya et al. 1999; Sladek and Seidel 2001). There is an HNF4β that is found only in Xenopus (Holewa et al. 1997).

There are multiple transcript variants (isoforms) of HNF4α in human, mouse, and rat (Hata S et al. 1992, 1995; Drewes T et al. 1996) (Fig. 1.1). Thus far, a total of twelve isoforms (HNF4α1-12) have been identified although it is not certain if all are expressed (Kritis et al. 1996; Nakhei H et al. 1998; Torres-Padilla et al. 2001; Eeckhoute et al. 2003; Ihara et al. 2005; Harries et al. 2008; Huang et al. 2008, 2009; Yusuf et al. 2012). The isoforms are generated through the utilization of two alternative promoters (proximal P1 and distal P2) that are forty kilobases (kb) apart in human and mouse and several 3' splicing events in the *HNF4A* gene that spans 30 kb and consists of at least twelve exons (Taraviras et al. 1994; Furuta et al. 1997). Variants HNF4α1-6 are expressed by the P1 promoter and include the exon 1A, whereas splice variants from HNF4α7-12 are expressed from the distal P2 promoter and contain exon 1D but not exon 1A. Splice variants within the P1 and P2 groups also have varying C-termini due to several splicing events: (1) an alternative splicing event in exon 9 that includes (or not) 10 amino acids and; (2) a terminal polyadenylation signal in the intron between exons 8 and 9 that leads to a truncated protein lacking the F domain (Huang et al. 2009) (Fig. 1.1).

P1- and P2-HNF4α isoforms differ only in the N-terminal A/B domain. They have distinct transactivation properties, as this region, the A/B, is important for interaction with coregulators (Torres-Padilla et al. 2002; Torres-Padilla and Weiss 2003). Others have shown that the role of the AF-1 module of NRs is to recruit coregulators; the AF-1 of P1-HNF4α has been shown to interact with the general transcription factors, cAMP response element-binding protein (CBP), glucocorticoid receptor-interacting protein 1 (GRIP-1), adenosine deaminase (ADA2), positive coactivator 4 (PC4), and many more

to mediate transcriptional regulations of target genes (Green et al. 1998; Hittelman et al. 1999; Ma et al. 1999; Yusuf et al. 2012) (see http://cisreg.cmmt.ubc.ca/cgi-bin/tfe/articles.pl?tfid=140&tab=interactions for additional list of other regulators that have been shown to interact with HNF4α) (Fig. 1.2).

Aside from the AF-1 module, an additional activation function, the AF-2 module in helix 12, is present in the C-terminus half of the protein in the LBD domain in both the P1- and P2-HNF4α isoforms. AF-2 recruits coactivators (e.g. GRIP-1 & p300) as well as corepressors (e.g. silencing mediator for retinoid and thyroid receptors (SMRT)) to regulate transcription (Torres-Padilla et al. 2002; Ruse et al. 2002; Maeda et al. 2002) (see http://cisreg.cmmt.ubc.ca/cgi-bin/tfe/articles.pl?tfid=140&tab=interactions for other coregulators). As a unique feature, HNF4α contains an F region that negatively regulates the AF-2 activity (Hadzopoulous-Cladaras et al. 1997). Interestingly, this inhibition is diminished by a 10 amino acid-insert in the center of the F domain that is found in certain splice variants of HNF4α (Sladek et al. 1999) (Fig. 1.2).

HNF4α as a tumor suppressor in hepatocellular carcinoma (HCC)

HNF4α is considered to be a master regulator of the liver-specific gene expression (Kaestner comment in Biolotin et al. 2010). It is the central mediator of gene expression and functions in the fetal and adult livers (Watt 2003). Aforementioned, knockout studies of *Hnf4* in the fetus and adult showed consequential effects (Duncan et al. 1997; Hayhurst et al. 2001; Watt 2003). In addition, more recent studies suggest that HNF4α might play a role in hepatocarcinogenesis (Tanaka et al. 2006; Lazarevich et al. 2008; Yin et al. 2008). Initially, it was shown that ectopic expression of HNF4α in dedifferentiated rat hepatoma variant H5 line was able to drive hepatic differentiation in

these cells by turning on epithelial morphogenic genes such as E-cadherin and intermediate filament cytokeratin proteins (Späth and Weiss 1997, 1998). Since then additional reports indicated that HNF4α has a more active and suppressive role in liver cancer (Lazarevich et al. 2004; Ning et al. 2010; Hatziapostolou et al. 2011; Bonzo et al. 2012; Walesky et al. 2013), None of these works, however, distinguished the two different promoter-driven HNF4α isoforms. Only Tanaka et al showed an increase in P2 (and concomitant decrease in P1) in human HCC (Tanaka et al. 2006). And there is one suggestion that it is HNF4α1, the P1-HNF4α isoform, that is a tumor suppressive but no data are presented (Lazarevich et al. 2004). Therefore, it is not clear, however, if the P2-HNF4α isoforms have a similar tumor suppressive role.

P1- and P2-HNF4α differential functions and ambiguous role in Colorectal Cancer (CRC)

Most studies on HNF4α focus on the P1-HNF4α isoforms; there is little emphasis on the P2-HNF4α isoforms. The few studies that have been done, however, showed some functional differences. For example, the Weiss groups at the Pasteur Institute in France, developed a knock-in mouse line in which only one isoform, either HNF4α1 or HNF4α7, is expressed under the control of both promoters and investigated the differential functions between the promoter-driven P1- and P2-HNF4α isoforms. They showed that these 'α7-only' and 'α1-only' mice are functionally redundant in development. However, they did note a discrepancy in the differential expression of some genes in the two lines; these genes are therefore dependent on the presence or absence of the AF-1 domain that recruits different coregulators (Briancon and Weiss 2006; Torres-Padilla et al. 2002; Eeckhoute et al. 2003). Another group compared the

HNF4α2 and HNF4α8 in rat β-cell line INS-1 and showed that they essentially regulate the same set of genes although more genes were affected by (P1) HNF4α2 and it is also a stronger transactivator than (P2) HNF4α8 (Erdmann et al. 2007).

Although P1- and P2-HNF4α isoforms are functionally redundant in normal development, there are a few reports suggesting that they may play differential roles in cancer development, especially in colorectal cancer (CRC). The P1-HNF4α isoforms, specifically HNF4α1, have been shown to inhibit proliferation and to interact with cell cycle players, most notably: c-Myc and cyclin D1; both have been shown to inhibit HNF4α activity (Hwang-Versules and Sladek 2008; Hanse et al. 2012). In contrast, the role of the P2-HNF4α in proliferation has not been extensively investigated, although several reports suggest that the P2-HNF4α may act as an oncogene in cancer. Immunohistochemical staining for P1- and P2-HNF4α in human cancer tissues of the colon, liver, stomach, and pancreas showed that P1-HNF4α is lost in the majority of the samples while the P2-HNF4α expression persists (Tanaka et al. 2006; Oshima et al. 2007; Takano et al. 2009). The Sladek lab showed a correlation between the loss of P1-HNF4α and the expression of the active Src in 80% of tumors from colon cancer patients. Mechanistically, the authors showed that Src kinase phosphorylates P1-HNF4α in a three-step mechanism that leads to protein instability and cytoplasmic localization, whereas the P2-HNF4α protein was not disturbed (Challeppa et al. 2012). At the same time, The Cancer Genome Atlas (TCGA) project identified a region surrounding the *HNF4A* gene in chromosome 20q13.12 is amplified in over 255 human colon cancers (The Cancer Genome Atlas Network 2012). Some of those tissues also showed HNF4α overexpression of HNF4α protein (Zhang et al. 2014). While these studies did not distinguished between the HNF4α isoforms, they could be explained by a persistent

expression of P2-HNF4α and loss of P1-HNF4α as others have noted (Tanaka et al. 2006; Takano et al. 2009; Oshima et al. 2007; Challeppa et al. 2012).

Another key contributing factor to colorectal cancer is the Wnt/β-catenin signaling pathway. This pathway is necessary for the self-renewal and preservation of adult stem and progenitor cells at the base of the crypt in the colon (Korinek et al. 1998). Upon injury or epithelial turnover, as cells at the top of the crypt undergo apoptosis, stem cells at the bottom of the crypt will differentiate and migrate up the crypt to replace damaged or older cells every 3-5 days (Barker 2014; Creamer et al. 1961) (Fig. 1.3). The Wnt pathway is a transduction cascade where, upon Wnt ligand stimulation of the cell, the Frizzled receptor recruits cytoplasmic Dishevelled (Dsh), which brings with it Axin, leading to the disassembling of the Axin/GSK3/APC complex and stabilization and nuclear translocation of β-catenin (Bienz and Clevers 2000). After translocation into the nucleus, β-catenin binds the transcription factor TCF4, one of four TCF/LEF family members (Fig. 1.4) and the predominant TCF family member expressed in intestinal epithelium; β-catenin/TCF4 stimulates transcription of Wnt target genes such as c-Myc in order to maintain the proliferative state of the progenitors in the crypts (Korinek et al. 1997, 1998) (Fig. 1.5). Disruption of the Axin/GSK3/APC complex results in uncontrolled activity of the β-catenin/TCF4 complex, and, consequentially, colon cancer (Munemitsu et al. 1995; Morin et al. 1997; Satoh et al. 2000; Liu et al. 2000; Bienz and Clevers 2000). The overactive Wnt/β-catenin signaling and the loss of HNF4α in colorectal cancer leads one to wonder whether HNF4α and the Wnt/β-catenin pathway could be linked in some way.

Cross-regulation between nuclear receptors and the Wnt/β-catenin pathway

Wnt signaling and nuclear receptors (NRs) are two very ancient systems (Teo et al. 2006; Cadigan and Waterman 2012; Sladek 2011) that play important roles in both cancer and development (Willert and Jones 2006; Nikolenko and Krasnov et al. 2007; Chung and Cooney 2003). Links between NRs and the Wnt/β-catenin cascade has been known for some time (Mulholland et al. 2005; Beildeck et al. 2010) but a full understanding of how the many diverse NRs interact with various components of the Wnt/β-catenin signaling is still lacking. Mulholland et al reviewed the work done by others on the potential crosstalk between NRs and the Wnt/β-catenin pathway. In their review, they mentioned that the Wnt/β-catenin signaling and NRs regulate each other with both repressive and stimulatory effects. For example, within the last decade, two groups reported a reciprocal crosstalk between Wnt/β-catenin signaling and the androgen receptor (AR). One group showed that the interaction between a mutated form of β-catenin and AR leads to androgen-dependent transcriptional activation in prostate cancer (Truica et al. 2000). A second group showed that AR represses β-catenin/TCF-related transcription in an androgen-dependent fashion (Schweizer et al. 2008). AR is one of the many NRs that crosstalk with the Wnt/β-catenin signaling. Another well-known NR is the retinoic acid receptor (RAR), which competes with TCF for binding to β-catenin and activates transcription on the RAR responsive elements (Easwaran et al. 1999). Interactions between other members of the NR superfamily and the Wnt/β-catenin/TCF pathway raise the issues of potential interactions also with HNF4.

HNF4α and the Wnt/β-catenin signaling pathway

Reports of links between HNF4α and the Wnt/β-catenin pathway start to make their appearance after the millennium. Several papers showed physical interactions between HNF4α and TCF4 or LEF1 in intestinal epithelium and hepatocytes (Benahmed et al. 2008; Cattin et al. 2009; Yang et al. 2013; Gougelet et al. 2013; Colletti et al. 2009). Benahmed et al (2008) showed that homeobox caudal 2 (*Cdx2*) gene expression is activated by a combination of HNF4, GATA6, β-catenin, and TCF4 in the adult intestine and identified putative binding sties for TCF4, HNF4, and GATA factors in a short 200 bp region – 8500 bp upstream of the transcription start site (TSS). Another group found that the Wnt signaling converges on the HNF4α-driven transcription for control of liver zonation (Colletti et al. 2009). In an immunoprecipiation assay, Colletti et al showed that HNF4α and LEF1 interact in resident liver stem cell derived hepatocytes (RLSC-dH). Colletti and co-authors also performed chromatin immunoprecipitation (ChIP) followed by polymerase chain reaction (PCR) on individual genes regulating liver zonation for HNF4α and LEF1 in RLSC-dH and found that HNF4α binds its own consensus site to activate genes in the periportal (PP) region and repress genes in the perivenular (PV) region. In contrast, when β-catenin is active, LEF1 binds its own and/or HNF4α consensus site, displacing HNF4α from its site, to active genes in the PV region and repress genes in the PP region (Colletti et al. 2009). At the same time, Cattin et al. demonstrated that HNF4α interferes with Wnt-TCF4/β-catenin transcriptional activities to modulate the balance between proliferation and differentiation in the intestinal epithelium. This suggested that loss of HNF4α or stimulation of Wnt/β-catenin activity could lead to an unbalance of that modulation and result in tumor formation (Cattin et al. 2009). In liver cells, Frietze et al found that HNF4α and TCF4 (*TCF7L2*) share similar

binding sites in a ChIP-Seq experiment in HepG2 cells, a well-differentiated hepatocellular carcinoma cell line (Frietze et al. 2012). Using a knockout model and genome-wide studies, another group unveiled an interaction between HNF4α and the Wnt/β-catenin/TCF4 pathway that patterns liver zonation (Gougelet et al. 2014). They showed that TCF4 could shift between two partners – HNF4α and β-catenin. In the PV compartment, where β-catenin is present, they found that TCF4 partners with β-catenin and binds Wnt responsive element (WRE). In contrast, in the PP zone, there is no β-catenin activity; hence, TCF4 is recruited to an HNF4 or a PPAR motif (HRE) (Gougelet et al. 2013). While the evidence is mounting that HNF4α, like other NRs, interacts with the Wnt/β-catenin/TCF pathway in an important and functional manner, none of these studies to date examined the effect of the different HNF4α isoforms on those interactions. All these studies focused on P1-HNF4α. Yang et al, however, did map the interaction and saw that the β-catenin-binding-domain of TCF4 is essential for interaction with HNF4α, while the DNA-binding-domain-containing N-terminal region of HNF4α directly interacts with TCF4 (Yang et al. 2013). P2-HNF4α, however, has not been examined.

Promoter-driven P1- and P2-HNF4α isoforms in development

HNF4α is one of the many key components necessary for early development. Knockout of *Hnf4a* in mouse is an embryonic lethal at day 9.5 (E9.5) (Chen et al. 1994). This is due to improper differentiation of the visceral endoderm and massive cell death that disrupts the gastrulation three days earlier at E6.5. Using in-situ hybridization, Duncan et al (1994) tracked the expression of HNF4α in early murine development. At the blastocyst stage (E4.5), HNF4α is found expressed in the primitive endoderm that

surrounds the inner cell mass (ICM). The expression persists as the cells differentiate to become the parietal endoderm and the visceral endoderm of the yolk sac (E8.0). By E9.5, low to moderate levels of HNF4α is seen in the hindgut and midgut, while a high level of expression is detected in the developing liver (Duncan et al. 1994).

Little is known about the spatial and temporal expression of the promoter-driven HNF4α isoforms in early development. However, others have used RT-PCR to show that P2-HNF4α variants are present in the fetal liver at E12.5 and peak at E17.5. With P2-HNF4α expression decreases to a lower level at birth and essentially becomes absent in the adult liver, P1-HNF4α variants expression increases significantly (Briancon et al. 2004). Not surprisingly, the temporal expression of the P1- and P2-HNF4α isoforms coincides with their transcriptional activities during development. The P2-HNF4α isoforms activate promoters of early genes in fetal development more effectively than the P1-HNF4α isoforms; likewise, the latter is an effective transactivator of promoters of genes expressed postnatally and in adult livers (Torres-Padilla et al. 2001).

The question remains, however, as to which promoter-driven HNF4α variant is present at the early stages of development, such as in the primary endoderm at E4.5 when HNF4α is first detected (Duncan et al. 1994). The primitive endoderm bifurcates into two distant lineages: parietal and visceral endoderms. By E5.5, the visceral endoderm forms a protective and selectively permeable barrier that allows for exchange of materials between the embryo and the maternal surroundings (Parr and Parr 1986; Jollie 1990). The visceral endoderm then becomes the yolk sac, which aids in the development of the embryo (Beckman et al. 1990). Its functions have been shown to be analogous to the adult liver, many of which include secretion, transcytosis, and digestion

(Jollie 1990; Duncan et al. 1994; Chen et al. 1994). With that said, however, the HNF4α isoforms have not been distinguished.

Understanding early development using embryonic and inducible pluripotent stem cells

The discovery of embryonic stem (ES) cells derived from early mouse and human embryos at the blastocyst stage made it possible to recapitulate and study early mammalian development ex vivo (Martin 1981; Thomson et al. 1998). These are pluripotent stem cells that can differentiate into any lineage of the embryonic body and have an unlimited proliferative capacity in vitro (Boiani and Scholer 2005; Suda et al. 1987). The pluripotent state of the cells is maintained through an integrated network of external signals, which include the canonical Wnt/β-catenin and cytokine leukemia inhibitory factor (LIF) pathways (Sato et al. 2003; Williams et al. 1988), and internal pluripotent factors, which involves Oct4, Sox2, Klf4, and c-Myc (Chen et al. 2008). Just a year ago, Rais et al found that these pluripotent factors (Oct4, Sox2, Klf4, c-Myc) may actually be lineage-specific factors in that they can drive pluripotent stem cells to differentiate into a particular lineage if one of the factors is not kept in check by the other (Rais et al. 2013). This finding goes against the very definition of pluripotency markers (Loh et al. 2013). Interestingly, these factors are the same ones of genes that can reprogram cells into inducible pluripotent, undifferentiated stem (iPS) cells when introduced into somatic, fully differentiated, cells (Takahashi and Yamanaka 2006). Other groups are also looking for other lineage-specific factors that potentially can reprogram fibroblasts into a pluripotent state (Shu et al. 2013; Montserrat et al. 2013).

The ability to differentiate or transform one type of cells into another type made it possible to study and understand factors or events that are involved in early development. In the last ten years, various hepatocyte differentiation protocols have been published that differentiate ES and iPS cells into fully mature hepatocytes with various cocktails of growth factors (Fair et al. 2003; Soto-Gutiérrez et al. 2007; Song et al. 2009; Hannan et al. 2013; Si-Tayeb et al. 2010; Mallanna and Duncan 2013). These hepatocytes could potentially be used to replace injured or damaged liver tissues and screen for better drugs to treat liver diseases (Soto-Gutiérrez et al. 2008).

More recently, two groups have developed methods that completely bypass ES to generate hepatocytes. Huang et al and Sekiya and Suzuki used liver-specific transcription factors (LTFs) to transdifferentiate mouse fibroblasts into functional hepatocyte-like cells (Huang et al. 2011; Sekiya and Suzuki 2011). Finally, two years later, Yamamizu et al showed that overexpression of HNF4α can alone direct differentiation into hepatocytes (Yamamizu et al. 2013), validating its essential role in differentiation and development (DeLaForest et al. 2011; Battle et al. 2006). With each improvement of these methods to achieve completely functional hepatocytes, we can begin to study HNF4α in early development, particularly the differential role of the P1 and P2-HNF4α.

Both the P1- and P2-HNF4α are expressed in mouse liver at 14.5 day post coitum (d.p.c) (Torres-Padilla et al. 2001). When signals from the surrounding mesenchyme induce the cells to proliferate, expand, and differentiate into the liver and ventral pancreas: only P1-HNF4α is expressed in the liver while P2-HNF4α is expressed in the pancreas (Gilbert 2000; Huang et al. 2008; Deutsch et al. 2001; Hansen et al. 2002; Eeckhoute et al. 2003; Briancon and Weiss 2006).

Aside from its role in the adult liver, HNF4α also plays a key role in β-cell function in the pancreas (Wang et al. 200; Miura et al. 2006), regulating about 11% of islet genes (Odom et al. 2004). Absence (knockout) or loss of function (heterozygous mutations) of the *HNF4A* leads to the impairment of the pancreatic β-cells to secrete glucose in response to insulin (Miura et al. 2006; Yamagata et al. 1996). The disease maturity-onset diabetes of the young one (MODY1) is the result of heterozygous mutation in the human *HNF4A* gene. This autosomal-dominant inheritance disease causes transient hyperinsulinemic hypoglycemia and macrosomia in infants (Stanescu et al. 2012; Pearson 2007; Kapoor et al. 2008). Many mutations however mapped in the HNF4α (HNF4A) coding region as well as the P2 but not P1 promoter.

It has been known for some time that pancreatic and hepatic cells can transdifferentiate into each other, and that the transcription factor C/EBPβ may be the switch between the two organs (Rao et al. 1995; Shen et al. 2000; Horb et al. 2003; Yi et al. 2012). However, little work has been done to understand how one of the promoter-driven HNF4α isoforms is lost during the differentiation of hepatic and pancreatic lineages from the hepatic diverticulum. With the advent of the ES and iPS research, dissecting out this mechanism may be possible.

Objectives of this dissertation

The overall goal of this dissertation is to understand how cells transition from a proliferating to a differentiating state in early embryogenesis, and then back again during cancer. What is the switch that controls proliferation and differentiation? To phrase it in another way, the more immediate goal of the project is to look at one potential mechanism that, hopefully, will shed light on this phenomenon that is the basis of all

multicellular organisms. The focus is on HNF4α and its role in cellular proliferation, looking specifically at the differential role between the P1- and P2-HNF4α promoter-driven isoforms in regulating proliferation and differentiation. Although others including our lab have reported that HNF4α inhibits proliferation, and some have worked out the potential mechanisms, as highlighted above, much remains to be learned. Also, little is known about the role of HNF4α, particularly the P2-HNF4α isoforms, in colorectal cancer and early development.

Both HNF4α and the Wnt/β-catenin/TCF4 pathway are known to be important in maintaining the epithelial and stem cell compartment in the colon; an imbalance in any one of those factors could lead to the loss of the epithelium or stem cells. We hypothesize that the interplay between HNF4α and TCF4 could be one switch that controls proliferation and differentiation in the colon.

To address these issues, we generated Tet-On inducible lines that express either the (P1) HNF4α2 or (P2) HNF4α8 under the control of doxycycline (DOX) in both a human colorectal cancer cell line, HCT116, and mouse embryonic stem (mES) cells. The generation of the clones is described in Chapter 2. Using high-throughput next generation sequencing (RNA-Seq and ChIP-Seq), in Chapter 3, we identify functional differences between HNF4α2 and HNF4α8 in the inducible HCT116 cells. In addition, we investigate a potential competition as well as co-regulation between HNF4α and TCF4 on certain target genes. Finally, in Chapter 4, we investigate the role of HNF4α in stem cell differentiation, asking how might HNF4α direct the differentiation of mouse embryonic stem cells, what type of cells are induced, and are there differences between HNF4α2 and HNF4α8? Could these differences help dissect out a mechanism that explains the loss of one promoter-driven isoform and the presence of the other in the

liver and pancreas, two lineages that come from the same precursor? At the end, in Chapter 5, I summarize the two projects and ask whether the results address our initial hypothesis and goals. I also touch on future directions for this on going project.

Fig. 1.1

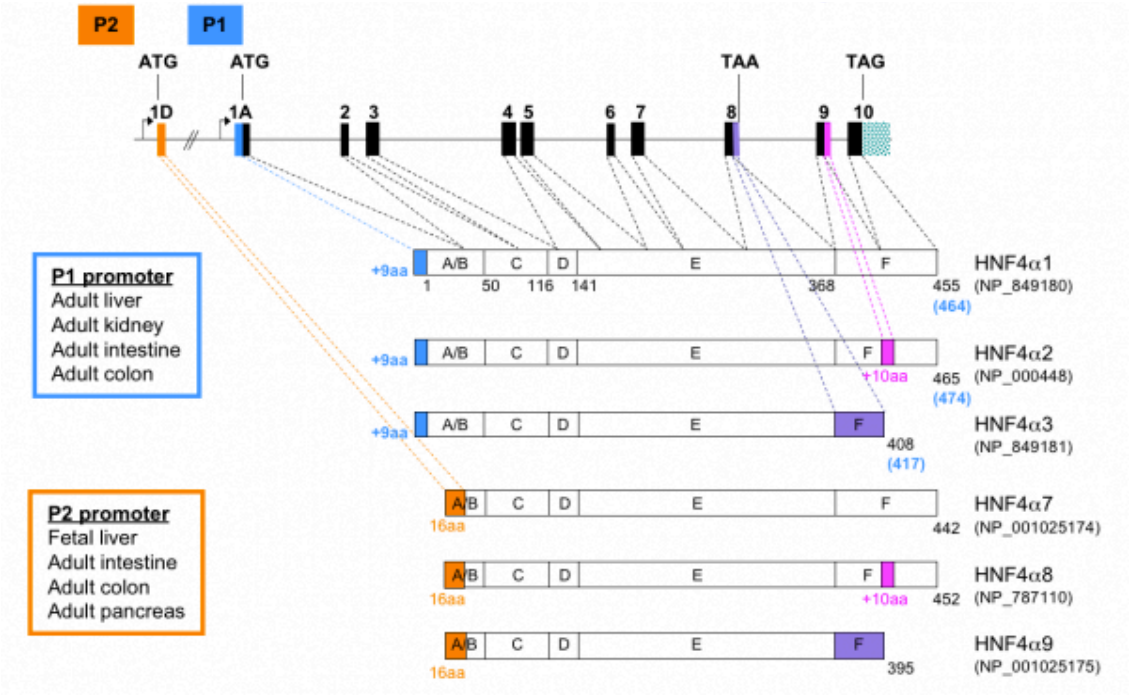


Figure 1.1 Schematic of the HNF4A gene structure.

The gene spans ~74 kb long with several splicing exons. The P1 promoter drives the expression of the HNF4α isoforms that contain the full length A/B domain while the P2 promoter drives the expression of the HNF4α isoforms that contains the truncated A/B domain. The box to the right shows the different tissues where P1 and P2 promoters are expressed.

(Bolotin, Schnable, Sladek: HNF4a. *In Yusuf et al.:* **Transcription Factor Encyclopedia,** *Genome Biology 2012*, 13:R24. http://cisreg.cmmt.ubc.ca/cgi-bin/tfe/articles.pl?tfid=140)

Fig. 1.2

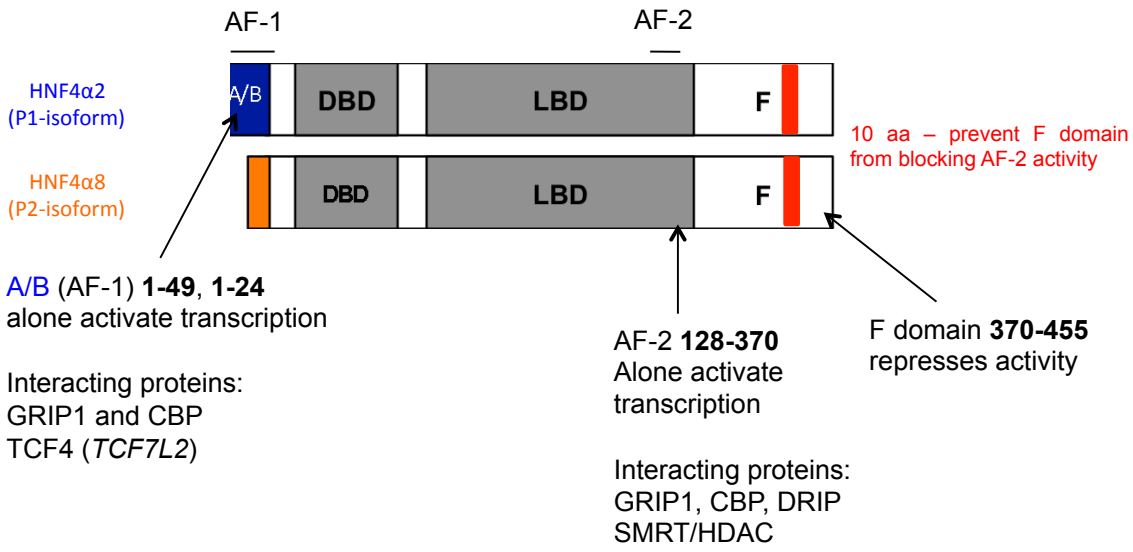


Figure 1.2 HNF4α structural domains and their functions.

AF-1 and AF-2 interact with co-activators (GRIP1, CBP, and DRIP) to activate transcription; AF-2 can also interact with co-repressor (SMRT/HDAC3) to repress transcription of target genes (Sladek et al. 1999; Torres-Padilla et al. 2002; Hadzopoulou-Cladaras et al. 1997; Ruse et al. 2002; Maeda et al. 2002) (http://cisreg.cmmt.ubc.ca/cgi-bin/tfe/articles.pl?tfid=140&tab=interactions list additional co-regulators, Yusuf et al. 2012). The F domain inhibits the activity of AF-2, and the addition of 10 amino acids (aa) in the F domain mitigates this inhibitory effect (Sladek et al. 1999; Suaud et al. 1999). TCF4 (TCF7L2) has been shown to interact with the DNA-binding-domain-containing N-terminal region of HNF4α (Yang et al. 2013).

Fig. 1.3

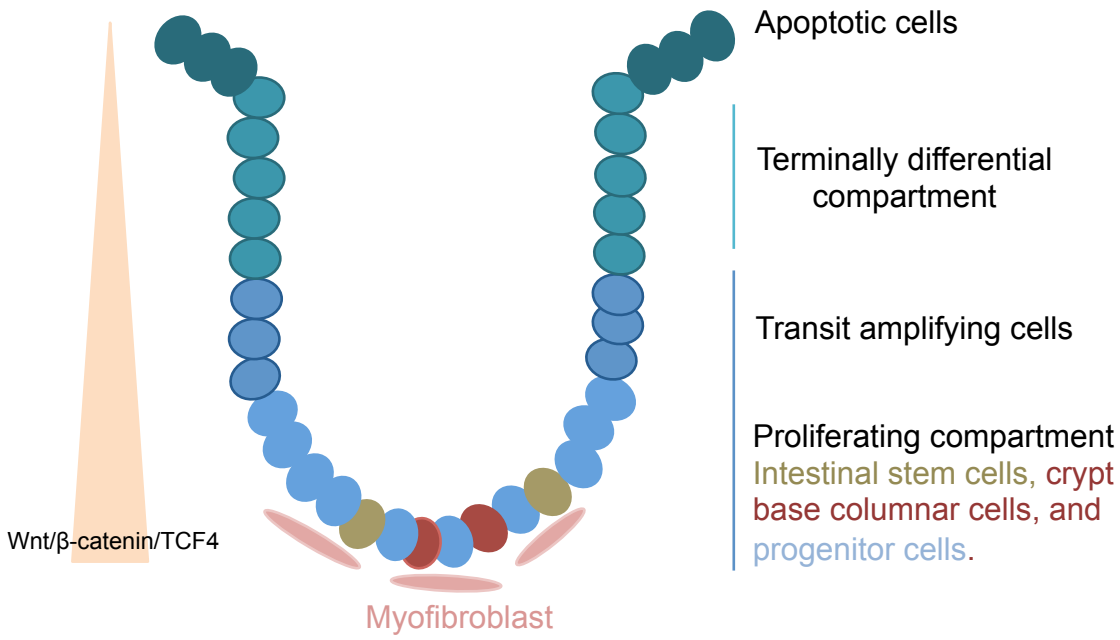


Figure 1.3 Basic structure of the colonic crypt.

Cells along the crypt progress through the different stages of proliferation and differentiation as they migrate up and eventually fall off the crypt structure as apoptotic cells. The bottom consists of stem cell niche, which include intestinal stem cells, crypt base columnar cells, progenitor cells, and mesenchymal cells of myofibroblast lineage (Medema and Vermeulen 2011; Humphries and Wright 2008). The expression and activity of the Wnt/β-catenin/TCF4 are highest at the bottom and decrease as the cells move up. Redrawn from van de Wetering et al. 2002 and Barker 2014.

Fig. 1.4

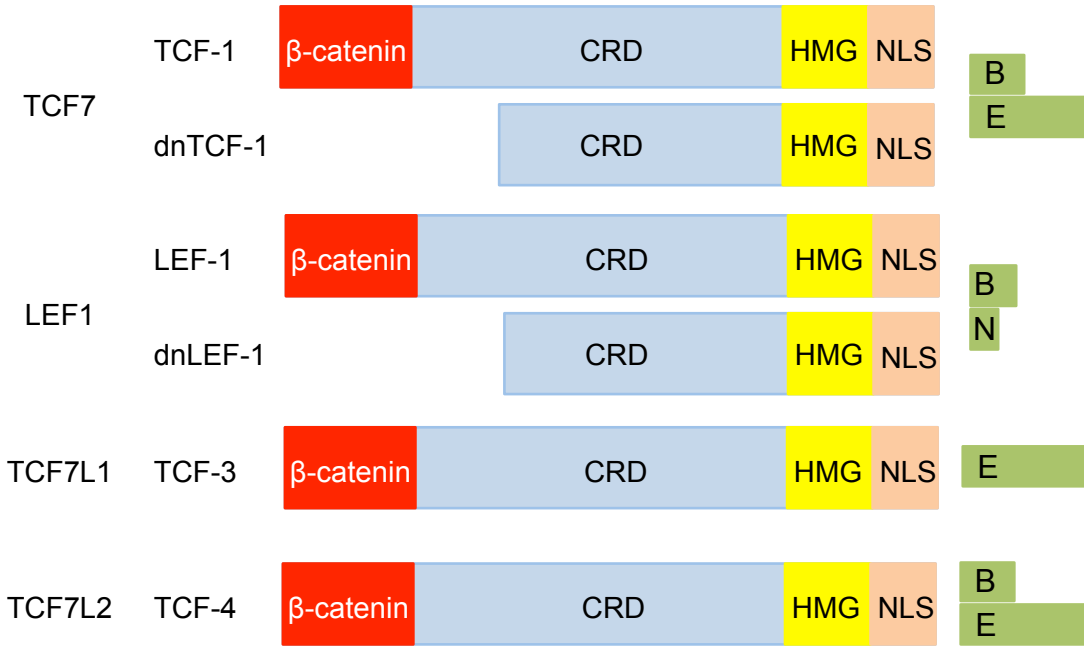


Figure 1.4 Schematic of the TCF/LEF gene structure.

LEF and TCFs are driven by alternative promoters and several splicing events. The dominant negative (dn) of TCF1 and LEF1 is driven by alternative promoters, while the context-dependent regulatory domain (CRD) and the C-termini (N, B, and E tails) are produced by alternative splicing. Upon Wnt signaling, the β-catenin-domain of TCF/LEFs interacts with β-catenin in the nucleus while the high-motility group (HMG) binds to sequence specific DNA motif to regulate Wnt target genes. The gene names are on the left and protein names are next to it. NLS – Nuclear localization signal. Redrawn from Waterman 2004 and Arce et al. 2006.

Fig. 1.5

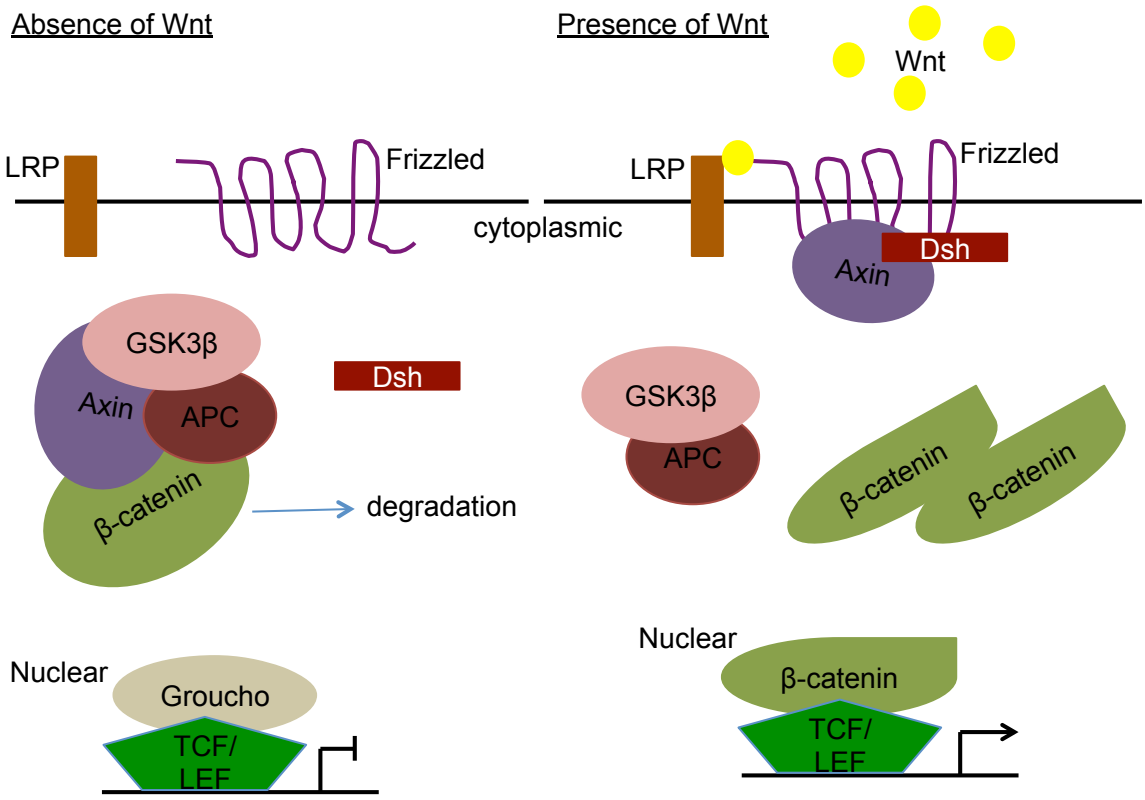


Figure 1.5 Schematic of the Wnt/β-catenin Signaling Pathway.

In the absence of Wnt, the Axin complex targets β-catenin for degradation. However, in the presence of a Wnt ligand, the ligand binds to the Frizzeled receptor; this then leads to a downstream signaling event that degrades the Axin complex. As a result, β-catenin is stabilized and gets translocated into the nucleus where it interacts with TCF/LEF to activate transcription of Wnt target genes. Redrawn from Bienz and Clevers 2000 and Clevers & Nusse 2012.

References

Arce L, Yokoyama NN, Waterman ML. 2006. Diversity of LEF/TCF action in development and disease. *Oncogene* **25**: 7492-7504.

Barker N. 2014. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature reviews Molecular cell biology* **15**: 19-33.

Battle MA, Konopka G, Parviz F, Gaggl AL, Yang C, Sladek FM, Duncan SA. 2006. Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 8419-8424.

Beckman DA, Koszalka TR, Jensen M, Brent RL. 1990. Experimental manipulation of the rodent visceral yolk sac. *Teratology* **41**: 395-404.

Beildeck ME, Gelmann EP, Byers SW. 2010. Cross-regulation of signaling pathways: an example of nuclear hormone receptors and the canonical Wnt pathway. *Experimental cell research* **316**: 1763-1772.

Benahmed F, Gross I, Gaunt SJ, Beck F, Jehan F, Domon-Dell C, Martin E, Kedinger M, Freund JN, Duluc I. 2008. Multiple regulatory regions control the complex expression pattern of the mouse Cdx2 homeobox gene. *Gastroenterology* **135**: 1238-1247, 1247 e1231-1233.

Bienz M, Clevers H. 2000. Linking colorectal cancer to Wnt signaling. *Cell* **103**: 311-320.

Boiani M, Scholer HR. 2005. Regulatory networks in embryo-derived pluripotent stem cells. *Nature reviews Molecular cell biology* **6**: 872-884.

Bonzo JA, Ferry CH, Matsubara T, Kim JH, Gonzalez FJ. 2012. Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4alpha in adult mice. *The Journal of biological chemistry* **287**: 7345-7356.

Briancon N, Bailly A, Clotman F, Jacquemin P, Lemaigre FP, Weiss MC. 2004. Expression of the alpha7 isoform of hepatocyte nuclear factor (HNF) 4 is activated by HNF6/OC-2 and HNF1 and repressed by HNF4alpha1 in the liver. *The Journal of biological chemistry* **279**: 33398-33408.

Briancon N, Weiss MC. 2006. In vivo role of the HNF4alpha AF-1 activation domain revealed by exon swapping. *The EMBO journal* **25**: 1253-1262.

Cadigan KM, Waterman ML. 2012. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harbor perspectives in biology* **4**.

Cancer Genome Atlas N. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**: 330-337.

Cattin AL, Le Beyec J, Barreau F, Saint-Just S, Houllier A, Gonzalez FJ, Robine S, Pincon-Raymond M, Cardot P, Lacasa M et al. 2009. Hepatocyte nuclear factor 4alpha, a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. *Molecular and cellular biology* **29**: 6294-6308.

Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CL, Chan C, Dent OF, Clarke SJ, Robertson GR et al. 2012. Src tyrosine kinase phosphorylation of nuclear receptor HNF4alpha correlates with isoform-specific loss of HNF4alpha in human colon cancer. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 2302-2307.

Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, Darnell JE, Jr. 1994. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes & development* **8**: 2466-2477.

Chen X, Vega VB, Ng HH. 2008. Transcriptional regulatory networks in embryonic stem cells. *Cold Spring Harbor symposia on quantitative biology* **73**: 203-209.

Chung AC, Cooney AJ. 2003. The varied roles of nuclear receptors during vertebrate embryonic development. *Nuclear receptor signaling* **1**: e007.

Clevers H, Nusse R. 2012. Wnt/beta-catenin signaling and disease. *Cell* **149**: 1192-1205.

Colletti M, Cicchini C, Conigliaro A, Santangelo L, Alonzi T, Pasquini E, Tripodi M, Amicone L. 2009. Convergence of Wnt signaling on the HNF4alpha-driven transcription in controlling liver zonation. *Gastroenterology* **137**: 660-672.

Creamer B, Shorter RG, Bamforth J. 1961. The turnover and shedding of epithelial cells. I. The turnover in the gastro-intestinal tract. *Gut* **2**: 110-118.

DeLaForest A, Nagaoka M, Si-Tayeb K, Noto FK, Konopka G, Battle MA, Duncan SA. 2011. HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. *Development* **138**: 4143-4153.

Deutsch G, Jung J, Zheng M, Lora J, Zaret KS. 2001. A bipotential precursor population for pancreas and liver within the embryonic endoderm. *Development* **128**: 871-881.

Drewes T, Senkel S, Holewa B, Ryffel GU. 1996. Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. *Molecular and cellular biology* **16**: 925-931.

Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, Darnell JE, Jr. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4

is a marker for primary endoderm in the implanting blastocyst. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 7598-7602.

Duncan SA, Nagy A, Chan W. 1997. Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of Hnf-4(-/-) embryos. *Development* **124**: 279-287.

Easwaran V, Pishvaian M, Salimuddin, Byers S. 1999. Cross-regulation of beta-catenin-LEF/TCF and retinoid signaling pathways. *Current biology : CB* **9**: 1415-1418.

Eeckhoute J, Formstecher P, Laine B. 2001. Maturity-onset diabetes of the young Type 1 (MODY1)-associated mutations R154X and E276Q in hepatocyte nuclear factor 4alpha (HNF4alpha) gene impair recruitment of p300, a key transcriptional co-activator. *Molecular endocrinology* **15**: 1200-1210.

Eeckhoute J, Moerman E, Bouckenooghe T, Lukoviak B, Pattou F, Formstecher P, Kerr-Conte J, Vandewalle B, Laine B. 2003. Hepatocyte nuclear factor 4 alpha isoforms originated from the P1 promoter are expressed in human pancreatic beta-cells and exhibit stronger transcriptional potentials than P2 promoter-driven isoforms. *Endocrinology* **144**: 1686-1694.

Erdmann S, Senkel S, Arndt T, Lucas B, Lausen J, Klein-Hitpass L, Ryffel GU, Thomas H. 2007. Tissue-specific transcription factor HNF4alpha inhibits cell proliferation and induces apoptosis in the pancreatic INS-1 beta-cell line. *Biological chemistry* **388**: 91-106.

Fair JH, Cairns BA, Lapaglia M, Wang J, Meyer AA, Kim H, Hatada S, Smithies O, Pevny L. 2003. Induction of hepatic differentiation in embryonic stem cells by co-culture with embryonic cardiac mesoderm. *Surgery* **134**: 189-196.

Frietze S, Wang R, Yao L, Tak YG, Ye Z, Gaddis M, Witt H, Farnham PJ, Jin VX. 2012. Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. *Genome biology* **13**: R52.

Furuta H, Iwasaki N, Oda N, Hinokio Y, Horikawa Y, Yamagata K, Yano N, Sugahiro J, Ogata M, Ohgawara H et al. 1997. Organization and partial sequence of the hepatocyte nuclear factor-4 alpha/MODY1 gene and identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* **46**: 1652-1657.

Gilbert SF. Developmental Biology. 6th edition. Sunderland (MA): Sinauer Associates; 2000. Endoderm. Available from: http://www.ncbi.nlm.nih.gov/books/NBK10107/

Gougelet A, Torre C, Veber P, Sartor C, Bachelot L, Denechaud PD, Godard C, Moldes M, Burnol AF, Dubuquoy C et al. 2014. T-cell factor 4 and beta-catenin chromatin occupancies pattern zonal liver metabolism in mice. *Hepatology* **59**: 2344-2357.

Green VJ, Kokkotou E, Ladias JA. 1998. Critical structural elements and multitarget protein interactions of the transcriptional activator AF-1 of hepatocyte nuclear factor 4. *The Journal of biological chemistry* **273**: 29950-29957.

Hadzopoulou-Cladaras M, Kistanova E, Evagelopoulou C, Zeng S, Cladaras C, Ladias JA. 1997. Functional domains of the nuclear receptor hepatocyte nuclear factor 4. *The Journal of biological chemistry* **272**: 539-550.

Hahn S. 1993. Structure(?) and function of acidic transcription activators. *Cell* **72**: 481-483.

Hannan NR, Segeritz CP, Touboul T, Vallier L. 2013. Production of hepatocyte-like cells from human pluripotent stem cells. *Nature protocols* **8**: 430-437.

Hanse EA, Mashek DG, Becker JR, Solmonson AD, Mullany LK, Mashek MT, Towle HC, Chau AT, Albrecht JH. 2012. Cyclin D1 inhibits hepatic lipogenesis via repression of carbohydrate response element binding protein and hepatocyte nuclear factor 4alpha. *Cell cycle* **11**: 2681-2690.

Hansen SK, Parrizas M, Jensen ML, Pruhova S, Ek J, Boj SF, Johansen A, Maestro MA, Rivera F, Eiberg H et al. 2002. Genetic evidence that HNF-1alpha-dependent transcriptional control of HNF-4alpha is essential for human pancreatic beta cell function. *The Journal of clinical investigation* **110**: 827-833.

Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, Steele A, Njolstad PR, Ellard S, Hattersley AT. 2008. The diabetic phenotype in HNF4A mutation carriers is moderated by the expression of HNF4A isoforms from the P1 promoter during fetal development. *Diabetes* **57**: 1745-1752.

Hata S, Inoue T, Kosuga K, Nakashima T, Tsukamoto T, Osumi T. 1995. Identification of two splice isoforms of mRNA for mouse hepatocyte nuclear factor 4 (HNF-4). *Biochimica et biophysica acta* **1260**: 55-61.

Hata S, Tsukamoto T, Osumi T. 1992. A novel isoform of rat hepatocyte nuclear factor 4 (HNF-4). *Biochimica et biophysica acta* **1131**: 211-213.

Hatziapostolou M, Polytarchou C, Aggelidou E, Drakaki A, Poultsides GA, Jaeger SA, Ogata H, Karin M, Struhl K, Hadzopoulou-Cladaras M et al. 2011. An HNF4alpha-miRNA inflammatory feedback circuit regulates hepatocellular oncogenesis. *Cell* **147**: 1233-1247.

Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ. 2001. Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Molecular and cellular biology* **21**: 1393-1403.

Hittelman AB, Burakov D, Iniguez-Lluhi JA, Freedman LP, Garabedian MJ. 1999. Differential regulation of glucocorticoid receptor transcriptional activation via AF-1-associated proteins. *The EMBO journal* **18**: 5380-5388.

Holewa B, Zapp D, Drewes T, Senkel S, Ryffel GU. 1997. HNF4beta, a new gene of the HNF4 family with distinct activation and expression profiles in oogenesis and embryogenesis of Xenopus laevis. *Molecular and cellular biology* **17**: 687-694.

Horb ME, Shen CN, Tosh D, Slack JM. 2003. Experimental conversion of liver to pancreas. *Current biology : CB* **13**: 105-115.

Huang J, Karakucuk V, Levitsky LL, Rhoads DB. 2008. Expression of HNF4alpha variants in pancreatic islets and Ins-1 beta cells. *Diabetes/metabolism research and reviews* **24**: 533-543.

Huang J, Levitsky LL, Rhoads DB. 2009. Novel P2 promoter-derived HNF4alpha isoforms with different N-terminus generated by alternate exon insertion. *Experimental cell research* **315**: 1200-1211.

Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L. 2011. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **475**: 386-389.

Humphries A, Wright NA. 2008. Colonic crypt organization and tumorigenesis. *Nature reviews Cancer* **8**: 415-424.

Hwang-Verslues WW, Sladek FM. 2008. Nuclear receptor hepatocyte nuclear factor 4alpha1 competes with oncoprotein c-Myc for control of the p21/WAF1 promoter. *Molecular endocrinology* **22**: 78-90.

Ihara A, Yamagata K, Nammo T, Miura A, Yuan M, Tanaka T, Sladek FM, Matsuzawa Y, Miyagawa J, Shimomura I. 2005. Functional characterization of the HNF4alpha isoform (HNF4alpha8) expressed in pancreatic beta-cells. *Biochemical and biophysical research communications* **329**: 984-990.

Jiang G, Nepomuceno L, Hopkins K, Sladek FM. 1995. Exclusive homodimerization of the orphan receptor hepatocyte nuclear factor 4 defines a new subclass of nuclear receptors. *Molecular and cellular biology* **15**: 5131-5143.

Jollie WP. 1990. Development, morphology, and function of the yolk-sac placenta of laboratory rodents. *Teratology* **41**: 361-381.

Kapoor RR, Locke J, Colclough K, Wales J, Conn JJ, Hattersley AT, Ellard S, Hussain K. 2008. Persistent hyperinsulinemic hypoglycemia and maturity-onset diabetes of the young due to heterozygous HNF4A mutations. *Diabetes* **57**: 1659-1663.

Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, Clevers H. 1998. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature genetics* **19**: 379-383.

Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B, Clevers H. 1997. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**: 1784-1787.

Kritis AA, Argyrokastritis A, Moschonas NK, Power S, Katrakili N, Zannis VI, Cereghini S, Talianidis I. 1996. Isolation and characterization of a third isoform of human hepatocyte nuclear factor 4. *Gene* **173**: 275-280.

Kyrmizi I, Hatzis P, Katrakili N, Tronche F, Gonzalez FJ, Talianidis I. 2006. Plasticity and expanding complexity of the hepatic transcription factor network during liver development. *Genes & development* **20**: 2293-2305.

Lazarevich NL, Al'pern DV. 2008. [Hepatocyte nuclear factor 4 (HNF4) in epithelial development and carcinogenesis]. *Molekuliarnaia biologiia* **42**: 786-797.

Lazarevich NL, Cheremnova OA, Varga EV, Ovchinnikov DA, Kudrjavtseva EI, Morozova OV, Fleishman DI, Engelhardt NV, Duncan SA. 2004. Progression of HCC in mice is associated with a downregulation in the expression of hepatocyte nuclear factors. *Hepatology* **39**: 1038-1047.

Loh KM, Lim B. 2013. Stem cells: Close encounters with full potential. *Nature* **502**: 41-42.

Ma H, Hong H, Huang SM, Irvine RA, Webb P, Kushner PJ, Coetzee GA, Stallcup MR. 1999. Multiple signal input and output domains of the 160-kilodalton nuclear receptor coactivator proteins. *Molecular and cellular biology* **19**: 6164-6173.

Maeda Y, Rachez C, Hawel L, 3rd, Byus CV, Freedman LP, Sladek FM. 2002. Polyamines modulate the interaction between nuclear receptors and vitamin D receptor-interacting protein 205. *Molecular endocrinology* **16**: 1502-1510.

Mallanna SK, Duncan SA. 2013. Differentiation of hepatocytes from pluripotent stem cells. *Current protocols in stem cell biology* **26**: Unit 1G 4.

Mangelsdorf DJ, Evans RM. 1995. The RXR heterodimers and orphan receptors. *Cell* **83**: 841-850.

Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P et al. 1995. The nuclear receptor superfamily: the second decade. *Cell* **83**: 835-839.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 7634-7638.

Medema JP, Vermeulen L. 2011. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* **474**: 318-326.

Miura A, Yamagata K, Kakei M, Hatakeyama H, Takahashi N, Fukui K, Nammo T, Yoneda K, Inoue Y, Sladek FM et al. 2006. Hepatocyte nuclear factor-4alpha is essential for glucose-stimulated insulin secretion by pancreatic beta-cells. *The Journal of biological chemistry* **281**: 5246-5257.

Montserrat N, Nivet E, Sancho-Martinez I, Hishida T, Kumar S, Miquel L, Cortina C, Hishida Y, Xia Y, Esteban CR et al. 2013. Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell stem cell* **13**: 341-350.

Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, Kinzler KW. 1997. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* **275**: 1787-1790.

Mulholland DJ, Dedhar S, Coetzee GA, Nelson CC. 2005. Interaction of nuclear receptors with the Wnt/beta-catenin/Tcf signaling axis: Wnt you like to know? *Endocrine reviews* **26**: 898-915.

Munemitsu S, Albert I, Souza B, Rubinfeld B, Polakis P. 1995. Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 3046-3050.

Nakhei H, Lingott A, Lemm I, Ryffel GU. 1998. An alternative splice variant of the tissue specific transcription factor HNF4alpha predominates in undifferentiated murine cell types. *Nucleic acids research* **26**: 497-504.

Nikolenko Iu V, Krasnov AN. 2007. [Nuclear receptors: structure and mechanisms of action]. *Genetika* **43**: 308-316.

Ning BF, Ding J, Yin C, Zhong W, Wu K, Zeng X, Yang W, Chen YX, Zhang JP, Zhang X et al. 2010. Hepatocyte nuclear factor 4 alpha suppresses the development of hepatocellular carcinoma. *Cancer research* **70**: 7640-7651.

Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK et al. 2004. Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**: 1378-1381.

Oshima T, Kawasaki T, Ohashi R, Hasegawa G, Jiang S, Umezu H, Aoyagi Y, Iwanari H, Tanaka T, Hamakubo T et al. 2007. Downregulated P1 promoter-driven hepatocyte nuclear factor-4alpha expression in human colorectal carcinoma is a new prognostic factor against liver metastasis. *Pathology international* **57**: 82-90.

Parr EL, Parr MB. 1986. Uptake of immunoglobulins and other proteins from serum into epithelial cells of the mouse uterus and oviduct. *Journal of reproductive immunology* **9**: 339-354.

Pearson ER, Boj SF, Steele AM, Barrett T, Stals K, Shield JP, Ellard S, Ferrer J, Hattersley AT. 2007. Macrosomia and hyperinsulinaemic hypoglycaemia in

patients with heterozygous mutations in the HNF4A gene. *PLoS medicine* **4**: e118.

Plengvidhya N, Antonellis A, Wogan LT, Poleev A, Borgschulze M, Warram JH, Ryffel GU, Krolewski AS, Doria A. 1999. Hepatocyte nuclear factor-4gamma: cDNA sequence, gene organization, and mutation screening in early-onset autosomal-dominant type 2 diabetes. *Diabetes* **48**: 2099-2102.

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M et al. 2013. Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**: 65-70.

Rao MS, Reddy JK. 1995. Hepatic transdifferentiation in the pancreas. *Seminars in cell biology* **6**: 151-156.

Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH. 2004. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature medicine* **10**: 55-63.

Satoh S, Hinoda Y, Hayashi T, Burdick MD, Imai K, Hollingsworth MA. 2000. Enhancement of metastatic properties of pancreatic cancer cells by MUC1 gene encoding an anti-adhesion molecule. *International journal of cancer Journal international du cancer* **88**: 507-518.

Schweizer L, Rizzo CA, Spires TE, Platero JS, Wu Q, Lin TA, Gottardis MM, Attar RM. 2008. The androgen receptor can signal through Wnt/beta-Catenin in prostate cancer cells as an adaptation mechanism to castration levels of androgens. *BMC cell biology* **9**: 4.

Sekiya S, Suzuki A. 2011. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475**: 390-393.

Shen CN, Slack JM, Tosh D. 2000. Molecular basis of transdifferentiation of pancreas to liver. *Nature cell biology* **2**: 879-887.

Shu J, Wu C, Wu Y, Li Z, Shao S, Zhao W, Tang X, Yang H, Shen L, Zuo X et al. 2013. Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**: 963-975.

Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, North PE, Dalton S, Duncan SA. 2010. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* **51**: 297-305.

Sladek FM. 2011. What are nuclear receptor ligands? *Molecular and cellular endocrinology* **334**: 3-13.

Sladek FM, and Seidel, S.D. 2001. In Nuclear Receptors and Genetic Diseases. *TP Burris and ERB McCabe, eds (London: Academic Press)*: pp.309-361.

Sladek FM, Ruse MD, Jr., Nepomuceno L, Huang SM, Stallcup MR. 1999. Modulation of transcriptional activation and coactivator interaction by a splicing variation in the F domain of nuclear receptor hepatocyte nuclear factor 4alpha1. *Molecular and cellular biology* **19**: 6509-6522.

Sladek FM, Zhong WM, Lai E, Darnell JE, Jr. 1990. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes & development* **4**: 2353-2365.

Song Z, Cai J, Liu Y, Zhao D, Yong J, Duo S, Song X, Guo Y, Zhao Y, Qin H et al. 2009. Efficient generation of hepatocyte-like cells from human induced pluripotent stem cells. *Cell research* **19**: 1233-1242.

Soto-Gutierrez A, Navarro-Alvarez N, Caballero-Corbalan J, Tanaka N, Kobayashi N. 2008. Endoderm induction for hepatic and pancreatic differentiation of ES cells. *Acta medica Okayama* **62**: 63-68.

Soto-Gutierrez A, Navarro-Alvarez N, Zhao D, Rivas-Carrillo JD, Lebkowski J, Tanaka N, Fox IJ, Kobayashi N. 2007. Differentiation of mouse embryonic stem cells to hepatocyte-like cells by co-culture with human liver nonparenchymal cell lines. *Nature protocols* **2**: 347-356.

Spath GF, Weiss MC. 1997. Hepatocyte nuclear factor 4 expression overcomes repression of the hepatic phenotype in dedifferentiated hepatoma cells. *Molecular and cellular biology* **17**: 1913-1922.

Spath GF, Weiss MC. 1998. Hepatocyte nuclear factor 4 provokes expression of epithelial marker genes, acting as a morphogen in dedifferentiated hepatoma cells. *The Journal of cell biology* **140**: 935-946.

Stanescu DE, Hughes N, Kaplan B, Stanley CA, De Leon DD. 2012. Novel presentations of congenital hyperinsulinism due to mutations in the MODY genes: HNF1A and HNF4A. *The Journal of clinical endocrinology and metabolism* **97**: E2026-2030.

Suaud L, Formstecher P, Laine B. 1999. The activity of the activation function 2 of the human hepatocyte nuclear factor 4 (HNF-4alpha) is differently modulated by F domains from various origins. *The Biochemical journal* **340 (Pt 1)**: 161-169.

Suda Y, Suzuki M, Ikawa Y, Aizawa S. 1987. Mouse embryonic stem cells exhibit indefinite proliferative potential. *Journal of cellular physiology* **133**: 197-201.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Takano K, Hasegawa G, Jiang S, Kurosaki I, Hatakeyama K, Iwanari H, Tanaka T, Hamakubo T, Kodama T, Naito M. 2009. Immunohistochemical staining for P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha may complement mucin phenotype of differentiated-type early gastric carcinoma. *Pathology international* **59**: 462-470.

Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C et al. 2006. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. *The Journal of pathology* **208**: 662-672.

Taraviras S, Monaghan AP, Schutz G, Kelsey G. 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mechanisms of development* **48**: 67-79.

Teo R, Mohrlen F, Plickert G, Muller WA, Frank U. 2006. An evolutionary conserved role of Wnt signaling in stem cell fate decision. *Developmental biology* **289**: 91-99.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. 1998. Embryonic stem cell lines derived from human blastocysts. *Science* **282**: 1145-1147.

Torres-Padilla ME, Fougere-Deschatrette C, Weiss MC. 2001. Expression of HNF4alpha isoforms in mouse liver development is regulated by sequential promoter usage and constitutive 3' end splicing. *Mechanisms of development* **109**: 183-193.

Torres-Padilla ME, Sladek FM, Weiss MC. 2002. Developmentally regulated N-terminal variants of the nuclear receptor hepatocyte nuclear factor 4alpha mediate multiple interactions through coactivator and corepressor-histone deacetylase complexes. *The Journal of biological chemistry* **277**: 44677-44687.

Torres-Padilla ME, Weiss MC. 2003. Effects of interactions of hepatocyte nuclear factor 4alpha isoforms with coactivators and corepressors are promoter-specific. *FEBS letters* **539**: 19-23.

Truica CI, Byers S, Gelmann EP. 2000. Beta-catenin affects androgen receptor transcriptional activity and ligand specificity. *Cancer research* **60**: 4709-4713.

van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, van der Horn K, Batlle E, Coudreuse D, Haramis AP et al. 2002. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**: 241-250.

Walesky C, Edwards G, Borude P, Gunewardena S, O'Neil M, Yoo B, Apte U. 2013a. Hepatocyte nuclear factor 4 alpha deletion promotes diethylnitrosamine-induced hepatocellular carcinoma in rodents. *Hepatology* **57**: 2480-2490.

Walesky C, Gunewardena S, Terwilliger EF, Edwards G, Borude P, Apte U. 2013b. Hepatocyte-specific deletion of hepatocyte nuclear factor-4alpha in adult mice results in increased hepatocyte proliferation. *American journal of physiology Gastrointestinal and liver physiology* **304**: G26-37.

Wang H, Maechler P, Antinozzi PA, Hagenfeldt KA, Wollheim CB. 2000. Hepatocyte nuclear factor 4alpha regulates the expression of pancreatic beta -cell genes implicated in glucose metabolism and nutrient-induced insulin secretion. *The Journal of biological chemistry* **275**: 35953-35959.

Waterman ML. 2004. Lymphoid enhancer factor/T cell factor expression in colorectal cancer. *Cancer metastasis reviews* **23**: 41-52.

Watt AJ, Garrison WD, Duncan SA. 2003. HNF4: a central regulator of hepatocyte differentiation and function. *Hepatology* **37**: 1249-1253.

Willert K, Jones KA. 2006. Wnt signaling: is the party in the nucleus? *Genes & development* **20**: 1394-1404.

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**: 684-687.

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI. 1996. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* **384**: 458-460.

Yamamizu K, Piao Y, Sharov AA, Zsiros V, Yu H, Nakazawa K, Schlessinger D, Ko MS. 2013. Identification of Transcription Factors for Lineage-Specific ESC Differentiation. *Stem cell reports* **1**: 545-559.

Yang M, Li SN, Anjum KM, Gui LX, Zhu SS, Liu J, Chen JK, Liu QF, Ye GD, Wang WJ et al. 2013. A double-negative feedback loop between Wnt-beta-catenin signaling and HNF4alpha regulates epithelial-mesenchymal transition in hepatocellular carcinoma. *Journal of cell science* **126**: 5692-5703.

Yi F, Liu GH, Izpisua Belmonte JC. 2012. Rejuvenating liver and pancreas through cell transdifferentiation. *Cell research* **22**: 616-619.

Yin C, Lin Y, Zhang X, Chen YX, Zeng X, Yue HY, Hou JL, Deng X, Zhang JP, Han ZG et al. 2008. Differentiation therapy of hepatocellular carcinoma in mice with recombinant adenovirus carrying hepatocyte nuclear factor-4alpha gene. *Hepatology* **48**: 1528-1539.

Yuan X, Ta TC, Lin M, Evans JR, Dong Y, Bolotin E, Sherman MA, Forman BM, Sladek FM. 2009. Identification of an endogenous ligand bound to a native orphan nuclear receptor. *PloS one* **4**: e5609.

Yusuf D Butland SL Swanson MI Bolotin E Ticoll A Cheung WA Zhang XY Dickman CT Fulton DL Lim JS et al. 2012. The transcription factor encyclopedia. *Genome biology* **13**: R24.

Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S et al. 2014. Proteogenomic characterization of human colon and rectal cancer. *Nature*

Chapter 2

Generation of the Tet-On Inducible Clones

Abstract

The Tet-On system is a great tool to control the expression and activity of an individual gene of interest, making it relatively easy to study the function of any given gene in a cell. We used this method to elucidate the role of the promoter-driven HNF4α isoforms in human colon cancer (HCT116) and mouse embryonic stem (mES) D3 cells neither of which express endogenous HNF4α. In this Chapter, we described the generation of Tet-On inducible HCT116 and mES lines that express either the human HNF4α2 or HNF4α8 under control of doxycycline (DOX). We characterized these lines to test for the efficacy of the inducible system. We observed that the Tet-On system produces a high level of HNF4α-mediated transactivation and gives a close-to-physiological expression of the protein. We also induced the parental lines (PL) with DOX and performed RNA-Seq to identify genes that were changed by DOX alone. There were a few genes that were dysregulated by DOX, but these genes have low fold changes (less than 2.0, non-log). We also found that the PL and two daughter HNF4α2 and HNF4α8 lines (α2 and α8) are different from one another. All told, these inducible lines are good model systems with which to study the function of HNF4α isoforms in cancer cells and early development.

Introduction

The tetracycline-controlled (Tet) transcriptional regulation is a commonly used inducible system to control the expression of transgenes in order to study their function under various cell-based conditions. There are three Tet-based systems: (1) the original Tet-controlled transcriptional activator (tTA); (2) the modified version of the tTA, reverse tetracycline-controlled transcriptional activator (rtTA); and (3) the recently developed

tetracycline-controlled transcriptional silencer (tTS). The tTA protein is a fusion of the tetracycline repressor (TetR) and the activation domain (VP16), from the bacterium *Escherichia coli* and Herpes Simplex Virus, respectively (Gossen et al. 1992).

There are two functional properties to this system each encoded in a separate vector. The first vector contains the tTA gene and an upstream promoter that drives the expression of tTA in a constitutive manner. The second vector has a multiple cloning region with which one can insert any gene of interest downstream of a tetracycline response element (TRE) that consists of the specific Tet operator sequences of E. coli (TetO) and a minimal promoter sequence of the human cytomegalovirus promoter (CMV), which is downstream of the TetO. The activity of the tTA is controlled by tetracycline (Tc) or its derivatives. In the absence of the ligand, tTA binds the TRE element and activates transcription. Upon addition of Tc, tTA activity is impaired leading to a Tet-Off system. There is also a Tet-On system in which a mutation is introduced in the tTA gene to become rtTA; and the TRE sequence is changed in such a way that the rtTA binds the TRE in the presence of the ligand (Gossen et al. 1995). The Tet-On system is faster and does not require continuous administration of Tc to silence the gene. Various tTA mutants have been screened to identify rtTA that produces low to no basal activity but has increased sensitivity to the ligand. In combination with the Tet-On system the Tet silencer (tTS) was generated which consists of a fusion of the Tet Repressor (TetR) protein and the KRAB-AB domain of Kid-1 protein, a transcriptional repressor (Freundlieb et al. 1999). In the absence of ligand, tTS bind TRE, preventing the rtTA from binding. In the presence of the ligand, tTS no longer binds the TRE element, allowing rtTA to bind and activate transcription of the transgene. This additional property provides a greater level of control of the transgene expression. These inducible

systems are constantly being improved to provide a greater control of gene expression (Zhu et al. 2002).

Here we demonstrate the use of the Tet-On system in two mammalian systems – the human colon cancer cell (HCT116) line and a mouse embryonic stem cell line D3. We show that the expression of HNF4α can be tightly regulated in a temporal fashion under physiological conditions and that DOX alone does not affect the cells significantly. Thus, this is a simple system that can be used to look at the differential role of the promoter-driven (P1 and P2) HNF4α isoforms in normal and cancer development.

Materials and Methods

Cell Culture

The Human colorectal cancer cell line HCT116 was cultured in McCoy's 5A (Iwakata & Grace Modification, with L-glutamine) modified medium supplemented with 10% Fetal Bovine Serum (FBS) (Benchmark Cat. #100-106 Lot. # A07D00C), and 1% penicillin and streptomycin (P/S). Mouse embryonic stem (mES) D3 cells (a gift from Dr. Prudence Talbot at the University of California, Riverside) were, first, cultured on irradiated mouse embryonic fibroblast (MEF) in mES medium that had the following compositions DMEM (Dulbecco's Modification of Eagle's Medium with 4.5 g/L glucose, L-glutamine, and pyruvate) supplemented with 15% FBS, 1% nonessential amino acids (NEAA), 1% P/S, 1% sodium pyruvate, 1% L-glutamine or glutmax, and 0.2% β-mercaptoethanol. After subsequent passages, the cells were cultured in MEF-conditioned medium (CM) (incubated with MEFs at 37^{o}C for 24 h, then stored at -20^{o}C), plus leukemia inhibitory factor (LIF) (100-1000 U/mL) (Millipore). Cells were maintained at 37^{o}C and 5% CO_2.

MEF Derivation

The MEFs, used to support mES culture and to generate CM, were derived from E13.5 embryos from a mixed background mice. MEF derivations were done using previously described protocols (Garfield et al. 2010; WiCell Protocol 2003, Research Institute, Inc.) with modifications. Mouse at 13.5-14 days of gestation was sacrificed via CO_2 asphyxiation. The abdomen was sterilized with 95% ethanol and cut open to remove the uterine horns. The horns were washed with 1 x PBS and embryonic sacs were separated in PBS, using forceps and scissors. Internal organs, heads, and limbs were removed from the embryos, while the rest of the tissues were minced in 2 mL of trypsin with a razor blade. An additional 5 mL of trypsin (0.5%) was added to the minced tissue and placed in 37^{o}C for 20-30 min. Cells were dislodged by pipetting up and down multiple times and cultured in MEF media (DMEM (Dulbecco's Modification of Eagle's Medium with 4.5 g/L glucose, L-glutamine, and pyruvate) supplemented with 10% FBS, 1% NEAA, 1% P/S, 1% sodium pyruvate, 1% L-glutamine) in 100-mm plates or T75 Flasks for expansion or freezing. Some MEFs were irradiated in the Faxitron machine at 90kVp for one hour to generate feeders. Aliquots of these cells were frozen and kept at -80^{o}C for later use. The University of California, Riverside, Institutional Animal Care and Use Committee (IACUC) guidelines were followed when handling mice.

Plasmids

The full-length human HNF4α2 (NM_000457) and HNF4α8 (NM_175914.3) in pcDNA3.1 (pcDNA3.1.HNF4α2 & pcDNA3.1.HNF4α8) were gifts from Dr. Christophe Rachez at Pasteur Institute, Paris, France (Chartier et al. 1994; Eeckhoute et al. 2001).

pTRE-HNF4α2 and pTRE-HNF4α8 were made by amplifying the human HNF4α2 and HNF4α8 constructs in the pcDNA3.1 (+) vector, using PCR primers that contain an EcoRI [plus a Kozak's sequence] and a BamHI site in the 5' and 3' end, respectively, and ligating the PCR fragment into the pTRE-Tight vector (Clontech) at the respective (EcoRI/BamHI) sites. The inserted HNF4αa2 and HNF4α8 cDNA from the pcDNA3.1 vector do not include the upstream start site in the N-terminal region. The reverse tetracycline-controlled transcriptional activator pCAG-rtTA (pCAG-rtTA-IR-PURO) and the pTRE-RFP (pTRE-Red Max C) plasmids were gifts from Dr. Chee-Gee Liew (Liew et al 2007; Xia et al. 2008). The pTRE-RFP reporter vector contains red fluorescence protein driven by a Tet-expression element, the construct of this plasmid as is the pTRE-HNF4α2/α8 (Fig. 2.1, bottom). The ApoB.-85-47.E4.Luc reporter (Maeda Y et al. 2006) is an HNF4α reporter target. The pTRE-EGFP-HNF4α2 and pTRE-EGFP-HNF4α8 were constructed in the same manner as the pTRE-HNF4α2 and pTRE-HNF4α8, except the 5' primer containing the KpnI site (underline), Kozak's sequence, and EGFP sequence (bold) was used to amplify the human EGFP-HNF4α2 and EGFP-HNF4αa8 constructs in the pcDNA3.1 (+) vector: 5'-GGGGTACCCCGCCACC**ATGGTG**-3'.

Generation of Tet-On HNF4α Inducible Cell Lines in mES and HCT116 cells

mES Stable Lines

mES cells seeded at 5 x 10^4 cells per well of a 6-well plate were co-cultured with MEFs in mES medium. The next day, 1 µg of linearized pCAG-rtTA DNA was transfected into the cells using HiFect (LONZA). Twenty-four hours later, mES medium was changed to CM with 2 µg/mL of puromycin. Small colonies were picked on the ninth day and transferred to a 0.1% gelatin-coated 12-well plate (all plates were gelatin coated unless

indicated otherwise). Confluent colonies were passaged into two wells of a 12-well plates; one plate was expanded, while the second plate was transfected with 1 µg of pTRE-RFP and treated with 1 µg/mL of doxycycline (DOX, dissolved in water, Clontech). Expression of Red fluorescent proteins (RFP) were observed under a Nikon fluorescence microscope. Clone 3.1 at passage 30 was selected to continue the second step of the cloning procedure. Cells were seeded at 2.5-3.5 x 10^5 cells per well of a 6-well plate. Linearized pTRE-HNF4α2 or pTRE-HNF4α8 plus a fragment of the Neo^R gene cut from the Tet-On vector (Clontech), at a 10:1 ratio respectively, were transfected into the cells using Lipofectamine 2000 (Invitrogen). Feeders (8 x 10^6) were added on top of the cells for support. Cells were selected with 50 µg/mL of G418 and 1.5 µg/mL of puromycin. Colonies were picked on the eleventh day of selection and transferred to MEF-coated 24-well plates. Cells were induced with 1 µg/mL DOX and protein extracts were analyzed on IBs to identify positive clones.

HCT116 Stable Lines

Cells were seeded at 3 x 10^6 cells per well of a 6-well plate. The next day, 1 µg of linearized pCAG-rtTA DNA was transfected into the cells using lipofectamine 2000. Twenty-four hours after transfection, cells were trypsinized and transferred to a 150-mm plate. Cells were selected with 0.50 µg/mL puromycin. Colonies were picked and tested for rtTA activity. Cells were seeded at 1.76 x 10^5 cells per well of a 24-well. The next day cells were transfected with 150 ng ApoB.-85-47.E4.Luc and 1.0 ng of pTRE-HNF4α2 expression vector, induced with 1 µg/mL DOX, and luciferase assay was performed; the luciferase assay is described in the Materials and Methods (M&M) of Chapter 3. The luciferase measurement was reported as mean of the relative luminescence units (RLU)

±SD, in triplicates from one independent study. Clones (cl11 & cl17) with the highest rtTA activity went through the second step. Linearized pTRE-HNF4α2 or pTRE-HNF4α8 plus a fragmented Neo^R gene cut out from the pTet-On vector (Clontech) at the XhoI sites, at a 10:1 ratio respectively, were transfected into the cells (1.3×10^6) using lipofectamine 2000. Cells were selected starting with 50 µg/mL of the G418 antibiotic and increased to 70 µg/mL, while using the same puromycin concentration as in the first step. Positive clones were expanded in 24-well and 6-well plates.

Immunoblot (IB) analysis

Protein extracts (~20 µg) from whole cell extracts (WCE) or nuclear extracts (NE) were separated by 10% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) as described in Maeda Y et al. 2002. The following primary (overnight, at 4^0C) and secondary (~2 h, at room temperature (RT)) antibodies were used: monoclonal anti-P1/P2 (R&D, Cat#PP-H1415-00) (1:30,000 to 1:60,000) that recognizes the C-terminus of both HNF4α isoforms, anti-OCT4 (Santa Cruz, sc9081) (1:5000), and horseradish peroxidase (HRP)-conjugated goat anti-rabbit (GαR-HRP) or goat anti-mouse (GαM-HRP) from the Jackson ImmunoResearch Laboratories (1:5000 to 1:20,000). SuperSignal West Dura Extended Duration Substrate Kit (Thermo Scientific) was used to develop the blots. Equal loading of the protein was verified by Coomassie staining of the blot. Blot was stained with 0.25% Coomassie in 45% methanol (MeOH) and 10% glacial acetic acid for 30 seconds and destained with 1 to 2 washes, 5 – 10 min each of 50% MeOH and 10% glacial acetic acid.

Immunofluorescence (IF)

Cells seeded at 1-3 x 10^5 cells per well of a 6-well or 12-well plate were fixed in 3.7-4% formaldehyde for 10-15 min at RT or 37^{o}C incubator, blocked with 10% goat and donkey serums (GS, DS) and 0.2% Triton X-100 in PBS at RT for 30 min, and then with 1% GS and DS in PBS at 4^{o}C overnight. Primary antibodies include HNF4α P1/P2 (R&D, cat#PP-H1415-00) at 1:1200 and OCT4 (Santa Cruz, sc9081) at 1:200 and secondary Alexa Fluor antibodies at 1:500 (Invitrogen); 4'6-diamidino-2-phenylindole (DAPI) (Invitrogen) was used at 1:1000 to visualize nuclei. The Nikon Eclipse Ti fluorescence microscope in the Stem Cell Core facility at UCR was used to visualize the cells.

RNA-Seq Analysis

See Chapter 3 M&M for full description. The pairwise comparisons of PL 0.3 μg/mL vs PL 0.0 μg/mL, α2 0.0 μg/mL vs PL 0.0 μg/mL, and α8 0.0 μg/mL vs PL 0.0 μg/mL in the RNA-Seq data are discussed here (concentrations refer to the amount of DOX). All non-log fold changes are significant with P-value $\leq$ 0.01 and Q-value $\leq$ 0.05, triplicate samples per condition.

Results

Generation of Tet-On inducible mES and HCT116 cell lines

To clone the HNF4α cDNA into the pTRE-Tight plasmid, we amplified the human HNF4α2 and HNF4α8 cDNA from pcDNA3.1 (+) vectors using primers that contain an EcoRI site and a Kozak's sequence at the 5' end and a BamHI site at the 3' end. We inserted the amplified fragments into the commercially available pTRE-Tight plasmid that had been digested with EcoRI and BamHI enzymes (Fig. 2.2). The final product consists

of the target cDNA (HNF4α2 or HNF4α8) downstream of the human CMV promoter and the TetO operon promoters (pTRE-HNF4α2 and pTRE-HNF4α8) (Fig. 2.1, bottom right).

In the reverse tetracycline-controlled transcriptional activator (rtTA), the gene that codes for the rtTA protein is under the control of the promoter CAG that consists of the chicken β-actin and the rabbit β-globin promoter (CAGG), polyoma virus mutant enhancer (PyF101) that is upstream of the CAGG, and an internal ribosomal entry site (IRES) that is downstream of the CAGG (pCAG-rtTA-IR-PURO) (Liew et al. 2007; Niwa et al. 1991; Attal et al. 1999) (Fig. 2.1, top). The composite promoter CAGG is made up of several elements: enhancer elements from HCMV-MIEP (the major immediate early promoter (MIEP) of human cytomegalovirus (HCMV)) coupled with promoter elements of the chicken β-actin promoter and rabbit β-globin gene (pCAGG) (Liew et al. 2007; Niwa et al. 1991). The IRES allows the CAG promoter to drive the expression of a bicistronic message (Attal et al. 1999). Liew et al showed that the CAGG and CAG (CAGG that is linked to the polyoma virus mutant enhancer (PyF101) and an IRES) promoters are more active than the CMV promoter in ES cells although the latter retain expression and activity longer even after puromycin selection (Liew et al. 2007). The CAG promoter does not get silenced as do the CMV and CAGG promoters in stable transfections and hence, the CAG promoter is more effective for generating stable transfections in mouse embryonic stem (mES) cells (Liew et al. 2007; Nelson et al. 1990).

Generation and characterization of the stable lines that express HNF4α variants

To generate Tet-On inducible stable HCT116 and mES lines that express either the human HNF4α2 or HNF4α8 under control of DOX, we performed a two-step cloning procedure (Fig. 2.3). In the first step we transfected in the pCAG-rtTA-IR-PURO plasmid

and obtained more than ~500 (HCT116) to ~1000 (mES) colonies that were resistant to puromycin. These clones were screened for high rtTA activity before proceeding to the second step. In mES clones, we transiently transfected in the pTRE-RFP and screened for high expression of the red fluorescence protein (RFP) (Fig. 2.4A). In HCT116 clones, we transfected in the pTRE-HNF4α2 expression vector and ApoB.-85-47.E4.Luc reporter and screened for high levels of luciferase in response to DOX induction (Fig. 2.4B). Clones 11 and 17 (HCT116) and clone 3.1 (mES) that had the highest rtTA activity or RFP expression were further processed. Finally in the second step we transfected in the pTRE-HNF4α2 or pTRE-HNF4α8 and obtained more than ~100-300 (mES) to ~1000 (HCT116) colonies that were resistant to the G418 antibiotic. About thirty clones of mES and ~200 clones of HCT116 were picked; protein extracts from some of these clones were analyzed on IBs to check for positive clones that express either HNF4α2 or HNF4α8 in the presence of doxycycline (DOX) (Fig. 2.5). We selected clones that have moderate to high level of HNF4α expression in the presence of DOX but no background expression in the absence of DOX: 3.1.2 (HNF4α2) and 3.1.8 (HNF4α8) – mES clones; 11.5 and 17.1 (HNF4α8), and 11.8 and 17.83 (HNF4α2) – HCT116 clones.

Next, we characterized these lines. We first determined if these clones express the correct form of the promoter-driven HNF4α isoforms. Immunoblots (IB) analysis using the P1/P2 HNF4α antibody that recognizes both HNF4α isoforms showed that clones 3.1.2 and 11.8 expressed HNF4α2 and clones 3.1.8 and 11.5 expressed HNF4α8. HNF4α2 with its full-length A/B domain migrates slower than HNF4α8, which has a truncated A/B domain (Fig. 2.6A). We also test different DOX concentrations to determine the optimal expression of HNF4α. IB analysis showed that 0.25 to 0.3 μg/mL DOX induces similar levels of HNF4α2 and HNF4α8 expression in the HCT116 and mES

inducible clones (Fig. 2.6B). Immunoflorescence (IF) staining showed that about 85% to 90% of the cells expressed HNF4α (Fig. 2.6C,D). Interestingly, the expression of the protein is loss over time even with continuous induction of fresh DOX, both in stable and transient transfection for HNF4α but not RFP (Fig.s 2.6C, 2.7, and 3.1C, Chapter 3); this would suggest that ectopic expression of HNF4α is unstable and does not turn on endogenous HNF4α expression in the cells. The HNF4α2/α8 cDNA is under the control of a modified tetracycline response element (TRE), which consists of seven direct repeats of the 19 bp TetO operator sequences and a downstream partial CMV mini promoter that lacks an enhancer (Clontech, cat#631059). As noted before, the CMV promoter is silenced in ES cells. Although the CAG promoter element could drive HNF4α expression for some time, at later time-points, CMV promoter somehow gets silenced, and HNF4α expression declines.

Effect of DOX on the inducible HCT116 parental line (PL)

To determine whether DOX alone induces changes in gene expression, the HCT116 parental cells (PL) were treated with or without 0.3 μg/mL DOX for 24 h and submitted for Next Generation Sequencing (RNA-Seq). We identified six genes (*ECM1, WFDC3 and MORN1*) whose expression increased 1.5 to 1.8-fold but none that increased at the 2-fold or more. There were, however, four genes (*MORN1, ASB18 (GBX2), CYP24A1, and DHRS2*) that were downregulated 1.6 to 2.0-fold. Since only a small number of genes changed by DOX, this would indicate that the concentration used in the induction did not have any long lasting adverse effect on the cells (See Table 2.1 for genes with their non-log fold changes).

Differences between the HCT116 Parental, HNF4α2, and HNF4α8 lines (PL, α2, and α8) in the absence of DOX

We next compared the HNF4α2 and HNF4α8 lines to the PL and to each other in the HCT116 lines. Although the daughter lines were derived from the parental line, after the selection process and random integration of the expression vector into the genome, these lines somehow acquired a slightly different mRNA profile from their parent. There were about 14 genes dysregulated in the α2 line while 39 genes were dysregulated in the α8 line $\geq$ 2.0-fold compared to the parental line (Table 2.1), even in the absence of DOX. The gene with the largest fold change in the α2 line is IZUMO3 family member 3 (chromosome 9 open reading frame 134). The gene with the largest fold change in the α8 line is keratin 23 (*KRT23*), a member of the keratin family, which are intermediate filament proteins that keep the integrity of the epithelial structure intact (http://www.ncbi.nlm.nih.gov/gene/25984). The genes that were downregulated the most in the α2 and α8 lines were solute carrier family 43 member (*SLC43A3*) and VGF nerve growth factor inducible (*VGF*), respectively. Interestingly, *SPTSSB* – a gene that codes for the serine palmitoyltransfersase (small subunit B), which catalyzes the first enzymatic step in sphingolipid biosynthesis – is 3.5-fold down in the α2 line while it is 2.9-fold up in the α8 line.

We are in the process of performing RNA-Seq for the inducible mES cell lines. We anticipate that there will be minor differences in gene expression between these lines (PL, α2, and α8) but those changes will result in functional differences. We know for now is that the mESrtTA lines express the pluripotency marker, OCT4, when cultured on mouse embryonic fibroblasts (MEFs), suggesting that these cells are still pluripotent

(Fig. 2.8B). Furthermore, neither the HCT116 and mES inducible systems show any morphological changes in the absence of DOX (Fig. 2.8A,B).

HNF4α mRNA and protein levels in HCT116 inducible lines in the absence or presence of DOX

The IB analysis did not detect HNF4α expression in HCT116 inducible lines (PL, α2, α8) (Fig.s 2.5 and 2.6) while the RNA-Seq analysis of HNF4α showed that there is background expression in the absence of DOX, slightly more in the α8 than α2 line (Fig. 2.9A; Table 2.2). This is probably because the absolute value of the HNF4α mRNA is $\leq$10, which is below the detectable range in an IB. Interestingly, the PL line in the absence and presence of DOX also has some expression but the HNF4α transcript is missing exons 1D and 1A (Table 2.2). Similarly, the HNF4α mRNA is also missing exons 1D and 1A in the α2 line (-DOX) while the 1A is missing in the α8 line (-DOX) (Table 2.2), suggesting that even though there is background expression, the protein might be non-functional and unstable. Another interesting observation is that the IB analysis showed that the HNF4α8 expression tends to be more than the HNF4α2 expression in the HCT116 lines (α2 and α8) in the presence of DOX (Fig. 2.6A and B). This is reflective in the mRNA level; in the presence of DOX, the absolute value of HNF4α8 is ~800 and the value for HNF4α2 is about half of that. One probable cause could be that HNF4α2 is prone to Src phosphorylation on the full length A/B domain; there is an appreciable amount of Src expressed in HCT116 cells (Chellappa et al. 2012; Welman et al. 2006). Chellappa et al (2012) showed that active Src targets the P1-HNF4α for degradation in human colorectal cancer cells.

Discussion

The system we currently are using is the mutant form of the tTA. It is a better-controlled system than the regular tTA. DOX is added to induce expression of the gene and not to block induction. We optimized the concentration of DOX to minimize inhibition of proliferation and promotion of apoptosis by DOX (Fife et al. 1997; Son et al. 2009). Son et al showed that (≥10 μg/mL) DOX inhibits proliferation and promotes apoptosis by inducing the expression of p53 and its downstream target gene *CDKN1A* in the human pancreatic tumor cell line (PANC-1) (Son et al. 2009). However, we did not see the *TP53* or *CDKN1A* gene changed in the HCT116 lines (PL, α2, α8) in the absence (0.0 μg/mL) or presence (0.3 μg/mL) of DOX. We did observe a decrease in cell proliferation and tumor growth in our cell count, MTT, and xenograft assays for the HCT116 PL in the presence of DOX (Fig.s 3.2 and 3.1.S2A and B, Chapter 3 and Appendix of Chapter 3), suggesting that DOX has a slight effect on the growth ability of HCT116 cells. In contrast, the mES PL in the presence (0.3 μg/mL) of DOX did not experience a decrease in cell number in the cell count assay (Fig.s 4.1.S1A,D and 4.1.S2C, Appendix to Chapter 4). Several reports indicated that p53 is inactive in ES cells despite high protein abundance and becomes activated in differentiated ES cells (Aladjem et al. 1998; Qin et al. 2007; reviewed in Solozobova and Blattner 2011). One group showed that ES cells response to DNA damage in a p53-indpendent apoptosis (Aladjem et al. 1998).

DOX can also alter the metabolism of the cell. One group found that DOX at 100 ng/mL to 5 μg/mL, which is the range commonly used, can cause the cells to produce more lactate and less oxygen consumption; high production of lactate or lactic acid can be toxic to the cells, but this change affects only some cell types (Ahler et al. 2013). We found that at 0.3 μg/mL, DOX does not cause major changes to global gene expression

patterns and the few genes altered are not involved in glycolysis; thus we used this concentration to turn on HNF4α expression in the HCT116 inducible cell lines and perform RNA-Seq and ChIP-Seq experiments. We also used the same concentration to turn on HNF4α expression in the mES inducible cell lines and perform RNA-Seq. RNA-Seq analysis of the HCT116 PL, α2 and α8 lines, at 0.0 μg/mL, we found that the parental and two daughter lines are slightly different from one another; this is to be expected since these cells underwent multiple transfection and selection processes. Despite the inherent differences, the induction of HNF4α still produced a greater change in gene expression in both the α2 and α8 lines than that observed in the absence of DOX (see Chapter 3). Therefore, we conclude that these HCT116 and mES inducible lines will provide important information on the role of the HNF4α2 and HNF4α8 in human colon cancer and early development.

Fig. 2.1

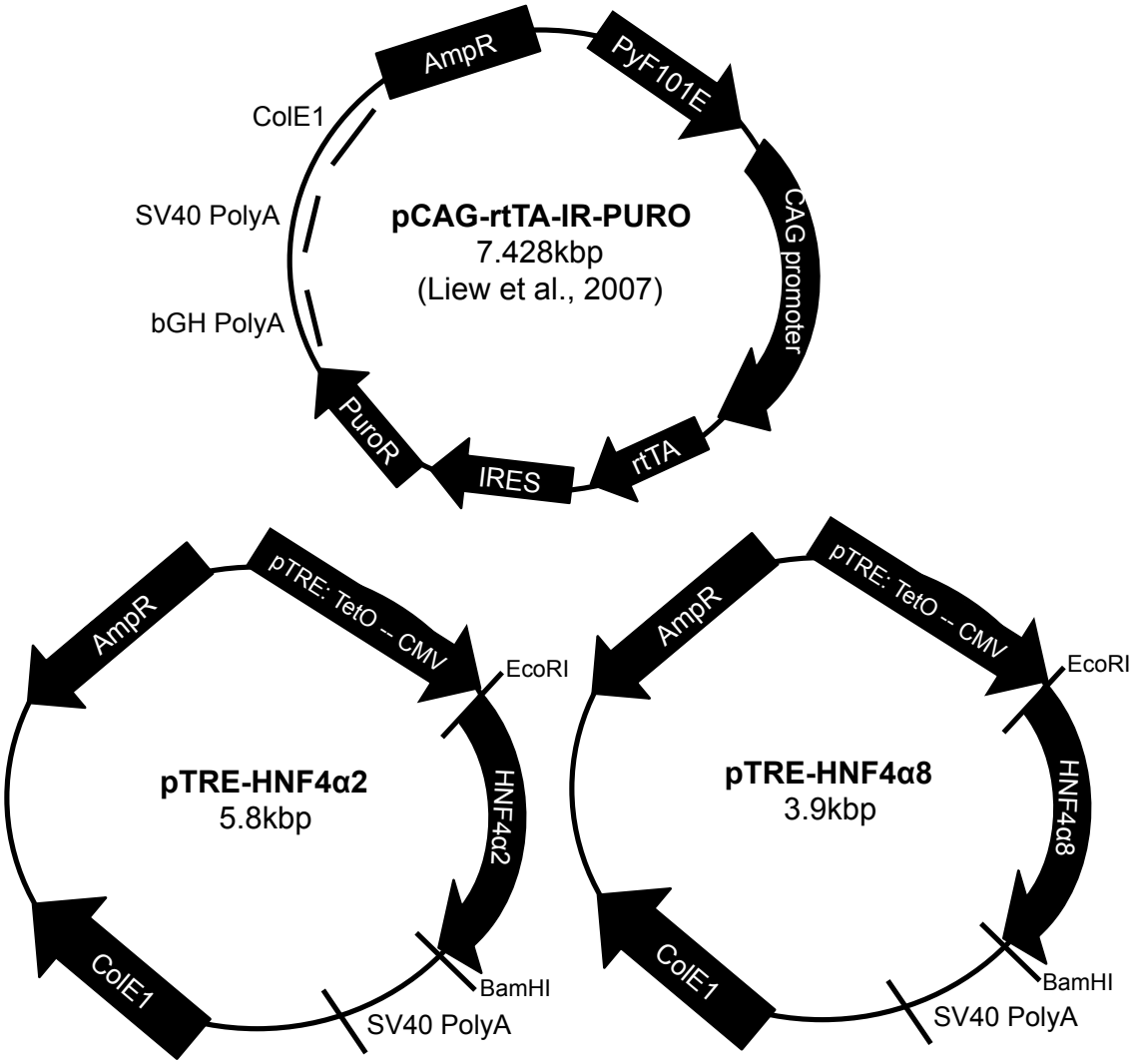

AmpR
PyF101E
ColE1
SV40 PolyA
pCAG-rtTA-IR-PURO
7.428kbp
(Liew et al., 2007)
CAG promoter
bGH PolyA
PuroR
IRES
rtTA
AmpR
pTRE: TetO -- CMV
EcoRI
pTRE-HNF4α2
5.8kbp
HNF4α2
ColE1
BamHI
SV40 PolyA
AmpR
pTRE: TetO -- CMV
EcoRI
pTRE-HNF4α8
3.9kbp
HNF4α8
ColE1
BamHI
SV40 PolyA

Figure 2.1 Maps of the Tet-On inducible vectors.

Generation of the pCAG-rtTA-IR-PURO plasmid is described in Liew et al. 2007. The rtTA and puromycin resistance (PuroR) genes are under the control of the CAG promoter, which consists of the CAGG promoter that is linked to an upstream PYF101E and a downstream IRES. The pTRE-Tight vector that was used to insert the human HNF4α2 or HNF4α8 construct was purchased from Clontech. The fusion EGFP-HNF4α2 and EGFP-HNF4α8 were also inserted into the pTRE-Tight vector in a similar fashion. The TRE promoter, which includes several TetO promoters and a downstream CMV promoter, drives the expression of the inserted cDNA between two restriction sites (EcoRI and BamHI). The pTRE-RFP (pTRE-Red Max C) from Dr. Chee-Gee Liew contains a red fluorescence protein (RFP) (Xia et al. 2008).

Dr. Karthikeyani Chellappa designed the primers and cloning procedures to generate these plasmids: pTRE-HNF4α2, pTRE-HNF4α8, pTRE-EGFP-HNF4α2, pTRE-EGFP-HNF4α8.

Fig. 2.2

Flow Chart of the Tet-On Inducible Expression Constructs

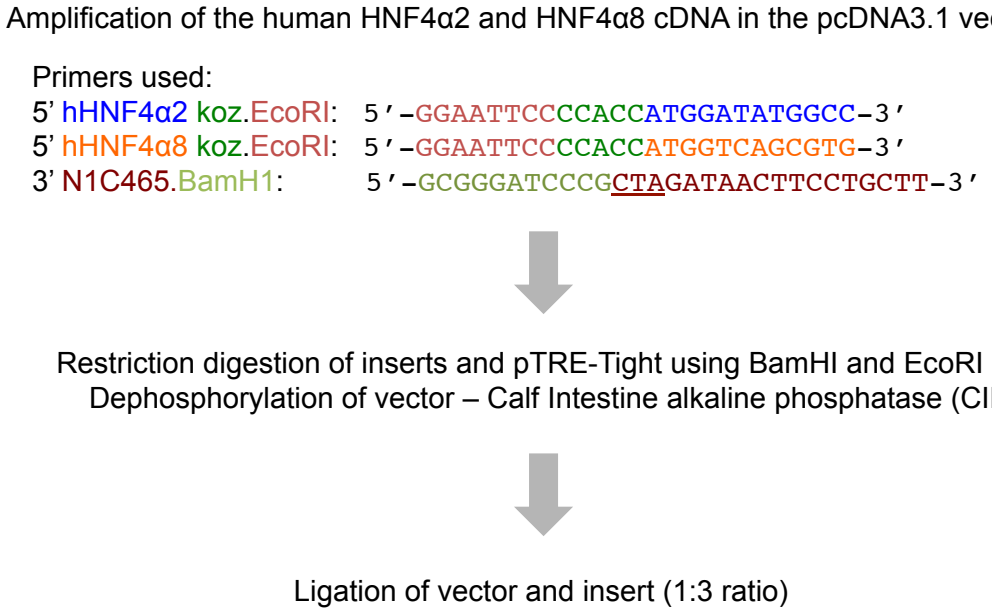


Figure 2.2 Workflow to construct Tet-On inducible human HNF4α2 and HNF4α8 vectors.

See Materials and Methods (M&M) and results for details. Restriction enzyme sites and consensus Kozak's sequences used for cloning are color-coded, as are the cDNA sequence of the N-terminus of the human HNF4α2 (blue), human HNF4α8 (orange), and the C-terminal region plus stop codon (underlined) common to both cDNAs (brown). The cDNAs were first amplified using primers shown above. Second, both the amplified inserts and pTRE-Tight vector were digested with BamHI and EcoRI restriction enzymes. The digested pTRE-Tight vector was then dephosphorylated using calf intestine alkaline phosphatase (CIP). Both the vector and insert were ligated at a 1 to 3 ratio.

Fig. 2.3

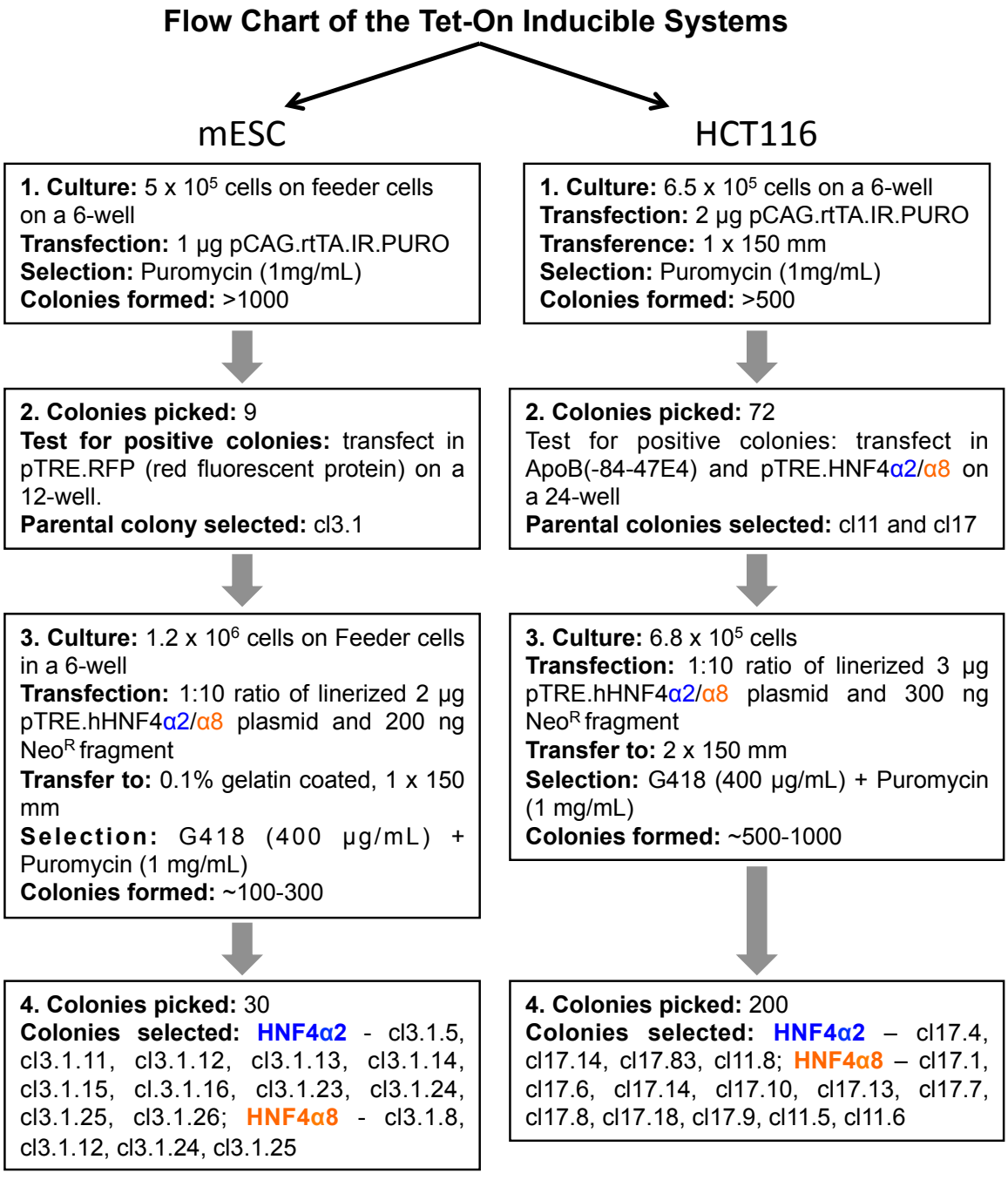
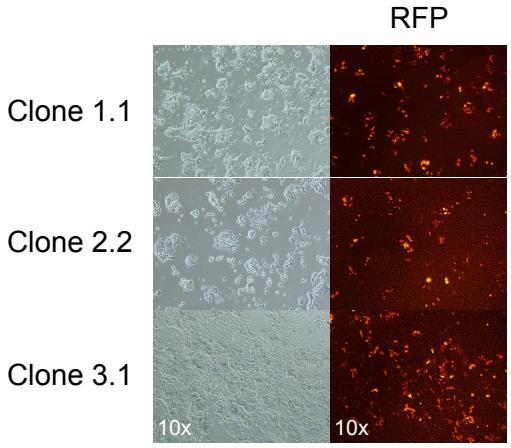

Flow Chart of the Tet-On Inducible Systems
mESC
HCT116
1. Culture: 5 x 10^5 cells on feeder cells on a 6-well
Transfection: 1 µg pCAG.rtTA.IR.PURO
Selection: Puromycin (1mg/mL)
Colonies formed: >1000
1. Culture: 6.5 x 10^5 cells on a 6-well
Transfection: 2 µg pCAG.rtTA.IR.PURO
Transference: 1 x 150 mm
Selection: Puromycin (1mg/mL)
Colonies formed: >500
2. Colonies picked: 9
Test for positive colonies: transfect in pTRE.RFP (red fluorescent protein) on a 12-well.
Parental colony selected: cl3.1
2. Colonies picked: 72
Test for positive colonies: transfect in ApoB(-84-47E4) and pTRE.HNF4α2/α8 on a 24-well
Parental colonies selected: cl11 and cl17
3. Culture: 1.2 x 10^6 cells on Feeder cells in a 6-well
Transfection: 1:10 ratio of linerized 2 µg pTRE.hHNF4α2/α8 plasmid and 200 ng Neo^R fragment
Transfer to: 0.1% gelatin coated, 1 x 150 mm
Selection: G418 (400 µg/mL) + Puromycin (1 mg/mL)
Colonies formed: ~100-300
3. Culture: 6.8 x 10^5 cells
Transfection: 1:10 ratio of linerized 3 µg pTRE.hHNF4α2/α8 plasmid and 300 ng Neo^R fragment
Transfer to: 2 x 150 mm
Selection: G418 (400 µg/mL) + Puromycin (1 mg/mL)
Colonies formed: ~500-1000
4. Colonies picked: 30
Colonies selected: HNF4α2 - cl3.1.5, cl3.1.11, cl3.1.12, cl3.1.13, cl3.1.14, cl3.1.15, cl.3.1.16, cl3.1.23, cl3.1.24, cl3.1.25, cl3.1.26; HNF4α8 - cl3.1.8, cl3.1.12, cl3.1.24, cl3.1.25
4. Colonies picked: 200
Colonies selected: HNF4α2 – cl17.4, cl17.14, cl17.83, cl11.8; HNF4α8 – cl17.1, cl17.6, cl17.14, cl17.10, cl17.13, cl17.7, cl17.8, cl17.18, cl17.9, cl11.5, cl11.6

Figure 2.3 Schematic of the workflow to generate Tet-On Inducible in mES and HCT116 cell lines.

A two-step cloning procedure was used to generate stable, inducible clones in mES (left) and HCT116 (right). See M&M and results for detail discussion of the procedure.

Fig. 2.4

A **mESC Parental Clones**



B **HCT116 Parental Clones**

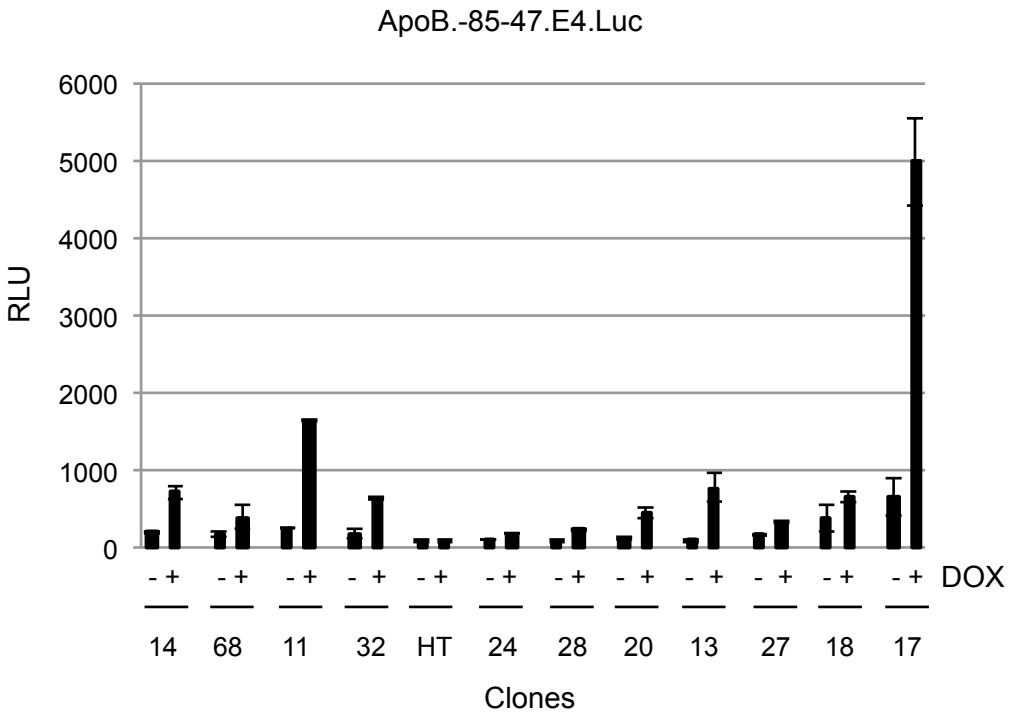


Figure 2.4 Testing for rtTA activity in the mES and HCT116 parental lines.

(A) mES clones transfected with 1 µg of pTRE-RFP and induced with DOX (1 µg/mL) for 24 h. (B) Cells were seeded at 1.76 x 10^5 cells per well of a 24-well. The next day, cells were transfected with 150 ng ApoB.-85-47.E4.Luc and 1.0 ng of pTRE-HNF4α2 expression vector, induced with 1 µg/mL DOX for 24 h, and luciferase assay was performed. Bar graphs represent RLU means±SD of triplicate samples from one independent study. Control: HT, HCT116 original line (no expression of the rtTA gene).

Fig. 2.5

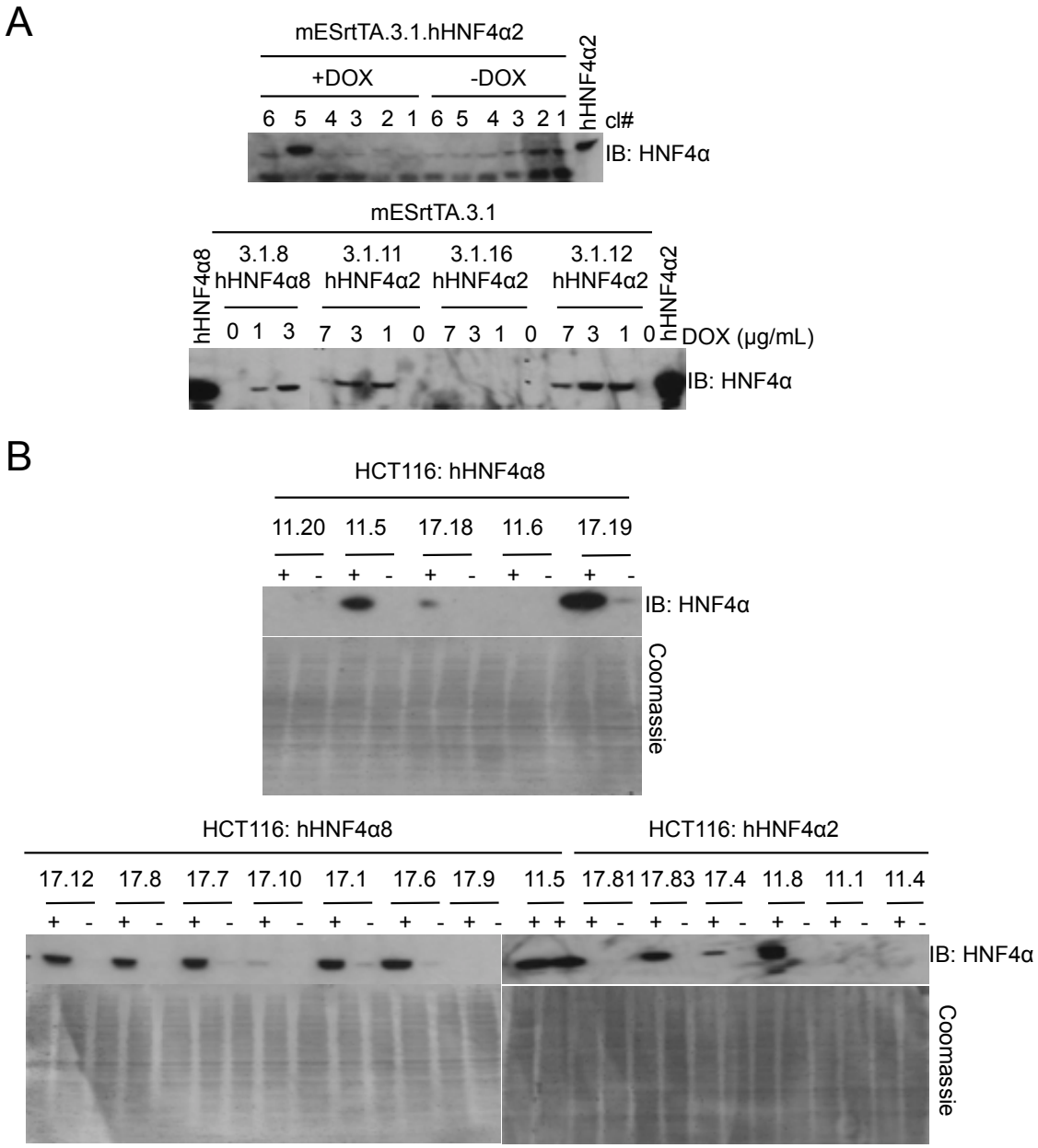

A
mESrtTA.3.1.hHNF4α2
+DOX
-DOX
hHNF4α2
6 5 4 3 2 1 6 5 4 3 2 1 cl#
IB: HNF4α
mESrtTA.3.1
hHNF4α8
3.1.8 hHNF4α8
3.1.11 hHNF4α2
3.1.16 hHNF4α2
3.1.12 hHNF4α2
hHNF4α2
0 1 3 7 3 1 0 7 3 1 0 7 3 1 0 DOX (μg/mL)
IB: HNF4α
B
HCT116: hHNF4α8
11.20 11.5 17.18 11.6 17.19
+ - + - + - + - + -
IB: HNF4α
Coomassie
HCT116: hHNF4α8
HCT116: hHNF4α2
17.12 17.8 17.7 17.10 17.1 17.6 17.9 11.5 17.81 17.83 17.4 11.8 11.1 11.4
+ - + - + - + - + - + - + - + + + - + - + - + - + - + -
IB: HNF4α
Coomassie

Figure 2.5 Identification of positive mES and HCT116 clones after puromycin and G418 selection.

Immunoblot (IB) with P1/P2 antibody of WCE (20 μg) from DOX-inducible mES (A) and HCT116 (B) clones to identify clones positive for HNF4α2 or HNF4α8 expression. These are examples of some of the IB analysis. We selected clones 3.1.5 (HNF4α2), 3.1.8 (HNF4α8), cl11.8 (HNF4α2), cl11.5 (HNF4α8), cl17.83 (HNF4α2), and cl17.1 (HNF4α8) since these clones do not show background expression of HNF4α in the absence of DOX but express moderate to high level of HNF4α in the presence of DOX.

Fig. 2.6

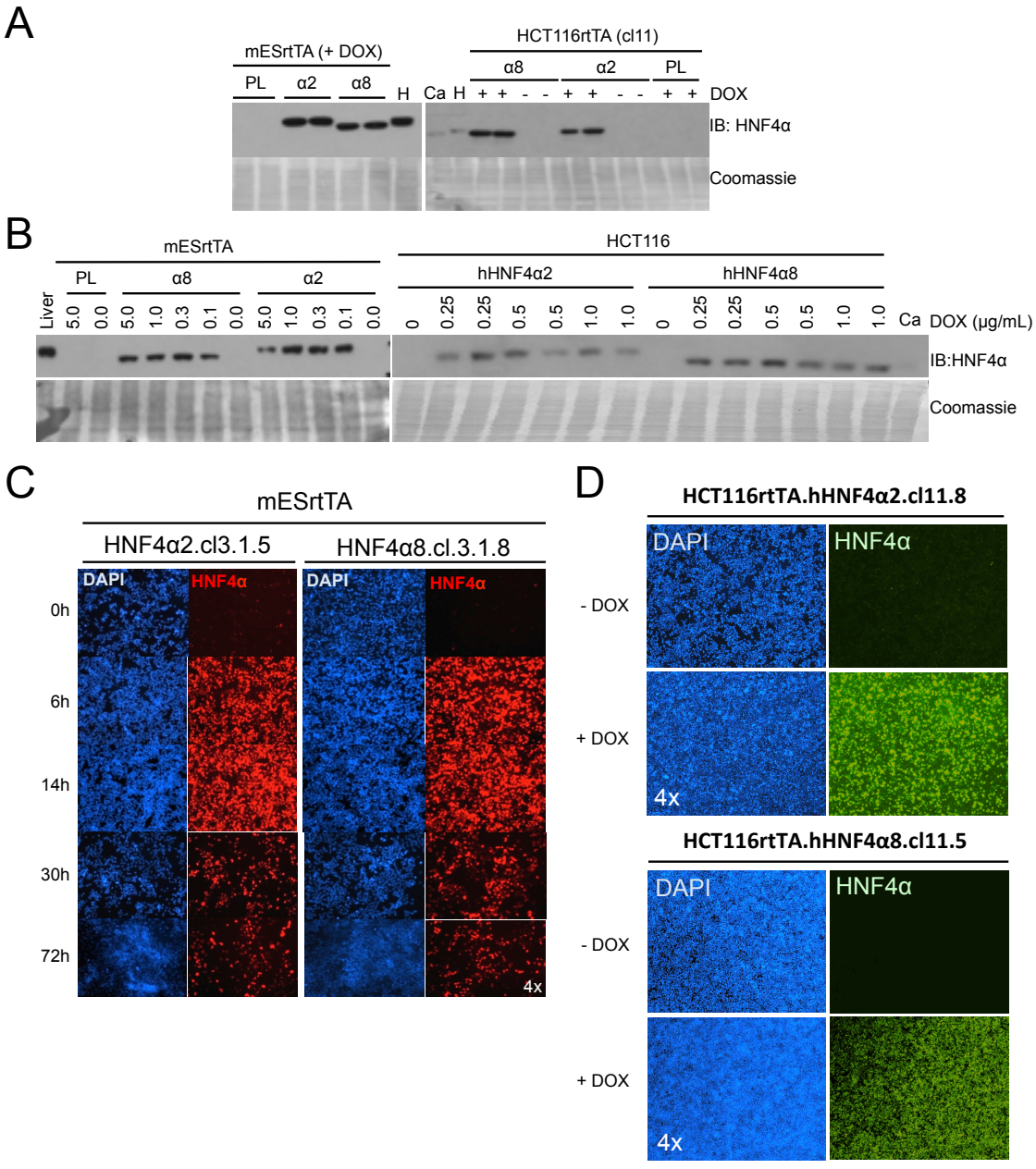

A
mESrtTA (+ DOX)
PL
α2
α8
H
HCT116rtTA (cl11)
α8
α2
PL
Ca H + + - - + + - - + + DOX
IB: HNF4α
Coomassie
B
mESrtTA
PL
α8
α2
Liver
5.0 0.0 5.0 1.0 0.3 0.1 0.0 5.0 1.0 0.3 0.1 0.0
HCT116
hHNF4α2
hHNF4α8
0 0.25 0.25 0.5 0.5 1.0 1.0 0 0.25 0.25 0.5 0.5 1.0 1.0
Ca DOX (μg/mL)
IB:HNF4α
Coomassie
C
mESrtTA
HNF4α2.cl3.1.5
HNF4α8.cl.3.1.8
DAPI
HNF4α
0h
6h
14h
30h
72h
4x
D
HCT116rtTA.hHNF4α2.cl11.8
DAPI
HNF4α
- DOX
+ DOX
4x
HCT116rtTA.hHNF4α8.cl11.5
DAPI
HNF4α
- DOX
+ DOX
4x

Figure 2.6 Characterization of the inducible mES and HCT116 lines.

(A and B) IB analyses of HNF4α2 and HNF4α8 in (~20 µg total protein) NE (A) and WCE (B) of inducible mES (cls 3.1.5 and 3.1.8) and HCT116 (cls 11.5 and 11.8) cells, using the anti-P1/P2 antibody. Cells were induced with 0.3 µg/mL of DOX for 24 h (HCT116) or 56 h (mES) (A) or 48 h (both) (B). Controls: H, HepG2; Ca, CaCO2; liver, mouse liver. (C and D) Immunofluorescence (IF) of HNF4α in the lines (mES: 3.1.5 and 3.1.8; HCT116: 11.5 and 11.8) treated with 0.3 µg/mL for 0, 6, 14, 24, 30, and 72 h.

Fig. 2.7

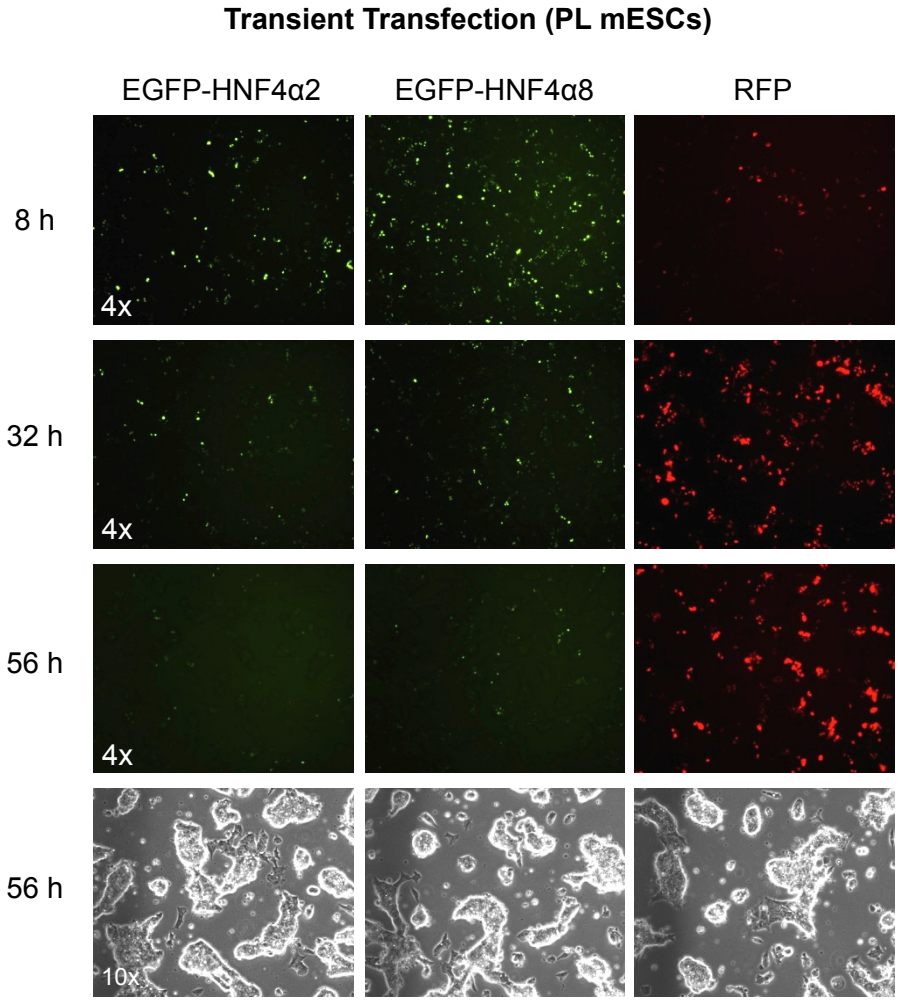


Figure 2.7 Expression of HNF4α2 and HNF4α8 but not RFP decrease over time in transient transfection.

Transient transfection of pTRE-EGFP-HNF4α2, pTRE-EGFP-HNF4α8, or pTRE-RFP in parental (PL) of the Tet-On inducible mES cell line, induced with 1 μg/mL DOX, and viewed under a confocal microscope at different time points (magnification, 4x and 10 x).

Fig. 2.8

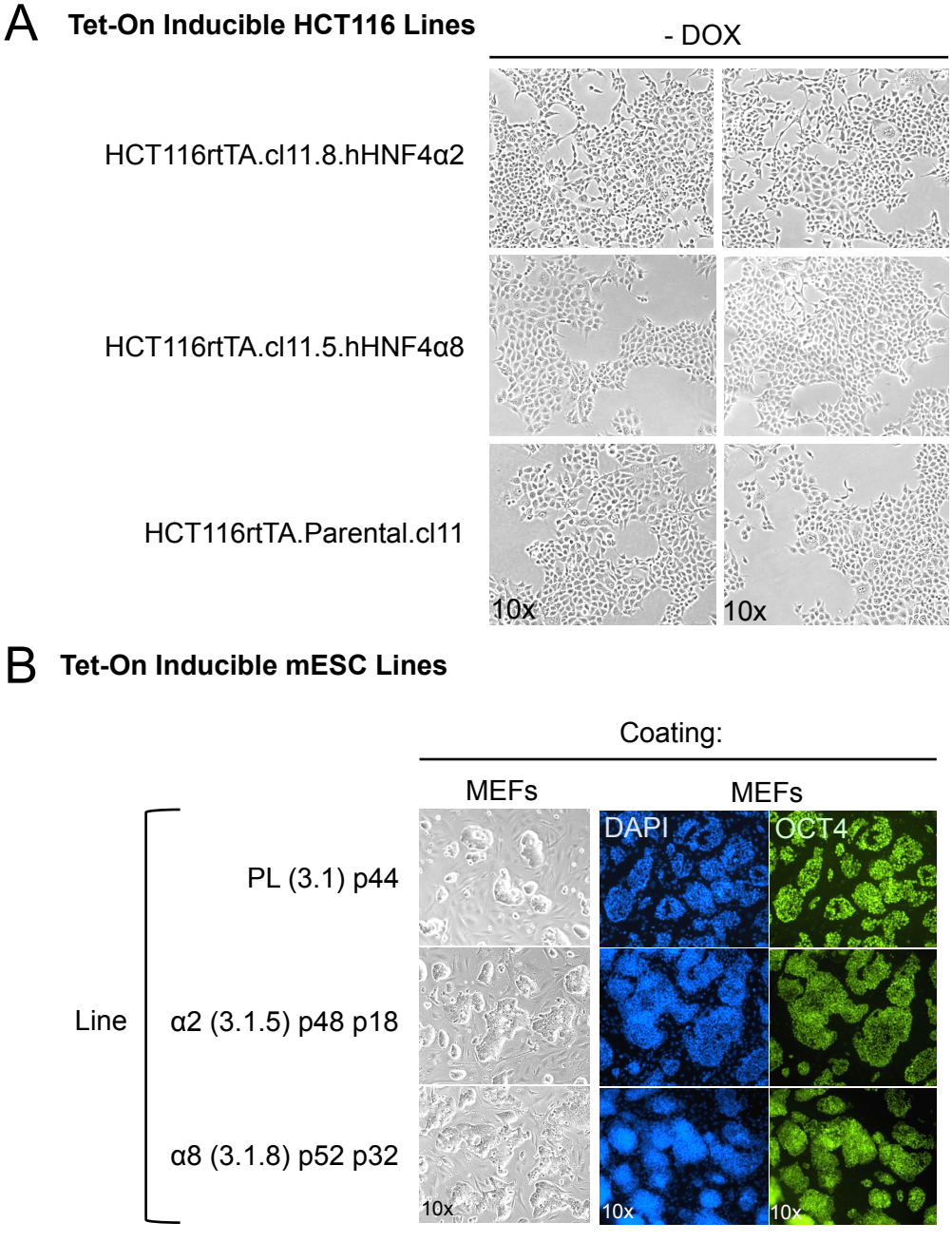

A
Tet-On Inducible HCT116 Lines
- DOX
HCT116rtTA.cl11.8.hHNF4α2
HCT116rtTA.cl11.5.hHNF4α8
HCT116rtTA.Parental.cl11
10x
10x
B
Tet-On Inducible mESC Lines
Coating:
MEFs
MEFs
DAPI
OCT4
Line
PL (3.1) p44
α2 (3.1.5) p48 p18
α8 (3.1.8) p52 p32
10x
10x
10x

Figure 2.8 Tet-On inducible lines have little or no morphological change in the absence of DOX.

(A) Phase contrast images of HCT116 inducible lines in the absence of DOX that have been cultured for 48 h after seeding (Magnification 10x). (B) Phase contrast and confocal imaging of DAPI and OCT4 of mES inducible lines in the absence of DOX (Magnification 10x).

Table 2.1 Genes dysregulated in HCT116 inducible lines: Parental (PL) (+/-DOX) and in α2 and α8 lines versus PL (-DOX)

PL 0.3 / PL 0.0	
FC	**Gene**
1.6	*ECM1*
1.5	*WFDC3*
-1.5	*MORN1*
-1.5	*ASB18, GBX2*
-1.6	*CYP24A1*
-2.0	*DHRS2*

α2 0.0 / PL 0.0	
FC	**Gene**
14.1	*IZUMO3*
6.0	*TMEM200A*
2.6	*IFIT2*
2.5	*OASL*
2.2	*IFIT3*
2.1	*KLRC4-KLRK1,KLRK1*
2.1	*KRTAP2-3,KRTAP2-4*
2.0	*GPR110*
-2.0	*CTHRC1*
-2.1	*FSD1*
-2.5	*MAGI1*
-2.5	*S100A4*
-3.5	*SPTSSB*
-7.0	*SLC43A3*

α8 0.0 / PL 0.0	
FC	**Gene**
6.7	*KRT23*
5.5	*KRTAP3-1*
4.9	*SEPP1*
3.9	*TNFSF18*
3.5	*APOBEC3G*
3.5	*BST2*
2.9	*SPTSSB*
2.8	*IFI6*
2.6	*IFITM1*
2.6	*TSPAN1*
2.4	*SLC16A6*
2.2	*SUSD2*
2.2	*TFAP2A*
2.2	*ID2*
2.0	*LINC00173*
2.0	*UTRN*
2.0	*IFI35*
2.0	*NR2F1*
-2.0	*C6orf141*
-2.0	*ECM1*
-2.0	*HIPK2*
-2.0	*STK39*
-2.0	*ANO1*
-2.1	*FN1*
-2.2	*TRIML2*
-2.2	*TIMP2*
-2.2	*ANTXR2*
-2.3	*HS3ST1*
-2.3	*WNT7B*
-2.3	*RPS6KA3*
-2.3	*CD24*
-2.4	*FRAS1*
-2.5	*FABP6*
-2.5	*HES1*
-2.6	*ACSL5*
-2.6	*S100A14*
-2.9	*EFNB2*
-2.9	*EMP1*
-3.4	*VGF*

Note:
Fold change (FC) non-log
Parental (PL), HNF4α2 line (α2), and HNF4α8 (α8) lines
0.0 μg/mL – no doxycycline (DOX)
0.3 μg/mL – DOX
Negative FC – genes downregulated
Positive FC – genes upregulated

Fig. 2.9

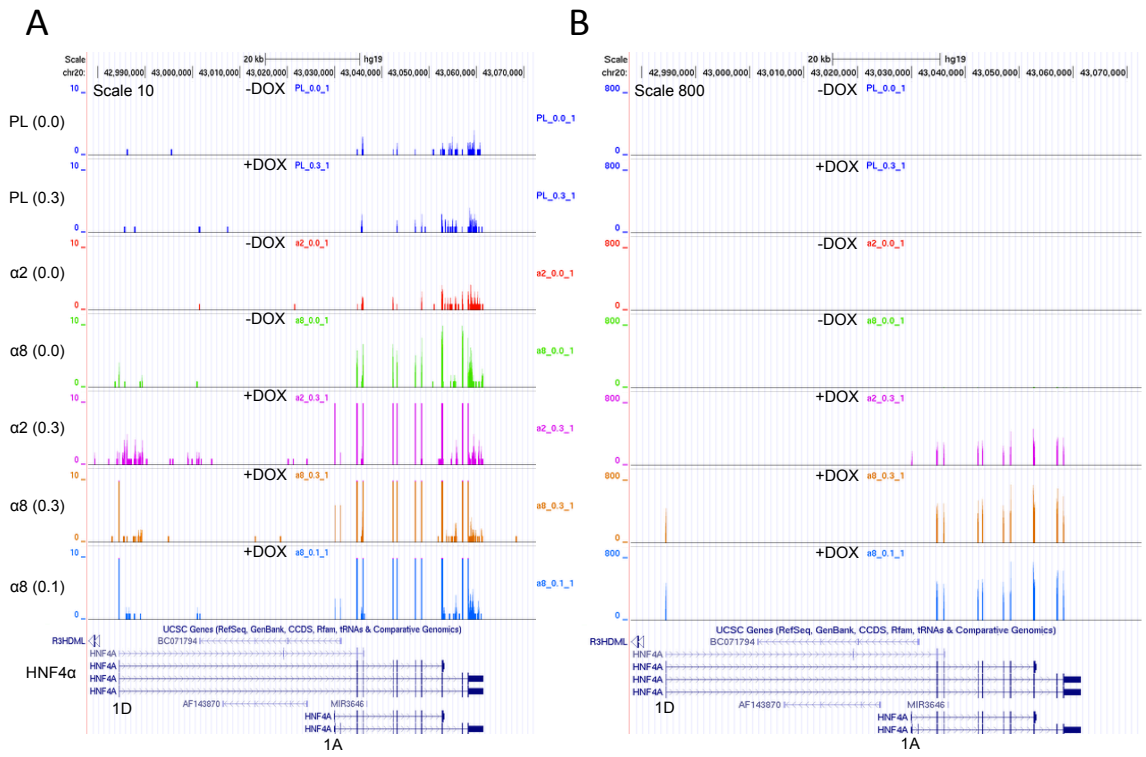


Figure 2.9 HNF4α mRNA levels in HCT116 inducible lines (PL, α2, and α8).

Snapshot of HNF4α transcript level for PL, α2, and α8 lines (-/+DOX) from the UCSC genome browser (Kent et al. 2002) set at scale 10 (A) or scale 800 (B). Below shows the exons and introns of HNF4α. The 1D and 1A exons encode for the different A/B domains, truncated and full-length A/B domains, respectively. HNF4α2 contains the 1A exon and HNF4α8 contains the 1D exon.

Table 2.2 HNF4α mRNA expression

Line	DOX	Absolute value	Which exon (1D or 1A) is expressed in the line?
PL	-	5	--
	+	5	--
HNF4α2	-	5	--
	+	400	1A; very little 1D (zoom in to the scale of 5-10)
HNF4α8	-	8-10	1D
	+	800	1D; very little 1A (zoom in to the scale of 5-10)

Note: absolute values are estimation from the UCSC genome browser of HNF4α as shown in Fig. 2.9. The genome browser was zoomed in (1kb to 0.2kb) to determine if there were any signal from the 1D and 1A positions in the PL, α2, and α8 (-/+DOX) conditions.

References

Ahler E, Sullivan WJ, Cass A, Braas D, York AG, Bensinger SJ, Graeber TG, Christofk HR. 2013. Doxycycline alters metabolism and proliferation of human cell lines. *PloS one* **8**: e64561.

Aladjem MI, Spike BT, Rodewald LW, Hope TJ, Klemm M, Jaenisch R, Wahl GM. 1998. ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. *Current biology : CB* **8**: 145-155.

Attal J, Theron MC, Houdebine LM. 1999. The optimal use of IRES (internal ribosome entry site) in expression vectors. *Genetic analysis : biomolecular engineering* **15**: 161-165.

Chartier FL, Bossu JP, Laudet V, Fruchart JC, Laine B. 1994. Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. *Gene* **147**: 269-272.

Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CL, Chan C, Dent OF, Clarke SJ, Robertson GR et al. 2012. Src tyrosine kinase phosphorylation of nuclear receptor HNF4alpha correlates with isoform-specific loss of HNF4alpha in human colon cancer. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 2302-2307.

Eeckhoute J, Formstecher P, Laine B. 2001. Maturity-onset diabetes of the young Type 1 (MODY1)-associated mutations R154X and E276Q in hepatocyte nuclear factor 4alpha (HNF4alpha) gene impair recruitment of p300, a key transcriptional co-activator. *Molecular endocrinology* **15**: 1200-1210.

Fife RS, Rougraff BT, Proctor C, Sledge GW, Jr. 1997. Inhibition of proliferation and induction of apoptosis by doxycycline in cultured human osteosarcoma cells. *The Journal of laboratory and clinical medicine* **130**: 530-534.

Freundlieb S, Schirra-Muller C, Bujard H. 1999. A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *The journal of gene medicine* **1**: 4-12.

Garfield AS. 2010. Derivation of primary mouse embryonic fibroblast (PMEF) cultures. *Methods in molecular biology* **633**: 19-27.

Gossen M, Bujard H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 5547-5551.

Gossen M, Freundlieb S, Bender G, Muller G, Hillen W, Bujard H. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**: 1766-1769.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome research* **12**: 996-1006.

Liew CG, Draper JS, Walsh J, Moore H, Andrews PW. 2007. Transient and stable transgene expression in human embryonic stem cells. *Stem cells* **25**: 1521-1528.

Maeda Y, Hwang-Verslues WW, Wei G, Fukazawa T, Durbin ML, Owen LB, Liu X, Sladek FM. 2006. Tumour suppressor p53 down-regulates the expression of the human hepatocyte nuclear factor 4alpha (HNF4alpha) gene. *The Biochemical journal* **400**: 303-313.

Maeda Y, Seidel SD, Wei G, Liu X, Sladek FM. 2002. Repression of hepatocyte nuclear factor 4alpha tumor suppressor p53: involvement of the ligand-binding domain and histone deacetylase activity. *Molecular endocrinology* **16**: 402-410.

Nelson JA, Gnann JW, Jr., Ghazal P. 1990. Regulation and tissue-specific expression of human cytomegalovirus. *Current topics in microbiology and immunology* **154**: 75-100.

Niwa H, Yamamura K, Miyazaki J. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**: 193-199.

Qin H, Yu T, Qing T, Liu Y, Zhao Y, Cai J, Li J, Song Z, Qu X, Zhou P et al. 2007. Regulation of apoptosis and differentiation by p53 in human embryonic stem cells. *The Journal of biological chemistry* **282**: 5842-5852.

Solozobova V, Blattner C. 2011. p53 in stem cells. *World journal of biological chemistry* **2**: 202-214.

Son K, Fujioka S, Iida T, Furukawa K, Fujita T, Yamada H, Chiao PJ, Yanaga K. 2009. Doxycycline induces apoptosis in PANC-1 pancreatic cancer cells. *Anticancer research* **29**: 3995-4003.

Welman A, Cawthorne C, Ponce-Perez L, Barraclough J, Danson S, Murray S, Cummings J, Allen TD, Dive C. 2006. Increases in c-Src expression level and activity do not promote the growth of human colorectal carcinoma cells in vitro and in vivo. *Neoplasia* **8**: 905-916.

Xia X, Ayala M, Thiede BR, Zhang SC. 2008. In vitro- and in vivo-induced transgene expression in human embryonic stem cells and derivatives. *Stem cells* **26**: 525-533.

Zhu Z, Zheng T, Lee CG, Homer RJ, Elias JA. 2002. Tetracycline-controlled transcriptional regulation systems: advances and application in transgenic animal modeling. *Seminars in cell & developmental biology* **13**: 121-128.

Chapter 3

Interplay between HNF4α and TCF4 (*TCF7L2*) in human colon cancer

This chapter has been prepared as a manuscript for publication.

The title will be: Interplay between HNF4α, TCF4, and AP-1 in human colon cancer

Authors will be: Linh M. Vuong, Karthikeyani Chellappa, Joseph M. Dhahbi, Jonathan Deans, Bin Fang, Eugene Bolotin, Nina Titova, Nathan P. Hoverter, Marian L. Waterman and Frances M. Sladek

Contributions by the other authors are noted in figure legends (Fig.s 3.2, 3.3, 3.5, 3.6, 3.7, 3.8, 3.9).

Abstract

HNF4α is a nuclear receptor and master regulator of hepatocyte gene expression with a merging role in the gut. It is tumor suppressive in the liver but amplified in colon cancer, suggesting it might also be oncogenic. To determine whether this ambiguity in HNF4α function is due to the fact that the *HNF4A* gene has two promoters (P1 and P2), both of which are expressed in normal colon, we generated Tet-On inducible human colon cancer (HCT116) cell lines that express either the P1- (HNF4α2) or the P2-driven (HNF4α8) isoform and analyzed them for tumor growth and global changes in gene expression (RNA-Seq, ChIP-Seq). The results show that while HNF4α2 acts as a tumor suppressor in HCT116 cells, HNF4α8 does not, and that the HNF4α isoforms regulate the expression of distinct sets of genes and recruit, co-localize and compete with TCF4 in the Wnt/β-catenin for chromatin occupancy at CTTTG motifs as well as AP-1 motifs (TGAxTCA) in a distinct fashion. Finally, protein binding microarrays (PBMs) show that SNPs in sites bound by both HNF4α and TCF4 can alter binding affinity in vitro, suggesting that they could play a role in cancer susceptibility in vivo. Thus, the HNF4α isoforms play distinct roles in colon cancer, which could be due to differential interaction with the Wnt/β-catenin/TCF/AP-1 pathway.

Introduction

Hepatocyte nuclear factor 4alpha (HNF4α) (NR2A1) is a highly conserved member of the nuclear receptor superfamily, and is found in all metazoans from sponges to humans (Sladek et al. 1990; Sladek 2011). In mammals, HNF4α is best known as a master regulator of tissue-specific gene expression in the adult liver (Hayhurst et al. 2001; Odom et al. 2004). HNF4α binds specific DNA sequences in regulatory regions of

target genes as a homodimer and regulates expression of genes involved in metabolism, homeostasis, differentiation, and immune response (Sladek 2001; Odom et al. 2004; Bolotin et al. 2010). It also plays a role in early development (Chen et al. 1994), as well as in the adult kidney, pancreas and gut (Garrison et al. 2006; Ahn et al. 2008; Babeu et al. 2009; Cattin et al. 2009; Darsigny et al. 2009). Mutations in the *HNF4A* gene or HNF4α binding sites have been linked to various human diseases, including an inherited form of type 2 diabetes (maturity onset diabetes of the young 1, MODY1) and hemophilia (Yamagata et al. 1996; Sladek 2001). Recently, HNF4α was shown to be involved in colon cancer but its precise role has not yet been elucidated (Garrison et al. 2006; Babeu et al. 2009; Cattin et al. 2009; Chellappa et al. 2012).

Several splice variants of HNF4α are generated through the utilization of two alternative promoters (proximal P1 and distal P2) and two distinct 3' splicing events (Taraviras et al. 1994; Furuta et al. 1997). Originally, P1-driven HNF4α1, which includes the full length N-terminal A/B domain, was cloned from adult rat liver (Sladek et al. 1990), while the P2-driven HNF4α7 with a truncated N-terminal domain was cloned from murine embryonic cell line (Nakhei et al. 1998) (Fig 3.1A). Alternative splicing at the C-terminal end of HNF4α1 and HNF4α7 produces isoforms HNF4α2-α3 and HNF4α8-α9, respectively. HNF4α2 and HNF4α8 appear to be the predominant forms in most tissues (Harries et al. 2008). The promoter-driven HNF4α isoforms exhibit tissue-specific expression patterns: the P1-driven HNF4α2 is expressed in the fetal and adult liver and kidney, whereas the P2-driven HNF4α8 is expressed in the fetal but not the adult liver and adult stomach and pancreas; both P1- and P2-HNF4α are expressed in the large and small intestines (Nakhei et al. 1998; Briancon and Weiss 2006; Tanaka et al. 2006). The HNF4α gene structure, promoter sequences and expression patterns are highly

conserved between human and mouse (Sladek 2011), suggesting that the P1- and P2-driven HNF4α play important and yet distinct functional roles. Indeed, exon-swap mice that express just a single HNF4α N-terminal isoform show subtle yet significant metabolic differences in unstressed animals (Briancon and Weiss 2006).

What is known about the role of HNF4α in cancer comes largely from studies on the P1- variants HNF4α1/2. Recently, it has become clear that P1-HNF4α acts as a tumor suppressor in the liver (Ning et al. 2010), inhibiting hepatocyte proliferation and inflammation (Hatziapostolou et al. 2011; Bonzo et al. 2012; Walesky et al. 2013; Saha K. Supriya 2014). Several key players in proliferation, including p53, c-Myc, T-cell factor (TCF) 4, lymphoid enhancer factor 1 (LEF1), and cyclin D1, have also been shown to physically interact and antagonize P1-HNF4α (Maeda et al. 2002; Maeda et al. 2006; Hwang-Verslues and Sladek 2008; Cattin et al. 2009; Colletti et al. 2009; Hanse et al. 2012; Sladek 2012).

Although no specific functional role has been established for P2-HNF4α in cancer, there is evidence suggesting that it may not have the potent tumor suppressive characteristics of P1-HNF4α. HNF4α7 was identified originally in the undifferentiated, pluripotent embryonic carcinoma cell line F9 (Nakhei et al. 1998). Immunohistochemical staining for P1 and P2-HNF4α in patients diagnosed with liver, colon, and stomach cancers showed that there is a dysregulation of the HNF4α isoforms, with P2-HNF4α typically being expressed at higher levels than P1-HNF4α in cancer samples (Tanaka et al. 2006; Oshima et al. 2007; Takano et al. 2009). We have also recently shown that in a cohort of 450 human colon cancer patients, there is a loss of nuclear P1-HNF4α in 80% of the samples (Chellappa et al. 2012), which we attributed to activation of Src tyrosine kinase, which preferentially phosphorylates P1 but not P2-driven HNF4α, leading to non-

nuclear localization as well as decreased transactivation function and protein stability of P1-HNF4a but not P2-HNF4α in human colon cancer cells (Chellappa et al. 2012).

In contrast, The Cancer Genome Atlas (TCGA) recently identified the region encompassing *HNF4A* (20q13.12) as being one of the most significantly amplified loci in over 255 human colon cancers (Cancer Genome Atlas 2012), and found an overexpression of the HNF4α protein in a subset of those samples (Zhang et al. 2014). While these findings suggest that HNF4α may act as an oncogene, the relative contribution of the different HNF4α isoforms was not distinguished in these studies.

There is also an increasing number of reports showing a cross-talk between HNF4α and the Wnt/β-catenin/TCF4 signaling pathway in liver zonation, hepatocellular carcinoma (HCC) development, and colorectal cancer. The Wnt/β-catenin/TCF4 signaling pathway is well known to promote proliferation while HNF4α, specifically the P1-HNF4α, drives differentiation. Physical interactions between HNF4α and TCF4 have been reported using immunoprecipitation (co-IP) and glutathione S-transferase (GST)-pull down assays as well as on isolated promoters (Cattin et al. 2009; Yang et al. 2013; Gougelet et al. 2014). TCF4 ChIP-on-chip and ChIP-Seq peaks have been found to be enriched for HNF4α binding motifs, suggesting a potential coregulation by these two transcription factors (TFs) (Hatzis et al. 2008; Frietze et al. 2012; Gougelet et al. 2014). Nevertheless, the interplay between TCF4 and the HNF4α isoforms, in colorectal cancer has not been examined in depth.

To distinguish the roles of P1- and P2-HNF4α in colon cancer, and to examine their interaction with the Wnt/β-catenin/TCF pathway, we established an inducible system in the human colon cancer cell line HCT116. Functional analysis of the lines, which express either P1-HNF4α2 or P2-HNF4α8 under control of doxycycline (DOX),

indicate that HNF4α2 reduces tumor growth more effectively than HNF4α8 in vivo, while HNF4α8 increases the invasive index in vitro. RNA-Seq and ChIP-Seq analyses revealed differences (as well as commonalities) in gene expression and binding locations between HNF4α2 and HNF4α8. TCF4 ChIP-Seq in the absence and presence of the HNF4α isoforms revealed recruitment by and co-localization with HNF4α on a common CTTTG core, as well as competition with HNF4α2, but not HNF4α8, via the AP-1 complex in vivo. Finally, common, as well as unique, TCF4 and HNF4α binding motifs are identified using protein binding microarrays (PBMs), which also showed that SNPs in TCF4/HNF4α binding sites can affect DNA binding in vitro often in different ways, revealing a remarkable complexity and specificity to DNA binding. Taken together, our results indicate that there are subtle yet important differences in the HNF4α isoforms, some of which involve distinct interactions with the Wnt/β-catenin/TCF pathway and which could be affected by individual genetic differences.

Materials and Methods

Plasmid Constructs

The full length human HNF4α2 (NM_000457) and HNF4α8 (NM_175914.3) cDNAs in pcDNA3.1 were gifts from Dr. Christophe Rachez at Pasteur Institute, Paris, France (Chartier et al. 1994; Eeckhoute et al. 2003). The pcDNA4/TO.Flag.dnTCF1, pcDNA3.1.Flag.TCF7L2 (Hoverter et al. 2012) and human ApoB (-85-47.E4) luciferase reporter construct (Maeda et al. 2006) have been previously described.

The doxyclcyine (DOX) inducible expression vectors, pTRE-HNF4α2 and pTRE-HNF4α8, were constructed by amplifying the human HNF4α2 and HNF4α8 cDNA, from the respective pcDNA3.1 vectors, with primers that contained an EcoRI site and a

Kozak's sequence (forward) or BamHI site (reverse) and cloning the PCR products into a EcoRI/BamHI-digested pTRE-Tight vector (Clontech). The sequences of the primers are as follows:

5'-HNF4α2_Koz.EcoRI: 5'-GGAATTCCCCACCATGGATATGGCC-3';

5'-HNF4α8_Koz.EcoRI: 5'-GGAATTCCCCACCATGGTCAGCGTG-3';

3'-N1C465.BamHI: 5'-GCGGGATCCCGCTAGATAACTTCCTGCTT-3'.

The expression vector containing the reverse tetracycline transcriptional activator (rtTA) pCAG-rtTA (pCAG-rtTA-IR-PURO) and the Tet-responsive red fluorescent protein (RFP) reporter construct pTRE-RFP (pTRE-Red Max C) were gifts from Dr. Chee-Gee Liew (Liew et al. 2007; Xia et al. 2008).

Cell culture and generation of the Tet-On inducible stable cell lines

The human colorectal cancer cell line HCT116 (American Type Culture Collection (ATCC), CCL-247) was maintained in McCoy's 5A (Iwakata & Grace Modification, with L-glutamine) (Corning Cellgro; cat# 10-050-CV) supplemented with 10% Fetal Bovine Serum (FBS) (BenchMark; cat# 100-106, lot# A07D00C) and 1% penicillin-streptomycin (P/S). Cells were passaged every third or fourth day at 85-95% confluency. The human embryonic kidney cell line 293T (ATCC, CRL-11268) and monkey kidney cell line COS-7 (ATCC, CRL-1651) were cultured in DMEM (Dulbecco's Modification of Eagle's Medium with 4.5 g/L glucose, L-glutamine, and pyruvate) supplemented with 10% FBS or Bovine Calf Serum (BCS) and 1% P/S. All cell lines were maintained at 37^{o}C and 5% CO_2.

To generate the stable lines, HCT116 cells were seeded at 3 x 10^6 cells per well in a 6-well plate and transfected 24 hours (h) later with 1 µg of linearized pCAG-rtTA-IR-PURO DNA using Lipofectamine 2000 (Invitrogen). The following day, cells were trypsinized

and transferred to a 150-mm plate; 24 h later, cells were selected in medium containing 0.50 μg/mL puromycin. Colonies resistant to the drug were screened for the highest constitutive expression of rtTA by transiently transfecting in pTRE-HNF4α2 and an HNF4α reporter construct (ApoB.-85-47.E4.Luc). Parental clone (clone 11) was transfected with linearized pTRE-HNF4α2 or pTRE-HNF4α8 plus an XhoI fragment containing the NeoR gene (10:1) from the pTet-On vector (Clontech) and the final HNF4α2 and HNF4α8 lines were selected with 50 μg/mL and then 70 μg/mL G418, along with 0.50 μg/mL puromycin.

The HCT116rtTA stable parental line (PL) was maintained in the modified McCoy's 5A supplemented with Tet-free 10% FBS (Benchmark Cat. #100-106 Lot. # A07D00C), 1% P/S, and 0.53 μg/mL puromycin. The HCT116rtTA HNF4α2 and HNF4α8 lines were maintained in a similar fashion with the addition of 70 μg/mL G418.

Migration and Invasion assay

TetOn inducible HCT116 (PL, α2, or α8) clones were seeded at 1.5-2.25 x 10^6 cells in a 100-mm plate and, 24 h later, treated with or without 0.5 μg/mL of Doxycycline (DOX, Clontech). Then 48 h after induction, cells were trypsinized, counted, and resuspended in serum free medium supplemented with 0.1% bovine serum albumin (BSA), with or without hydroxyl urea (2 mM), for migration and invasion assay. Here, 5 x 10^4 cells were transferred to the upper chamber of an invasion or migration transwell plate (BD Biocoat). McCoy 5A media supplemented with 20% FBS was added to the lower chamber; 48 h later, the top chambers were removed, stained with hematoxylin, and viewed under a microscope. Images were taken at 20x magnification and printed out for manual counting of the cells that had invaded or migrated to the other side of the upper

chamber. Invasive indexes were calculated by dividing the number of cells invaded over the number of cells migrated.

Xenograft assay

One day prior to injection, TetOn inducible HCT116 cells (PL, HNF4α2, or HNF4α8) were seeded at 4 x 10^6 cells per 150-mm plate. On the day of injection, 3 x 10^6 cells were trypsinized and subcutaneously injected into the flank of 7-8 weeks old athymic nude male mice, strain 01B74 (National Cancer Institute, Frederick, MD). Eight days later, after the tumors reached about 48 mm^3 (measured with calipers), mice were switched to a diet either lacking DOX (7012 Teklad LM-485, Harlan Laboratories) or with 625 mg/kg DOX (TD.05125, Harlan). Food was changed every other day and tumor size was monitored weekly for ~3 weeks, at which point the mice were sacrificed via CO_2 asphyxiation. Tumors were removed from the inner side of the skin with a scalpel and the interior of the mouse was checked for any visible metastasis. Tumors were weighed and snap frozen for subsequent analysis. Xenografts using Matrigel were performed in a similar fashion except that Matrigel (BD Matrigel Matrix, high concentration, Cat# 354248, BD Biosciences) was added to the cells at 25% final volume immediately prior to injection. The care and handling of the mice were in accordance with the guidelines from the University of California, Riverside, Institutional Animal Care and Use Committee (IACUC).

Immunoblot (IB) analyses

Protein extracts were separated by 10% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to PVDF membrane, Immobilon (Millipore)

as previously described (Maeda et al. 2002). Signals were detected using the SuperSignal West Dura Extended Duration Substrate Kit (Thermo Scientific). Bradford assay (Bio-Rad) was used to measure protein concentration: 20-30 µg of whole-cell lysates (WCE) (Maeda et al. 2006) or nuclear extracts (NE) (Jiang et al. 1995) were loaded per lane. Coomassie staining of the blot verified equal loading of protein. Primary antibodies (Ab) were mouse monoclonal anti-HNF4α P1/P2 (R&D, Cat#PP-H1415-00) and affinity purified anti-α445 (Sladek et al. 1990) that recognize the C-terminus of both P1- and P2-HNF4α isoforms; monoclonal anti-HNF4α P1 and P2 (R&D, Cat#PP-K9218-00 and Cat#PP-H6939-00, respectively) that recognize the different N-termini of HNF4α; anti-Flag (Sigma, M2), and anti-TCF7L2 (Millipore, Cat#6H5-3). Secondary antibodies were horseradish peroxidase (HRP)-conjuated goat anti-rabbit (GαR-HRP) or goat anti-mouse (GαM-HRP) from the Jackson ImmunoResearch Laboratories.

HNF4α and TCF1/4 Protein Binding Microarray (PBM)

Protein binding microarrays (PBMs) were carried out essentially as previously described (Bolotin et al. 2010; Fang et al. 2012). In Figure 3.6A, the HNF4α-centric PBM2 design described in Bolotin et al. (Bolotin et al. 2010) was used with NE of human HNF4α2 or dnTCF1 from transfected COS7 cells. In Figure 3.7, another custom-designed array was ordered from Agilent (SurePrint G3 Custom GE 1X1M), which contained a ~60 base oligonucleotide corresponding to sequences within 100 bp of the center of published HNF4α (and RXRα) ChIP-Seq peaks from HepG2 and CaCo2 cells (Wallerman et al. 2009; Verzi et al. 2010). A total of ~125,000 DNA sequences including SNP alleles from dbSNP Version were spotted in quadruplicated (~125k loci x ~2 alleles x 4 replicates = 1

million spots of DNA) on the slide as single-stranded DNA, which was subsequently made double-stranded using a primer to a common linker sequence, dNTPs (GE Healthcare) and Thermo Sequenase (Cat#78500 Affymetrix). For hybridization, ~6 µg of human HNF4α2 or HNF4α8 in NE from transfected Cos-7 cells were processed with a 30 kDa cut-off Amicon column to a final concentration of 110 mM KCl and then applied to the arrays diluted 1:10 in binding buffer (110 mM KCl, 20 mM Hepes, 8 mM EDTA, 8 mM EGTA, 0.1% Tween 20, 20 µg salmon sperm DNA, pH 7.8). TCF WCE (~600 ng Flag.tagged TCF4) was applied to the array diluted 1:10 in casting buffer (25 mM KCl, 60 mM NaCl, 0.75 mM MgCl2, 35 mM Tris-HCl pH 7.5, 2.5 mM beta-mercaptoethanol, 0.25% NP-40, 20 µg salmon sperm DNA). Arrays were incubated for 100 min at RT, after washing the bound TFs were detected using anti-HNF4α P1/P2 or anti-Flag M2 Ab (Sigma-Aldrich, Cat#F1804) for dnTCF1 or TCF4 at a 1:100 dilution in 2% milk plus 0.1% Tween 20 in PBS overnight (ON) at RT, followed by a conjugated secondary Ab (goat anti-mouse IgG H&L DyLight 550, Cat#84540, Pierce) diluted 1:50 in 2% milk as above for 90 min. Secondary Abs were detected using an Agilent G2565CA Microarray Scanner at the UCLA DNA Microarray Core. Extraction and normalization of the data were as described previously (Bolotin et al. 2010). Position weight matrices (PWM) were generated using SeqLogo (Schneider and Stephens 1990) and Weblogo v2.8.2 (Crooks et al. 2004).

HNF4α and TCF Reporter Constructs and Assay

Luciferase reporter constructs containing one of three HNF4/TCF binding sites driving a minimal promoter were generated by cloning the appropriate synthetic oligonucleotides (Integrated DNA Technologies, IDT) containing NheI and HindIII overhangs into pGL4.23

(Luc2/minP) (Promega); a third site (AseI) was incorporated to screen for positive clones (5'-CTAGTAGGC[motif-sequence]GCGCGATTAAT.AGCT-3'; restriction sites NheI, AseI, and HindIII are underlined; see Fig. 3.6A for motif sequences).

Cells (293T) were seeded at 1.6-2.0 x 10^5 per well of a 12-well plate and 24 h later transfected with 80 ng of expression vector (pcDNA3.1.HNF4α2, pcDNA3.1, and pcDNA4/TO.Flag.dnTCF1), 0.5 µg of reporter, and 0.1 µg of CMV.β-gal, using Lipofectamine 2000. The following day, cells were harvested with Lysis Buffer (25 mM glycylglycine pH 7.8, 15 mM MgSO4, 4 mM EGTA, 1% (v/v) Triton X-100) and luciferase and β-galactosidase activities were measured as previously described (Maeda et al. 2002). All transfections were performed in triplicate, normalized to β-gal and performed at least four times.

HNF4α and TCF Fluorescent Gel Shift

Double-stranded oligonucleotides (dsDNA) (5 µg) with 5' adenine overhangs were fluorescently labeled using 5-fold molar excess of Cy3-dUTP (GE Healthcare) and 5 Units Klenow (New England BioLabs) in a 100-µL reaction. Unincorporated label was removed using Mini Quick Spin DNA Columns (Roche). The shift probes contained the motifs shown in Fig. 3.6A (5'-AAAACGCGC[motif sequence]GCCTA-3'). NEs were prepared from COS-7 cells transfected with ~12 µg of Flag.dnTCF1 or pcDNA3.1.HNF4α2 via CaPO4 precipitation as previously described (Jiang et al. 1995). IB analysis was used to normalize the amount of HNF4α2 and dnTCF-1 protein using as standards recombinant HNF4α ligand binding domain plus F domain (LBD/F) (Ruse et al. 2002) and the carboxy-terminal Flag fusion protein (Flag-BAP) from Sigma-Aldrich.

Gel shifts were performed using a 10% non-denaturing polyacrylamide gel as previously described (Jiang et al. 1995). Briefly, each shift reaction contained: 1.5 µL 50 mM EDTA, 1-2 µL purified, labeled probe at 0.3-5.0 ng/µL, 4.0 µL 5 x Shift Buffer, and 1.5 µL poly(dIdC) or sonicated salmon sperm DNA at 1 µg/uL. NE (5-10 µg total protein) was added to achieve the indicated amounts of dnTCF1 and HNF4α2 protein; BSA or NE from mock-transfected (pcDNA3.1) cells was used to bring up the volume to 20 µL and total protein to 10 µg. After 20-30 min at RT, six µL of the reaction were loaded per lane and the gel was run at 12 mA constant current for 45 to 60 min. A Typhoon 9410 imager was used to visualize the bands on the gel.

RNA-Seq analysis

Tet-On inducible HCT116 clones (PL, HNF4α2, and HNF4α8) were seeded at 5.5-5.6 x 10^5 cells per well of 6-well plates. Six hours later, cells were treated with 0, 0.1, or 0.3 µg/mL of DOX. Twenty-four hours after induction, cells were harvested by adding 700 µL QIAzol Lysis Reagent (Qiagen) to the adherent cells. The miRNeasy Mini kit (Qiagen) was used to extract and purify total RNA; 4 µg of each RNA sample was used to generate a polyA+ RNA library using the Illumina TruSeq RNA Sample Prep v2 Kit (Cat# RS-122-2001). Libraries were submitted for 50bp-paired end next generation sequencing by Illumina HiSEQ 2000 at the Genomics Core in the UCR Institute of Integrated Genome Biology (IIGB). A total of 21 samples (seven different conditions, each condition in triplicates) were multiplexed and sequenced in two lanes, each of which yielded ~442 M (million) reads or ~42 M reads per sample.

Paired-end sequencing reads were aligned to the human reference genome (GRCh37/hg19) with Tophat v1.2 (Trapnell et al. 2009) using the default parameters with

the exception of allowing up to 10 alignments to the reference for a given read, instead of the default value of 20. The data aligned by Tophat were processed by Cufflinks (Trapnell et al. 2010) to assemble transcripts and to measure their relative abundance in FPKM (fragments per kilobase of exon per million fragments mapped). Assembled transcripts from experimental samples were compared with the RefSeq annotated transcriptome downloaded from the UCSC Genome Browser and examined for differential expression using the Cuffcompare and Cuffdiff utilities included in the Cufflinks package. Cuffdiff was run with FPKM upper-quartile normalization and a false discovery rate (FDR) threshold of 5%. Cufflinks calculates differential expression at the transcript, primary transcript, and whole gene levels. The following criteria were used to select differentially expressed genes: (1) a fold change (FC) of at least 1.5 or more; (2) at least two replicates have an FPKM $\geq$5; (3) the triplicates for a given condition have a coefficient variant (CV) $\leq$0.5.

Chromatin immunoprecipitation sequencing (ChIP-Seq) analysis

Tet-On inducible HNF4α2 and HNF4α8 HCT116 clones were seeded at 5-6 x 10^6 or 7-9 x 10^6 cells per 100-mm or 150-mm plate, respectively. Six hours after seeding, cells were treated with 0 or 0.3 μg/mL of DOX. Twenty-four hours after induction, cells were harvested as previously described (Li et al. 2007; Hwang-Verslues and Sladek 2008) with minor modification. Cells were fixed with 1% formaldehyde in 1 x phosphate buffered saline (PBS) for 10 min at RT; the cross-linking was stopped with 0.125 M glycine in PBS for 10 min at RT. All subsequent steps were performed at 4^{o}C, using ice-cold buffers. Cells were scraped in PBS plus 1 mM phenylmethylsulfonyl fluoride (PMSF) and dithiothreitol (DTT) and centrifuged at 6,000g for 5 min. The pellet was

resuspended in 0.5 mL hypotonic buffer (10 mM HEPES-KOH pH 7.9, 10 mM KCl, 1.5 mM MgCl2) plus 1 mM PMSF and DTT for 10 min. The nuclei were pelleted and resuspended in 0.34 mL nuclei lysis buffer (1% Triton X-100, 50 mM Tris-HCl pH 8.0, 10 mM EDTA) plus 1 mM PMSF and DTT and 2 µg/mL leupeptin and aprotinin. The samples were sonicated using a Sonic Dismembrator Model 500 (Fisher Scientific) to obtain DNA fragments of about 200-500 bps and then diluted 1:1 in immunoprecipitation (IP) dilution buffer (0.01% SDS, 20 mM Tris-HCl pH8.0, 1.1% Triton X-100, 167 mM NaCl, 1.2 mM EDTA) and pre-cleared with 20 µL of packed Protein G Agarose (Pierce) beads that were preblocked with 100 µg/µL BSA for 30 min. The IP was performed (~1-3 x 10^7 cell equivalents per IP) with the following antibodies for 2 h: 3-5 µg of affinity-purified anti-HNF4α (α-445) (Sladek et al. 1990), anti-TCF7L2 (TCF4) (Millipore, Cat#6H5-3), or mixed equal amount of mouse and rabbit IgG controls (Millipore Cat#12-371, and Santa Cruz Cat#sc-2027, respectively). Protein G beads (in PBS (30-40 µL 1:1 slurry) and the sample were nutated overnight at 4^{o}C. The IP sample was washed with three sequential buffers for 5 min each at RT: TSE I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl), TSE II as TSE I but with 500 mM NaCl, and TSE III (0.25 mM LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0). At the final wash, the IP sample was washed two times with 1 x TE for 5 min at RT. The precipitated material was eluted with 150 µL IP Elution Buffer (1% SDS and 0.1 M $NaHCO_3$) RT for 20 min. The pellet was transferred to a new tube and eluted a second time with 1 min of boiling preceding the 20 min incubation. The two elutions were combined and incubated at 65^{o}C for 4 to 5 h to reverse the crosslinks. The DNA was precipitated with 1 mL 100% ethanol ON at -20^{o}C, washed and resuspended in 50-100 µL TE. RNA and protein digestions were performed for 25 min at RT with 1 µL of 10

μg/uL RNAse A in water and 1 h at 55^{o}C with 11 μL of 10 x proteinase K buffer (100 mM Tris-HCl pH 8.0, 50 mM EDTA pH 8.0, 500 mM NaCl) and 1 μg Proteinase K (IBI Scientific), respectively. GeneJET PCR Purification Kit (Thermo Scientific) was used to clean and purify the DNA material and Qubit fluorometer in the UCR Genomics Core was used to measure the DNA concentration; 5-20 ng of ChIP'd material was used to generate libraries using a BIOO Scientific ChIP-Seq DNA library kit (NEXTflex ChIP-Seq kit cat# 5143-02 and barcodes cat# 514122). Libraries were submitted for 50bp single-end next generation sequencing as noted above.

Sequencing reads (~30M per sample) were pre-processed and mapped to the human reference genome (GRCh37/hg19) with Bowtie v1.8 (Langmead et al. 2009). The alignment files generated by Bowtie were analyzed with MACS v2 (Zhang et al. 2008). We first used the 'callpeak' function to generate bedGraph files for all conditions. The generated bedGraph files were further analyzed with the function 'bdgdiff' to identify differential binding events between any two conditions. A log10 likelihood ratio cutoff was applied in this step to obtain 3 types of differential regions: i) more enrichment in condition 1 over condition 2; ii) more enrichment in condition 2 over condition 1; and iii) similar enrichment in both conditions. Higher values of the log10 likelihood ratio reflect bigger differences in binding. We also ran the 'bdgcmp' function which deducts noise by comparing two signal tracks and generate a fold-enrichment bedGraph. The peaks identified by ChIP-Seq were further analyzed with the R Bioconductor package, ChIPpeakAnno (Zhu et al. 2010), to retrieve Ensembl genes that are nearest to transcription start site (TSS).

Bioinformatics and statistical analyses

VENNY (Oliveros 2007) was used to compare lists of genes in the different cell lines. Gene Ontology (GO) analysis was performed using DAVID Bioinformatics Resources 6.7 (Huang da et al. 2009). Cistrome (Liu et al. 2011) was used to generate the heatmap for HNF4α and TCF4 ChIP-Seq peaks in α2 and α8 lines. Cisgenome (Ji et al. 2011) was used to identify HNF4α and TCF4 target genes and extract DNA sequence from ChIP peaks. Nearby genes were defined as 50 kb or less from the peak center. The fold enrichment bigwig files of the ChIP-Seq data were uploaded to UCSC Genome Browser (Kent et al. 2002) and Integrated Genome Viewer (IGV) (Thorvaldsdottir et al. 2013) for visualization. R was used to identify overlapping HNF4α2 and HNF4α8 peaks, defined as at least 13 nucleotides (NT). MEME analysis was used to mine motifs (Bailey and Elkan 1994). Line and bar graphs are plotted as means±SEM or SD as indicated. A Student's t-test was used to calculate P-values: $P < 0.05$ was considered significant.

All RNA-Seq and ChIP-Seq data have been submitted to GEO, accession number GSE62889 and GSE62890, respectively. All PBM data are uploaded onto The Nuclear Receptor DNA Binding Project website (http://nrdbs.ucr.edu).

Results

Generation and characterization of the HNF4α inducible cell lines

To separate the function of the HNF4α promoter-driven isoforms (Fig. 3.1A) in human colon cancer, we generated Tet-On inducible system in HCT116 (colon cancer) cells that express either the human HNF4α2 or HNF4α8 in a controlled manner. The parental line, HCT116, is poorly differentiated and considered to have stem cell-like

properties (Yeung et al. 2010). The line does not express any endogenous HNF4α protein (Fig. 3.1B); it does, however, contain a mutant allele of the β-catenin gene (*CTNNB1*) resulting in a constitutively active Wnt/β-catenin pathway (Morin et al. 1997). TCF4 is the primary member of that family expressed in HCT116 cells (Korinek et al. 1997) (Fig. 3.1D). The chromatin occupancies of TCF4 and β-catenin in HCT116 line is well characterized by the ENCODE project (Bottomly et al. 2010; Frietze et al. 2012), making it a good model to examine the role of the HNF4α isoforms in colon cancer.

Induction and appropriate expression of the HNF4α isoforms were verified by immunoblotting (IB) with isoform-specific antibodies (Fig. 3.1A,B). HNF4α8 was detected as early as 2 h post DOX induction and the HNF4α2 protein was typically expressed at slightly lower levels then HNF4α8 (Fig. 3.1C). Expression of both HNF4α2 and HNF4α8 peaked at 24 h and began to decline after 72 h (Fig. 3.1C), possibly due to the fact that HNF4α expression tended to decrease cell viability with time after induction (data not shown). The expression of HNF4α did not significantly affect β-catenin or TCF4 expression (Fig. 3.1D), confirming the suitability of the lines for our studies.

HNF4α2 is more tumor suppressive than HNF4α8

In order to determine the effect of the HNF4α isoforms on tumor growth, in mouse, we employed a xenograft assay in which we subcutaneously injected the HCT116 HNF4α2 and HNF4α8 lines into immunocompromised mice and allowed the tumors to develop for 8 days (to ~48 mm^3) before placing the mice on a diet containing DOX (Fig. 3.2A). Tumor growth was monitored for another ~23 days at which point the mice were sacrificed. The HNF4α2 line in the presence of DOX resulted in a significantly

decreased tumor volume but not the HNF4α8 line (Fig. 3.2C): HNF4α2 and HNF4α8 were appropriately expressed in the majority of tumors (Fig. 3.2C, right).

A second xenograft assay in which Matrigel was added to the cells at the time of injection in order to enhance tumor growth showed again that the HNF4α2 line in the presence of DOX resulted in statistically smaller tumors ($P \leq 0.03$) compared to the absence of DOX (Fig. 3.2D). IB analysis confirmed that the injected cells expressed the appropriate HNF4α isoforms (Fig. 3.2B). While the parental line lacking the HNF4α transgene also showed somewhat smaller, although not statistically significant ($P \leq 0.07$) tumors in the presence of DOX, there was no difference in tumors weights +/-DOX in the HNF4α8 line, despite the expression of HNF4α8 in the DOX-treated tumors (Fig. 3.2D, bottom). Finally, HNF4α2 displayed a lower and HNF4α8 a higher invasive index in an in vitro invasion/migration assay (Fig. 3.2E). These results suggest that while HNF4α2 is clearly tumor suppressive in human colon cancer, HNF4α8 plays a different, apparently less suppressive role.

HNF4α2 and HNF4α8 regulate unique sets of genes relevant to tumor growth

To determine the basis for the difference on tumor growth between HNF4α2 and HNF4α8, we preformed next generation sequencing (RNA-Seq) on the poly A+ RNA isolated from the HCT116 lines treated with DOX for 24 h. We used two different concentrations of DOX (0.1 and 0.3 μg/mL) for the HNF4α8 line since the HNF4α8 protein tended to be expressed at a higher level than HNF4α2 (Fig. 3.3A,B). IB analysis verified similar levels of expression of HNF4α2 and HNF4α8, which were comparable to those expressed in the normal mouse colon (Fig. 3.3B). At least 635 known genes were differentially expressed 1.5-fold or more in the HNF4α2 and HNF4α8 lines after DOX

induction (0.3 μg/mL) ($P < 0.01$). Of those, 512 genes were up regulated while 123 were down regulated (Fig. 3.3C), consistent with HNF4α acting as a positive regulator of gene expression (Yuan et al. 2009). Nearly half of the upregulated genes (156) were common in both the HNF4α2 and HNF4α8 lines, while 200 genes were uniquely regulated by the different isoforms (Fig. 3.3D). Interestingly, HNF4α8 downregulated more genes than HNF4α2 (98 v. 25 genes). However, when a 2.0-fold cut off was applied, again there was more genes upregulated than were downregulated (Fig. 3.3E). The parental rtTA line had only four genes dysregulated at the 1.5-fold cut off (not shown). Since the two concentrations of DOX used for the HNF4α8 line (0.1 and 0.3 μg/mL) showed very consistent results, we combined the two datasets. A heatmap shows some of the most highly dysregulated genes with their non-log fold change -/+ DOX (Fig. 3.3G).

To determine the functions of these genes regulated by the HNF4α isoforms, we performed Gene Ontology (GO) and found that both HNF4α2 and HNF4α8 upregulated genes involved in drug metabolism, oxidative stress, wound healing, negative regulation of phosphorylation and glycoprotein metabolism (Fig. 3.4A). HNF4α2 specifically upregulated genes involved in cell death, angiogenesis, response to extracellular stimuli while the only category of genes upregulated only by HNF4α8 was cell adhesion, although several of those genes actually promote cell growth (*ALCAM, BCAR1, LAMB1, and CIB1*). GO analysis did not yield a distinct category of HNF4α2 downregulated genes, but three categories were noted for HNF4α8 only genes – kidney development (HNF4α8 is not expressed in the adult kidney), enzyme-linked receptor protein signaling, and anti-apoptosis (Fig. 3.4A). The GO categories most relevant to tumor growth are noted in Figure 3.4B. Interestingly, ninjunin 1 (*NINJ1)*, an inhibitor of the cyclic dependent kinase gene p21 (*CNDK1*) and inducer of cellular senescence in tumor cells

(Toyama et al. 2004), and one of the cell adhesion genes upregulated by HNF4α8, was the gene most highly downregulated by a knockdown of HNF4α2 in HepG2 cells (Bolotin et al. 2010). Fifteen genes involved in growth promotion or inhibition that showed significant differences between the two lines where at least one line had a ≥2.0 fold change (Fig. 3.4B). HNF4α8 upregulated eight genes involved in growth promotion to a greater degree than HNF4α2, while HNF4α2 upregulated seven genes involved in growth inhibition to a greater degree than HNF4α8. Two of the genes significantly more upregulated by HNF4α8 than HNF4α2 were Ras homolog family member U (*RHOU*) and Dickkopf (Dkk) 4, which regulates focal adhesion formation and cell migration (Berenguer-Daize et al. 2013) and enhances migration and invasion potential (Pendas-Franco et al. 2008), respectively (Fig. 3.4B,C), consistent with the higher invasive index for HNF4α8 (Fig. 3.2F). Also of note are *PIM1*, an oncogenic serine/threonine kinase (Weirauch et al. 2013) and *ERBB3*, a member of the epidermal growth factor receptor (EGFR) family of receptor tyrosine kinases (Sollome et al. 2014) (Fig. 4B,C). Activation of *PIM1* promotes proliferation and inhibits apoptosis and leads to upregulation of the ERBB3 (Siu et al. 2011). Likewise, the two growth suppressive genes most upregulated by HNF4α2 than HNF4α8 were pleckstrin homology domain containing family O member 1 (*PLEKHO1*) and myosin binding protein H (*MYBPH*), both of which suppress tumor progression in vivo (Tokuda et al. 2007; Hosono et al. 2012). A group identified 25 tag SNPs associated with a high probability risk for CRC, some of which are located in regulatory elements that are active in normal colon and/or colon cancer cells (Yao et al. 2014). One SNP (rs10411210) in particular that is located in the exon of *RHPN2,* but was not changed significantly between tumor and cancer, overlaps with HNF4α2 peak and similarly was not dysregulated in the RNA-Seq by HNF4α2. They also identified

several genes near promoter SNPs that are highly expressed in tumor versus normal colon tissue. Two of these genes (*FOXM1* and *RAD51AP1*), which are involved in cell proliferation and homologous DNA repair, respectively, are downregulated only by HNF4α2, while one gene (*TERC*), which is involved in maintaining telomeres, is upregulated only by HNF4α8. Taken together, the differential gene expression observed between the HNF4α isoforms is consistent with their effect on tumor growth and invasion/migration (Fig. 3.2).

HNF4α2 and HNF4α8 exhibit somewhat distinct chromatin occupancy in vivo

Considering that HNF4α2 and HNF4α8 have identical DNA (and ligand) binding domains (Fig. 3.1A), it was rather surprising that they regulate so many genes in a distinct fashion. In order to determine how many of the dysregulated genes in the RNA-Seq are direct targets of the HNF4α isoforms, we performed ChIP-Seq analysis on the individual cell lines 24 h after DOX induction. Although HNF4α8 continued to be expressed at a somewhat higher level than HNF4α2 (Fig. 3.5A), there were many more ChIP-Seq peaks in the HNF4α2 line compared to HNF4α8 (14,897 v 2,417), with ~2,200 of the peaks common among the two lines (Fig. 3.5B, left). Cisgenome analysis identified ~6,600 genes within 50 kb of HNF4α2 peaks and ~1,700 genes within 50 kb of HNF4α8 peaks (Fig. 3.5B, right). Cross-referencing the ChIP-Seq data with the RNA-Seq data ($\geq$1.5 FC) showed that among the genes with unique HNF4α2 ChIP-Seq peaks, 101 were differently upregulated by HNF4α2 (Fig. 3.5C, top left). A substantial number of the HNF4α2 unique genes (80) had ChIP-Seq peaks for both HNF4α2 and HNF4α8 while another 67 genes had no HNF4α peaks at all. All but eleven genes in Figure 3.4B (*MYBPH, HTRA2, CTSD, SFN, GP1BB, COL6A1, CIB1, ANXA1, FOXC1,*

NRG1, and MSX2) had an HNF4α ChIP-Seq within 50 kb of TSS. Interestingly, for the HNF4α8 up regulated genes, only one gene (*PLXNB1*) was occupied solely by HNF4α8, while 109 genes had peaks for both isoforms and 94 had peaks only for HNF4α2 (Fig. 3.5C top right). Examples of dysregulated genes with HNF4α ChIP peaks (*VTN, LAMB2, APOBEC1, PLXNB1, IDH1, and RHOU*) are shown in the lower half of Figure 5C. Interestingly, a mutation in isocitrate dehydrogenase 1/2 (*IDH1/2)* was recently shown to inhibit the ability of HNF4α2 to differentiate hepatocytes, thereby causing an increase in biliary cancer (Saha K. Supriya 2014). Some of the genes upregulated by both HNF4α2 and HNF4α8, although not necessarily to the same degree, that are involved in wound healing have HNF4α binding sites nearby (*KLK6, SDC1, LAMB2, ADM, ERBB3, MAP2K3, HBEGF, NFE2L1, CDH3*), consistent with HNF4α playing a role in protecting against colonic epithelium from inflammatory bowel disease (Ahn et al. 2008; Babeu et al. 2009; Darsigny et al. 2009; Hatziapostolou et al. 2011; Chahar et al. 2014). All told, there were many more genes with HNF4α ChIP-Seq peaks than were dyregulated in the RNA-Seq. While this is not uncommon in genomics analysis (Gougelet et al. 2014), a lower cut off (1.2-fold) showed more overlap (not shown) and a longer induction time (>24 h DOX) would have presumably increased the number of dysregulated genes without significantly altering the HNF4α peaks.

Identification of shared and unique binding motifs for TCF/LEF and HNF4α

With the increasing reports linking HNF4α and TCF4 in colon cancer and since there was only a partial overlap between the HNF4α RNA-Seq and ChIP-Seq genes, we hypothesized that TCF4, which is expressed endogenously in HCT116 cells (Fig.s 3.1D, 3.3B), might come into play. It was noted previously that the core of the canonical TCF

binding motif (CCTTTGA) is the reverse complement of the center of the HNF4α consensus sequence (AGGTCAaAGGTCA) (Hatzis et al. 2008) (Fig. 3.6A, right). While the overlap contains an A/T base pair that is classically considered to be a spacer sequence between the two more important AGGTCA half sites, we recently showed that the A/T spacer is in fact very important for HNF4α DNA binding (Fang et al. 2012).

To determine the extent of the overlap in binding specificity between HNF4α and TCF, we employed a high throughput DNA binding assay called protein binding microarray (PBMs) with which we could examine binding to ~5000 DNA sequences, in triplicate, in a single experiment, and identified 90 unique DNA sequences that were bound by both dominant negative (dn) TCF1 and HNF4α2 (Fig. 3.6A, left). (PBM assays with 45,000 unique sequences showed that all the TCFs – dnTCF1, TCF3, LEF1, and TCF4 – have nearly identical DNA binding specificity, data not shown.) Position weight matrices (PWMs) derived from 156 sequences bound only by dnTCF1, 523 sequences bound only by HNF4α2, as well as the 90 sequences bound by both factors, revealed highly related binding motifs, all of which contain the CTTTG with variations in flanking sequence (Fig. 3.6A, center).

We selected three sequences that had weak or strong affinity for HNF4α and TCF (HstrgTstrg, HstrgTstrg, HstrgTwk) (Fig. 3.6A, right) and designed fluorescent probes for gel shift analysis. The results confirmed the relative affinity of these sequences for TCF1 and HNF4α2, with HwkTstrg yielding a more intense shift band for TCF1 than HNF4α2 and, conversely, HstrgTwk yielding an intense band for HNF4α but not TCF1; HstrgTstrg yielded shift bands of similar intensity for both (Fig. 3.6B). Further characterization of the HstrgTstrg site showed that HNF4α2 bound with a higher affinity than TCF1 in the gel shift assay (Fig. 3.6C).

In order to determine whether TCF1 and HNF4α2 could compete for binding in vivo, we performed transient transfection assays with luciferase constructs containing one of the three motifs used in the gel shift assay. The results show that HNF4α2 competes for transcriptional control of the constructs containing high affinity HNF4α sites (HstrgTstrg, HstrgTwk), but not the one containing the low affinity site (HwkTstrg) (Fig. 3.6D,E), suggesting that competition occurred in vivo. IB analysis of the transfected cells confirmed appropriate expression of the two TFs (Fig. 3.6E, bottom). Interestingly, dnTCF1 activated all three of the luciferase constructs under the conditions used here. While dnTCF1 typically acts as a repressor of transcription, others have noted that it can also activate transcription (Van de Wetering et al. 1996). Taken together, these results indicate not only that HNF4α and TCF recognize many of the same binding elements, but also that they can compete in vivo to regulate gene expression.

To identify additional binding motifs shared by HNF4α and TCFs, we designed a second PBM that contained one million spots of DNA corresponding to 250,000 sequences (in four replicates) that we mined from published HNF4α ChIP-Seq data (-/+ 100 bp from peak center) from a human liver cancer cell line, HepG2 (Wallerman et al. 2009), that expresses predominantly P1-HNF4α, and another colon cancer cells line, CaCO2 (Verzi et al. 2010), that expresses predominantly P2-HNF4α (Chellappa et al. 2012). We compared binding by ectopically expressed TCF4 to that of HNF4α2 and HNF4α8 and found 741 DNA sequences bound by both TCF4 and HNF4α8, the majority of which also bound HNF4α2 (Fig. 3.7A, red spots). We then divided the sequences that bound all three TFs into three categories – TCF4 preferred (green), HNF4α preferred (red), similar affinity (blue) – and determined the most prevalent motif in each category using MEME. Again, the CTTTG core was the defining feature for all three categories

with some variations in the flanking nucleotides (Fig. 3.7A, top right); the same was found for sequences bound only by TCF4, only by HNF4α2 or by HNF4α8 and TCF4 (Fig. 3.7B,C).

To better distinguish DNA binding specificity, we examined individual sequences. The PBM design incorporated all the SNPs found in the ChIP-Seq peaks, thus providing natural single base mutations. Of the 741 sequences bound by both HNF4α isoforms and TCF4 there were 107 SNPs that altered the affinity of either TCF4 or HNF4α (Fig. 3.7D, black spots). Of those, 35 SNPs prevented both HNF4α2 and TCF4 from binding DNA (dual non binders, red spots) while another 72 interfered only with TCF4 binding (HNF4α only binders, purple spots). Interestingly, the majority of the dual non binders (22/35) had the SNP in the CTTTG core while only 10/75 of HNF4α only binders did (Fig. 3.7D, lower left). In fact, the majority of the HNF4α only binders (45/75) did not contain a CTTTG anywhere in the sequence. In contrast, both groups had a similar proportion of SNPs in the flanking sequence (6/35 and 17/72, respectively). Examples of individual SNPs and associated genes are given in Fig. 3.7E, including three that were dysregulated by the HNF4α isoforms (*RHOU*, *DKK1*, and *SPTBN1*). These results indicate that TCF4 cannot tolerate any mutations in the CTTTG core while HNF4α can, depending on the flanking sequence, revealing a subtle yet important difference in DNA binding between these TFs.

Interplay between HNF4α, TCF4 and AP-1 in vivo

To examine HNF4α and TCF4 binding in vivo, we performed ChIP-Seq on TCF4 in the absence or presence of DOX in the HCT116 HNF4α2 and HNF4α8 lines (Fig. 3.8A). A heatmap comparing all TCF4 and HNF4α peaks in both lines showed that the

majority of the peaks were centered around the TSS and had similar clustering patterns for TCF4 (-/+DOX) and HNF4α (+DOX) (Fig. 3.8B), suggesting that, on a genome-wide level, there were no large-scale alterations in TCF4 peaks when either HNF4α2 or HNF4α8 was expressed. Nevertheless, since the TCF4 patterns -/+ DOX were not 100% identical and since the PBM analysis showed that global views can obscure important differences in individual binding sites, we identified all differential binding peaks between TCF4 (-DOX) and TCF4 (+DOX) and divided them into three categories: the TCF4 (-DOX) peak is significantly smaller than the TCF4 (+DOX) peak; TCF4 (-DOX) is roughly equal to TCF4 (+DOX); and TCF4 (-DOX) is larger than TCF4 (+DOX) (Fig. 3.9A). We then asked how many of those peaks contained an HNF4α overlapping peak in the +DOX samples and identified three distinct binding patterns: HNF4α recruits, co-localizes, or competes with TCF4 (Fig. 3.9B). While the total percent of peaks in these different categories was relatively low ($\leq$2.4%), differences between the HNF4α2 and HNF4α8 lines were noted. For example, there were 41 TCF4 peaks that appeared only when an HNF4α2 peak was present (0.18% of all HNF4α2 peaks) but 75 TCF4 peaks when an HNF4α8 peak was present (1.4% of all HNF4α8 peaks), suggesting a preferential recruitment of TCF4 by HNF4α8. There were many more TCF4 peaks that co-localized with the HNF4α peaks (485 for HNF4α2 and 126 for HNF4α8), of which 68 were common to the two cell lines. GO analysis showed that HNF4α2-unique recruiting and co-localizing peaks regulate genes involved in metabolism processes and apoptotic mitochondrial changes, whereas HNF4α8-unique recruiting and co-localizing peaks regulate genes involved in cellular signaling, including the Wnt signaling pathway (not shown), underscoring a functional difference between the HNF4α isoforms in how they interact with TCF4.

An even more remarkable difference was observed between the HNF4α isoforms in the competing peaks, where the TCF4 peak was reduced in the +DOX sample at the same location that an HNF4α peak appeared. There were 43 such peaks in the HNF4α2 line but none called by the MACS program in the HNF4α8 line. Finally, the vast majority of the overlapping peaks in all three categories were found close to the TSS (Fig. 3.9C), suggesting potential functional relevance in regulating gene expression.

Cross referencing the genes with TCF4/HNF4α overlapping ChIP-Seq peaks with the RNA-Seq data using a 1.2 FC cut off showed six out of 33 genes within 50 kb of the competing HNF4α2 peaks had altered RNA-Seq levels (Table 3.1). For example, *SSH1* is a member of the slingshot homolog family of phosphatases that regulate the activity of cofilin in cell migration; knockdown of the gene has been shown to impair directional cell migration (Nishita et al. 2005). The only HNF4α2 ChIP peak in the vicinity of the *SSH1* gene is one where HNF4α2 competes with TCF4 (Fig. 3.10C) and *SSH1* was down regulated in the HNF4α2 RNA-Seq, consistent with the HNF4α2 line having a lower invasive index (Fig. 3.2E). In contrast, HNF4α2 appears to compete with TCF4 to activate *FGGY,* a member of a kinase family that phosphorylates carbohydrates such as L-ribulose, erythritol, L-fuculose, and D-glycerol (Zhang et al. 2011) (Fig. 3.10C). Although the number of dysregulated genes near a competing peak is small, if we extend the search from 50 kb to 500 kb with 3 or more neighboring genes or do a longer time point for the RNA-Seq (48 h oppose to 24 h) we would have a larger list to look at. It has been shown that some enhancers can regulate genes far from the transcription start site, as far as 350 kbp (Yao et al. 2014).

There were additional genes in the recruited and co-localized peaks that were altered in the RNA-Seq with more upregulated genes near recruited peaks (14 up vs. 2

down) and similar numbers of up- and downregulated genes near co-localized peaks (14-16 up and down), including several in the Wnt pathway (*WNT3, DKK1, LRP5*) (Table 3.1). Motif analysis using MEME found the canonical CTTTG core motif as the most enriched sequence in the recruited and co-localized peaks (Fig. 3.9D), consistent with the PBM analysis (Fig. 3.7). However, surprisingly, the CTTTG core was not found in any of the 43 HNF4α2 competing peaks. Instead, the only significantly enriched motif was TGAxTCA (1.3 e-14) (Fig. 3.9E). The TGAxTCA motif was also found as the second or third most enriched peak in the co-localized peaks (3.2 e-89 for HNF4α2 only peaks, 4.4 e-11 for HNF4α8 only peaks and 7.3e-2 for HNF4α2 and HNF4α8 peaks). It was not found in any of the HNF4α2 or HNF4α2 and HNF4α8 recruited peaks, although it was enriched in the HNF4α8 recruited peaks (7.1e -10) (Fig. 3.9D and not shown).

Since TGAxTCA is an AP-1 motif, we compared the HNF4α and TCF4 ChIP-Seq data with JunD and FosL1 ChIP-Seq in HCT116 cells from ENCODE and found that 50-60% of co-localized peaks and 70% of the competing peaks overlapped with AP-1 peaks (Fig. 3.9DE, AP-1 ChIP). Figure 3.10B and 3.10C show examples of co-localizing and competing peaks with overlapping FosL1 and JunD peaks; no overlap was found in the recruited peaks, consistent with the MEME analysis (Fig. 3.10A). On some genes TCF4 co-localized well with both HNF4α2 and HNF4α8 (*ALK*) while on others TCF4 co-localized much better with HNF4α2 (*ACSL1*) or HNF4α8 (*DKK1*) (Fig. 3.10B). Figure 3.10C shows examples of competition of TCF4 by HNF4α2. No competition was observed for HNF4α8 at these peaks; this is because HNF4α8 is found co-localized with TCF4 (e.g., *ABHD2*, *12 peaks total*), TCF4 is not bound there to begin with (e.g., *NUDT13*, *25 peaks*), or the HNF4α8 binding is weak (e.g., *FGGY* and *SSH1*, *4 peaks*).

Discussion

HNF4α is a nuclear receptor that exhibits disparate roles in liver and colon cancer, which we propose here is due to the two promoters (P1 and P2) that drive the expression of different HNF4α isoforms (HNF4α2 and HNF4α8, respectively). While there are a few reports in the literature on differential activity of the HNF4α isoforms (Torres-Padilla et al. 2002; Torres-Padilla and Weiss 2003; Briancon and Weiss 2006; Erdmann et al. 2007), to our knowledge, this is the first in depth functional comparison of the HNF4α isoforms in colon cancer. Likewise, while there are earlier reports of interactions between HNF4α and TCF4 (Cattin et al. 2009; Colletti et al. 2009), this is the first report of a potential three-way interaction between HNF4α, TCF4 and AP-1.

Using human colon cancer cell lines (HCT116) with inducible expression of a single HNF4α isoform, we show for the first time that while P1-driven HNF4α2 clearly suppresses the growth of tumors in colon cancer cells, P2-driven HNF4α8 does not (Fig. 3.2). RNA-Seq analysis suggests that this functional difference is due to differential expression of certain target genes, with HNF4α8 having a tendency to up regulate genes involved in cell proliferation and anti-apoptosis while HNF4α2 is more likely to up regulate genes involved in growth suppression and cell death (Fig.s 3.3, 3.4).

In vitro DNA binding assays and luciferase reporter assays indicated that TCF4 and HNF4α share a common CTTTG core and that the two TFs can compete for transcriptional control of target genes (Fig. 3.6), consistent with results noted previously by others (Hatzis et al. 2008; Cattin et al. 2009; Gougelet et al. 2014). A high throughput in vitro DNA binding assay (PBMs) that examined 250,000 unique DNA sequences derived from HNF4α ChIP'd genomic DNA identified 741 unique sequences that bound both HNF4α and TCF4 (Fig. 3.7). Not unexpectedly those sequences contained the

common CTTTG core found in the TCF and HNF4α consensus sequences (Fig.s 3.6A, 3.7A), but so did an even larger number of sequences bound by either TCF4 or HNF4α alone (Fig. 3.7BC), indicating that consensus PWMs cannot distinguish binding by these TFs, which could lead to false positives when mining ChIP-Seq data for motifs. However, analysis of individual sequences revealed 107 instances in which a SNP in the sequence significantly altered the binding of one or both TFs (Fig. 3.7D), indicating that DNA binding specificity is remarkably complex and can be exquisitely sensitive to sequence alterations. It also suggests that SNPs in individuals can alter binding by TCF4 and HNF4α, and hence potentially gene expression.

Comparison of HNF4α and TCF4 chromatin binding *in vivo* also showed that the vast majority of binding events of the HNF4α isoforms and TCF4 are independent. However, we did find ~770 TCF4/HNF4α overlapping peaks that fell into three categories – peaks where TCF4 is recruited by HNF4α, where TCF4 co-localizes with HNF4α or where TCF4 is competed by HNF4α (Fig. 3.9-11). This is within the range of sequences that were bound by HNF4α and TCF4 in the first and second PBM. Those peaks represent a small but significant fraction of the total HNF4α peaks (2.6% to 3.8%), and an even larger fraction of the total TCF4 peaks (12.3% in the HNF4α8 line and 19.6% in the HNF4α2 line). The overlapping peaks were relatively close to genes (Fig. 3.9D), which were frequently dysregulated in RNA-Seq (Table 3.1), suggesting that the interaction is indeed functional. While others have found previously that TCF4 and HNF4α often co-occupy promoter regions, they did not examine the effect of HNF4α on TCF binding, as we do here (Cattin et al. 2009; Frietze et al. 2012; Gougelet et al. 2014),

HNF4α recruits TCF4 to the chromatin

The recruited peaks for both HNF4α2 and HNF4α8 harbor the common CTTTG core which raises two issues – why does TCF4 not bind on its own and how do the two factors bind to the same site at the same time? Both questions could be explained by DNA bending (Fig. 3.11A). The crystal structure of TCF family member Lef1 shows that TCF binds in the minor groove and bends DNA by ~130^0, consistent with it being an HMG box protein (Giese et al. 1992; Love et al. 1995). In contrast, the HNF4α crystal structure shows HNF4α binding in the major groove (Lu et al. 2008; Chandra et al. 2013). While no bending was noted in that structure, we showed previously in a circular permutation gel shift assay that HNF4α bends DNA by ~80^0 and estimated that the bend centered around the central A in the TGGGCAaAGGTCA site, which is equivalent to the middle T in the common core CTTTG (Jiang et al. 1997). Therefore, we propose that HNF4α, upon binding the common CTTTG core in the major groove bends the DNA and thereby facilitates TCF4 binding in the minor groove (Fig. 3.11A). While the circular permutation assay could not indicate the direction of the bend, it did show that when the N-terminal (and C-terminal) domain of P1-HNF4α was removed, the protein bent DNA somewhat less well (~60^0), suggesting that HNF4α8, which has a truncated N-terminus, may also bend DNA in a different fashion. If that is the case, then it could explain why HNF4α8 yielded more TCF4 recruited peaks than did HNF4α2 (1.4% versus 0.18%) (Fig. 3.9B).

This model of recruitment also raises the issue of why we did not see co-occupancy of HNF4α and TCF4 in the gel shift assay (Fig. 3.6B,C). One possibility is that the putative trimeric complex – HNF4α, TCF4 and DNA – was not stable enough to

survive migration through the gel. Another possibility is that the excess probe in the shift reaction allowed for separate binding events. The final issue for the recruited peaks is the effect on gene expression. Of the 86 total recruited TCF4 peaks for both HNF4α2 and HNF4α8, 25 of those genes (29%) showed an upregulation in the RNA-Seq (≥1.2-fold) whereas only 4 genes were down regulated (4.6%) (Table 3.1), consistent with HNF4α strong transactivation ability. What is not known, however, is whether the presence of TCF4 enhances or mitigates the ability of HNF4α to activate transcription: TCF4 recruits both co-repressors, such as Groucho, as well as co-activators such as β-catenin (Clevers and Nusse 2012)

HNF4α2 competes with TCF4 for chromatin binding

While the recruitment of TCF4 by HNF4α at CTTTG sequences was unexpected, even more surprising were the TCF4 peaks that were competed by HNF4α2, none of which contained the CTTTG core. Rather, all 43 HNF4α2 competing peaks contained an AP-1 binding motif (TGCxTCA) and 30 of the 43 peaks were bound by both FosL1 and Jun in HCT116 cells (Fig. 3.9E, 3.10C, 3.11C). While we could find only one report in the literature of HNF4α interacting with AP-1 (Yang et al. 2009), other nuclear receptors, such as GR and ER have long been known to interact with AP-1 bound to DNA (Paech et al. 1997; Pearce et al. 1998). There are also two papers that show that both TCF4 and β-catenin interact with AP-1 at TGAxTCA motifs (Nateri et al. 2005; Blahnik et al. 2010). While our ChIP-Seq data suggest that HNF4α2 displaces TCF4 in the AP-1 complex, it remains to be determined whether β-catenin is still present, whether it interacts with HNF4α2 and how it might affect gene expression: of the six genes with competing peaks that were altered in the RNA-Seq, two had increased expression and

four had decreased expression. Bottomly et al. explored β-catenin (and TCF4) chromatin binding in HCT116 cells (Bottomly et al. 2010), but none of those peaks overlapped our HNF4α/TCF4 peaks.

Interestingly, the MACS program did not yield any TCF4 peaks that were competed by HNF4α8. While in some instances that was due to the fact that the HNF4α8 line did not have a TCF4 peak at the chromosomal location that was competed in the HNF4α2 line (see *NUDT13* in Fig. 10), there were at least 12 other examples where a TCF4 peak was present, but the induction of HNF4α8 did not decrease the peak signal (e.g., *ABHD2* in Fig. 3.10). Since those peaks lacked a CTTTG motif, but had a TGAxTCA motif as well as an AP-1 complex, we propose that the interaction between HNF4α8 and TCF4 and the AP-1 complex may be fundamentally different than that of HNF4α2, TCF4 and AP-1 (Fig. 3.11B).

HNF4α co-localizes with TCF on the chromatin

The final category of overlapping HNF4α and TCF4 peaks appears to be a combination of the recruitment and competition scenarios as both the CTTTG core and the TGAxTCA motif are observed in the majority of the peaks (Fig. 3.9D). There are two features, however, that distinguish the co-localization category from the recruitment and competition category. The first is that there were many more genes down regulated in the co-localization category than in the recruitment category: 28/63 genes (44%) altered in RNA-Seq versus 4/29 (14%) in the recruitment category. The second is that we observed a similar percentage of HNF4α8 peaks in this category as the HNF4α2 peaks (2.4% and 2.2%, respectively), in contrast to the competition category where we found only HNF4α2 peaks and the recruitment category that had more HNF4α8 than HNF4α2

peaks. The co-localized peaks for both HNF4α2 and HNF4α8 frequently contained TGAxTCA motifs, as well as overlapping AP-1 peaks and CTTTG motifs, making it difficult to determine which of the motifs are relevant for the co-localization.

Finally, it is possible that other sequences are driving binding: while the vast majority of the binders in the PBM had a CTTTG motif, we did observe isolated examples of binding in the PBM to sequences that contained neither a CTTTG nor a TGAxTCA motif (Fig. 3.7D), suggesting an additional level of complexity. (There were also 2316 TGAxTCA motifs present in the PBM but only four were bound by both TCF4 and HNF4α; 126 were bound by TCF4 alone while 63 were bound by HNF4a alone. This could be due to insufficient Fos/Jun present in the Cos-7 extracts used in the PBM.)

Perhaps the most important question raised by our model is that, given the remarkable consensus of the common core for HNF4α and TCF4, why are there not more overlapping peaks? One potential answer is that even one nucleotide change can make a big difference in binding (Fig. 3.7D). This sort of fine tuning ensures that even though both HNF4α and TCF regulate many hundreds of genes, they will interact only on a subset of them, thereby allowing the cell to maintain tight control on gene expression and hence homeostasis.

While this work was in progress three additional papers were published that support our findings. Frietze et al (Frietze et al. 2012) and Gougelet et al (Gougelet et al. 2014) both observed co-localization of TCF4 and HNF4α in HepG2 and isolated mouse hepatocytes, respectively, while Yang et al found that HNF4α overexpression compromised Wnt/β-catenin-induced epithelieal-to-mesenchymal transition (EMT) (Yang

et al. 2013). While none of those works examined the HNF4α-TCF4 interplay on the chromatin in detail, or the effect of the different HNF4α isoforms, they nonetheless support our findings and reveal the importance of the HNF4α and the Wnt/b-catenin/TCF4 interaction. Going forward, additional biochemical and structural studies will be required to determine whether TCF4 and HNF4α can bind to the same CTTTG sequence at the same time and whether DNA bending is involved and differs between HNF4α2 and HNF4α8. The mechanism by which HNF4α2 but not HNF4α8 displaces TCF4 from chromatin bound AP-1 also needs to be investigated. Finally, the role of SNPs that alter the ability of TCF4 and HNF4α to bind DNA needs to be considered as personalized medicine is incorporated into predicting disease susceptibility.

Acknowledgement: We thank Z. Chen at UCLA for help with the scanner and J. Evans for injecting nude mice.

Fig. 3.1

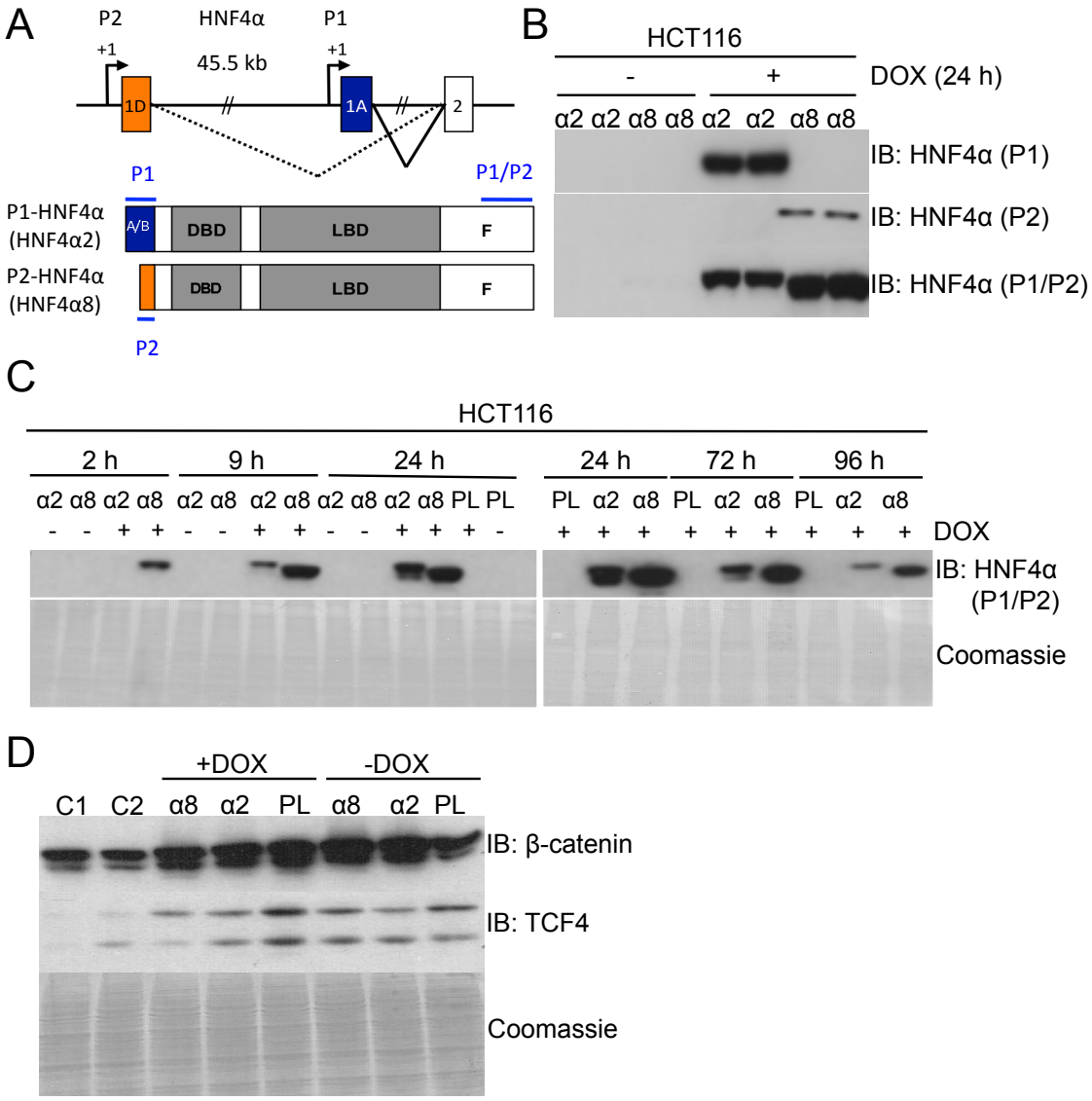

A
P2
HNF4α
P1
+1
45.5 kb
+1
1D
1A
2
P1
P1/P2
P1-HNF4α
(HNF4α2)
A/B
DBD
LBD
F
P2-HNF4α
(HNF4α8)
DBD
LBD
F
P2
B
HCT116
-
+
DOX (24 h)
α2 α2 α8 α8 α2 α2 α8 α8
IB: HNF4α (P1)
IB: HNF4α (P2)
IB: HNF4α (P1/P2)
C
HCT116
2 h
9 h
24 h
24 h
72 h
96 h
α2 α8 α2 α8 α2 α8 α2 α8 α2 α8 PL PL PL α2 α8 PL α2 α8 PL α2 α8
- - + + - - + + - - + + + - + + + + + + + + +
DOX
IB: HNF4α
(P1/P2)
Coomassie
D
+DOX
-DOX
C1 C2 α8 α2 PL α8 α2 PL
IB: β-catenin
IB: TCF4
Coomassie

Figure 3.1 Establishment of stable inducible HCT116 lines expressing human HNF4α2 or HNF4α8. (A) Schematic of a portion of the human *HNF4A* gene and the isoforms generated by its two promoters (P1 and P2). Epitopes to the P1, P2, and P1/P2 antibodies used to distinguish the isoforms are indicated. DBD, DNA binding domain; LBD, ligand binding domain. The P1-HNF4α isoforms contain a full length A/B domain while the P2-HNF4α contains a truncated A/B domain lacking 38 amino acids (aa) encoded by exon 1A (blue) but containing 16 aa from exon 1D (orange). (B, C) IB with antibodies described in (A) of NE (B) and WCE (C) (20 μg per lane) from inducible HCT116 lines HNF4α2 (α2), HNF4α8 (α8) or parental (PL, contains rtTA) treated +/- 0.3 μg/mL DOX for the indicated times. (D) IB for β-catenin and TCF4 in the indicated NE prepared 24 h after the addition of DOX. C1, C2: NE from HEK293T and HepG2 cells, respectively. Coomassie staining verified equal loading.

Fig. 3.2

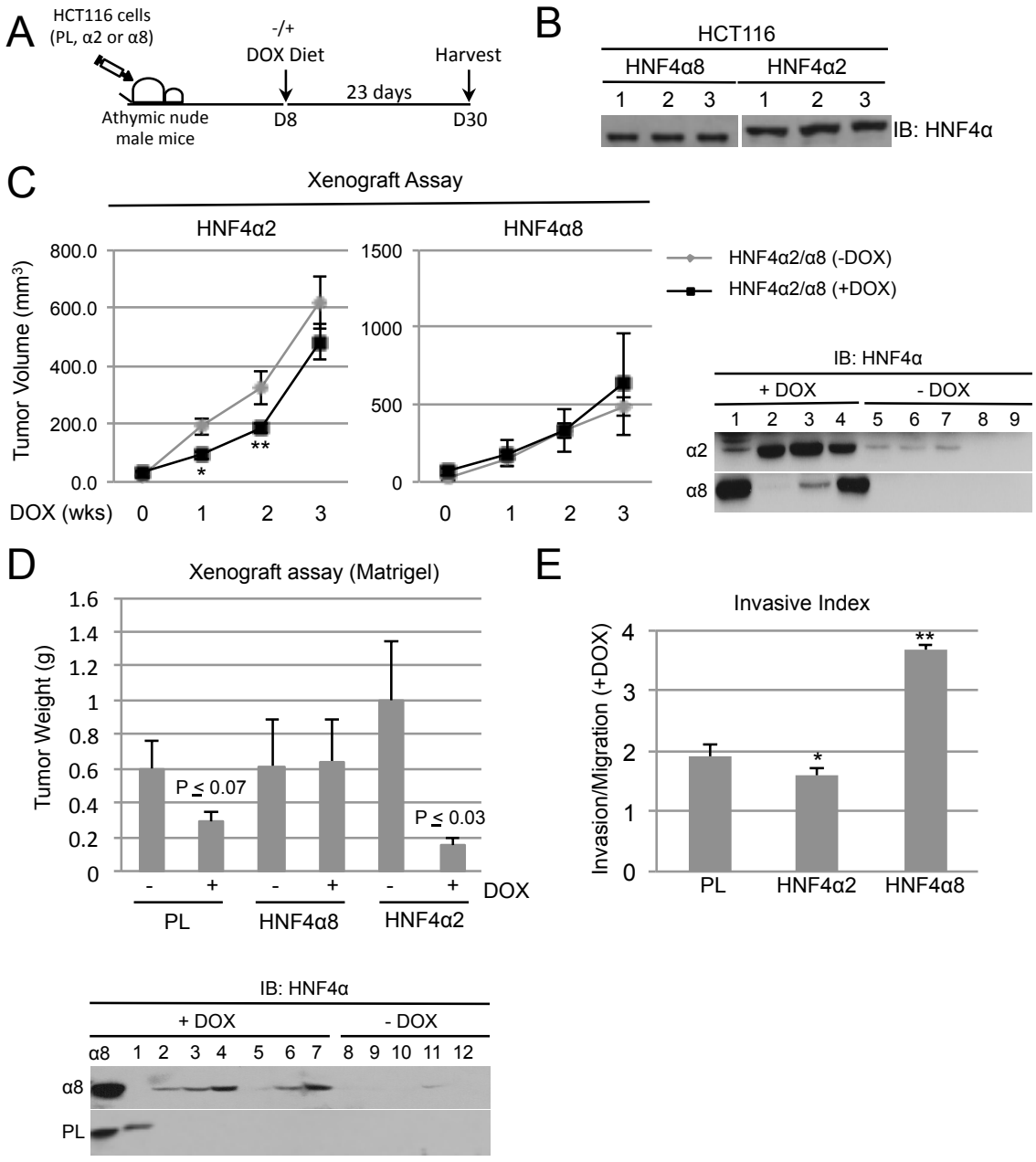
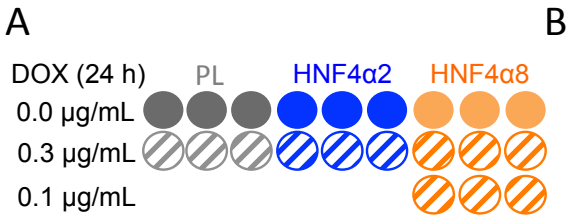
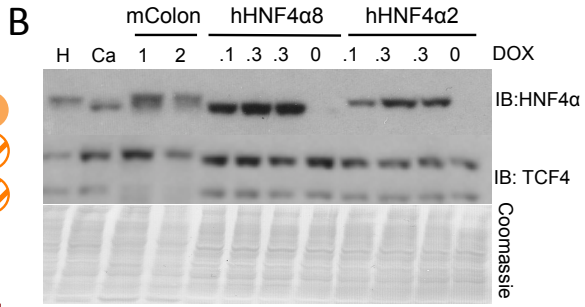

A
HCT116 cells
(PL, α2 or α8)
-/+
DOX Diet
Harvest
23 days
Athymic nude
male mice
D8
D30
B
HCT116
HNF4α8
HNF4α2
1 2 3 1 2 3
IB: HNF4α
C
Xenograft Assay
HNF4α2
HNF4α8
Tumor Volume (mm3)
800.0
600.0
400.0
200.0
0.0
1500
1000
500
0
DOX (wks) 0 1 2 3
HNF4α2/α8 (-DOX)
HNF4α2/α8 (+DOX)
IB: HNF4α
+ DOX
- DOX
1 2 3 4 5 6 7 8 9
α2
α8
D
Xenograft assay (Matrigel)
Tumor Weight (g)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
P ≤ 0.07
P ≤ 0.03
- + - + - +
DOX
PL
HNF4α8
HNF4α2
IB: HNF4α
+ DOX
- DOX
α8 1 2 3 4 5 6 7 8 9 10 11 12
α8
PL
E
Invasive Index
Invasion/Migration (+DOX)
4
3
2
1
0
PL
HNF4α2
HNF4α8

Figure 3.2 HNF4α2 is more tumor suppressive than HNF4α8 in colorectal cancer cells. (A) Schematic of xenograft assay. After injection of the indicated HCT116 lines, half of the immuno compromised mice were switched to a DOX diet on day 8. (B) IB of WCE (20 μg) from triplicate samples of the same batch of HCT116 cells used for injection, induced with 0.3 μg/mL DOX for 24 h. (C, D) HCT116 cells without (C) or with Matrigel (D) were injected as shown in (A). Tumor volume was measured weekly. (D) Tumor weight was determined at the time of harvest. (C,D) Error bars are mean±SEM for each condition: (C) HNF4α2 n = 5 each for – and + DOX *$P \leq 0.01$, **$P \leq 0.05$. HNF4α8 n = 4 each for – and + DOX. (D) Parental (PL) n = 8-9, HNF4α8 n = 7-9, HNF4α2 n = 10 for each condition. (C, right; D, below) HNF4α IB (P1/P2 antibody) of WCE (20-60 μg) of individual tumors. α8, HCT116 HNF4α8 line 24 h after DOX. Coomassie stain verified equal loading (not shown). (E) Invasive index of the parental (PL), HNF4α2, and HNF4α8 lines as described in Materials and Methods. Bars are ±SD; *$P \leq 0.04$, **$P \leq 0.000003$ versus PL. Number of times migration and invasion assay was performed: HNF4α8 (4x); HNF4α2 (1x); PL (2x).

Dr. Karthikeyani Chellappa performed the xenograft and the migration and invasion assays.

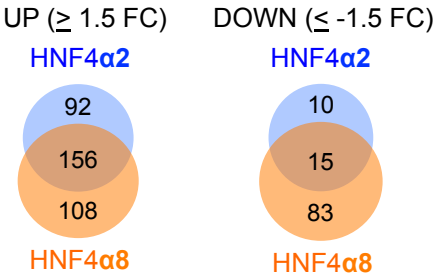
Fig. 3.3

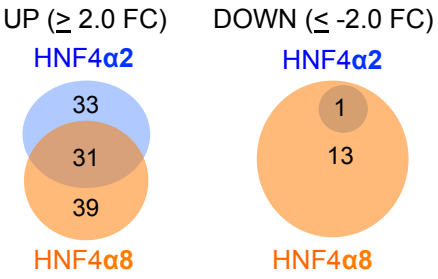


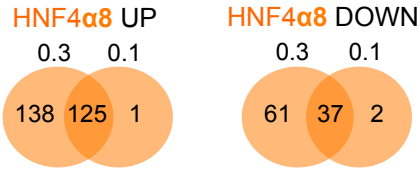


C RNA-Seq — Known Genes

	HNF4α2	HNF4α8	Total
UP (≥ 1.5 FC)	248	264	512
DOWN (≤ -1.5 FC)	25	98	123
Total	273	362	635







G

Genes	α2 (0.3)	α8 (0.3)	α8 (0.1)	PL (0.3)
C11orf86	21.2	20.1	13.33	0
CREB3L3	12.8	13.7	8.5	1.21
APOBEC1	8.4	0	0	0
EPS8L3	7.2	6.3	5.5	0
HPD	6.3	4.8	3.68	0
IL32	6.3	7.2	5	0
AQP3	6.1	-1.31	-1.42	0
RBP2	5.8	5	3.15	0
VTN	5.8	1.92	0	0
MYO15B	5	5.6	4.82	0
SEPP1	4.6	2.9	2.7	1.23
PIPOX	4.3	6.1	3.61	0
ACSL5	4	8.1	5.6	0
KLK6	4	2.5	1.87	0
TUBAL3	3.1	5.5	3.76	0
PLEKHO1	2.4	1.5	1.19	0
CIDEC	2.3	2.4	0	0
ISG20	2.3	1.7	1.33	1.36
METTL7B	2.3	2.1	1.75	0
OSGIN1	2.3	1.6	1.33	1.17
USH1C	2.2	1.8	1.55	0
KLC4	2.1	0	0	0
LAMB2	2.1	1.6	1.27	1.2
OAF	2.1	1.9	1.39	0
FBXO2	2	1.8	1.34	0
OAS1	2	4.7	2.83	0
SERINC2	1.9	2.1	1.78	0
ADM	1.8	2	1.63	0
RHOU	1.8	3.4	2.7	0
ZDHHC9	1.8	2.2	1.75	0
AGPAT9	1.6	2.2	1.88	0
DKK4	1.6	2.1	1.82	0
ERBB3	1.5	2	1.63	0
PIM1	1.4	2.5	1.83	0
VIPR1	1.3	2.6	2.34	0
MYBPH	1.2	-2.1	-1.31	-1.77
GP1BB	0	2.7	1.7	0
NRG1	0	-2.6	-1.62	-1.44
TNFSF18	-1.4	-2.7	-1.7	-1.4
RPS17	-42.9	-3.6	4.26	1.76

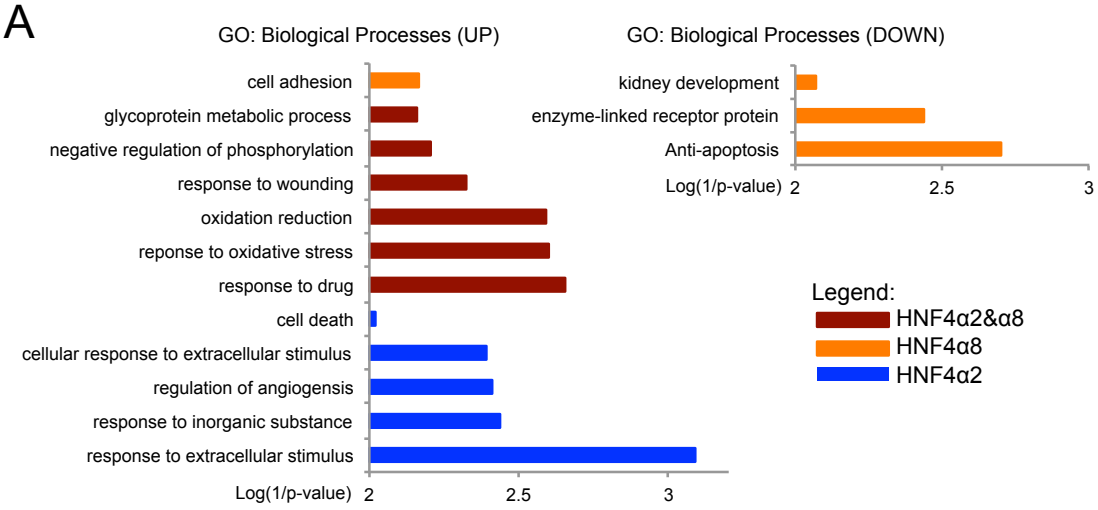
Key: 20 10 5 0 -5 -20 -42.9

Figure 3.3 Differential expression of genes in HCT116 cells by HNF4α2 and HNF4α8. (A) Schematic of the samples submitted for RNA-Seq. Each circle represents one well of a 6-well plate. (B) IB of NEs (20 µg) prepared from the same set of cells used for RNA-Seq verified the level of HNF4α and TCF4 protein. H, HepG2; Ca, CaCO2; mColon, mouse colon, lane 1 (40 µg), lane 2 (20 µg). (C) Total number of known genes dysregulated in the HNF4α2 and HNF4α8 HCT116 lines $\geq$ 1.5 fold change (FC) upon 24 h of DOX. (D) Venn diagram of the genes in (C). (E) As in (D) but with a 2-fold cut-off. (F) As in (D) but for the two sets of HNF4α8 samples (0.1 and 0.3 µg/mL DOX). Genes in both the 0.3 and the 0.1 µg/mL DOX samples were combined for HNF4α8 (D and E). (G) Heat map showing unique and common genes with the highest FC compared across the different lines and treatments. All values are statistically significant ($q \leq 0.05$), 0 is a placeholder for genes without any FC.

Dr. Joseph M. Dhahbi analyzed the HCT116 RNA-Seq data.

Fig. 3.4

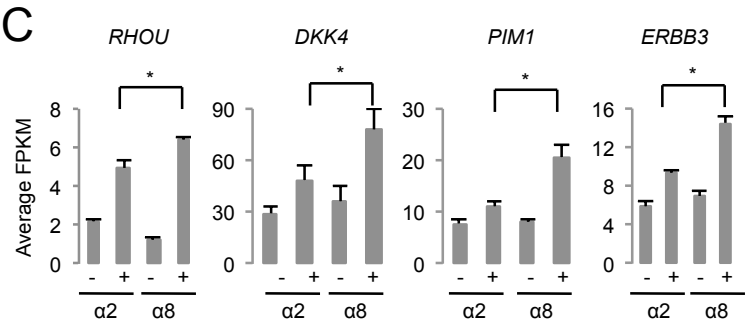
A



B

Line	Category	Genes
HNF4α2 (UP)	Cell Death	*HYAL1, TNFRSF1A, KLK8, HTRA2, MCL1, HMOX1, KRT8, CTSD, SFN, MAP3K11*
HNF4α8 (UP)	Cell Adhesion	*ALCAM, COL7A1, PKP2, GP1BB, BCAR1, COL27A1, CD58, COL6A1, NINJ1, LAMB1, CIB1*
HNF4α8 (DOWN)	Anti-apoptosis	*IER3, BDNF, ANXA1, FOXC1, NRG1, TNFSF18, MSX2*
HNF4α8 (More UP)	Growth promotion	***RHOU***, ***PIM1***, *ZDHHC9, AGPAT9,* ***DKK4***, ***ERBB3***, *CD58, ADM*
HNF4α2 (More UP)	Growth inhibition	***PLEKHO1***, *OSGIN1, ISG20, KLC4, LAMB2, FBXO2,* ***MYBPH***

C



D

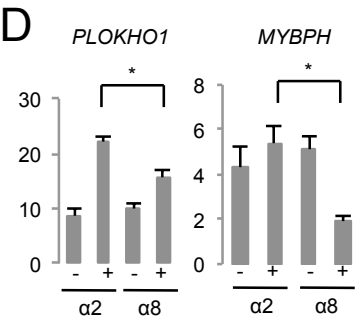


Figure 3.4 HNF4α2 and HNF4α8 regulate different biological processes in HCT116 cells. (A) Gene Ontology of genes up or down regulated by the HNF4α isoforms ($\geq$ 1.5 fold) that were not dysregulated in the parental line by DOX. (B) Select biological processes related to proliferation and differentiation with the corresponding genes. (C) Average FPKM from triplicate samples (mean$\pm$SD) of select growth promoting genes in the HNF4α2 and HNF4α8 lines -/+ DOX (0.3 μg/mL). (D) as in (C) but for growth inhibiting genes. All FC within a given line have $P \leq 0.05$, * $P \leq 0.05$ across cell lines.

Fig. 3.5

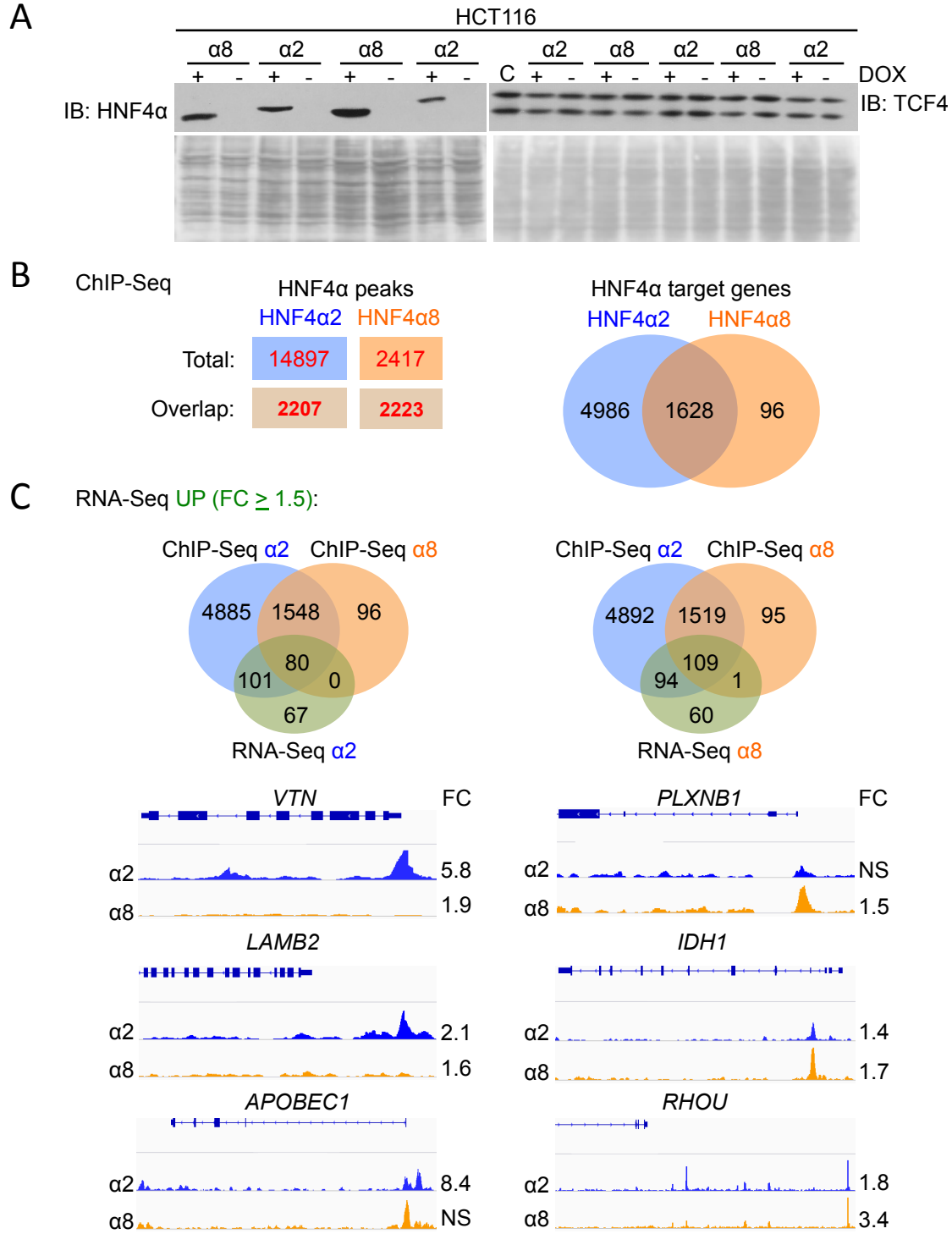

A
HCT116
α8 α2 α8 α2 C α2 α8 α2 α8 α2
+ - + - + - + - + - + - + - + - + - DOX
IB: HNF4α
IB: TCF4
B
ChIP-Seq
HNF4α peaks
HNF4α2 HNF4α8
Total: 14897 2417
Overlap: 2207 2223
HNF4α target genes
HNF4α2 HNF4α8
4986 1628 96
C
RNA-Seq UP (FC ≥ 1.5):
ChIP-Seq α2 ChIP-Seq α8
4885 1548 96
80
101 0
67
RNA-Seq α2
ChIP-Seq α2 ChIP-Seq α8
4892 1519 95
109
94 1
60
RNA-Seq α8
VTN
FC
α2 5.8
α8 1.9
PLXNB1
FC
α2 NS
α8 1.5
LAMB2
α2 2.1
α8 1.6
IDH1
α2 1.4
α8 1.7
APOBEC1
α2 8.4
α8 NS
RHOU
α2 1.8
α8 3.4

Figure 3.5 Integration of transcriptomic (RNA-Seq) with cistromic (ChIP-Seq) data in HCT116 HNF4α inducible cell lines. (A) IB of HNF4α and TCF4 in WCE (20 μg) 24 h after DOX (0.3 μg/mL) induction; shown are representative samples of those used in ChIP-Seq. (B) Venn diagram comparing ChIP peaks and nearby (-/+50 kbp) genes for HNF4α2 and HNF4α8. Overlapping HNF4α2 and HNF4α8 peaks were defined as those with 13 or more overlapping nucleotides. (C) Venn diagram comparing genes upregulated in RNA-Seq $\geq$ 1.5 fold to genes near ChIP-Seq peaks. Below, examples of genes with nearby HNF4α ChIP-Seq peaks and FC for HNF4α2 and HNF4α8.

Dr. Joseph M. Dhahbi analyzed the HCT116 ChIP-Seq data.

Fig. 3.6

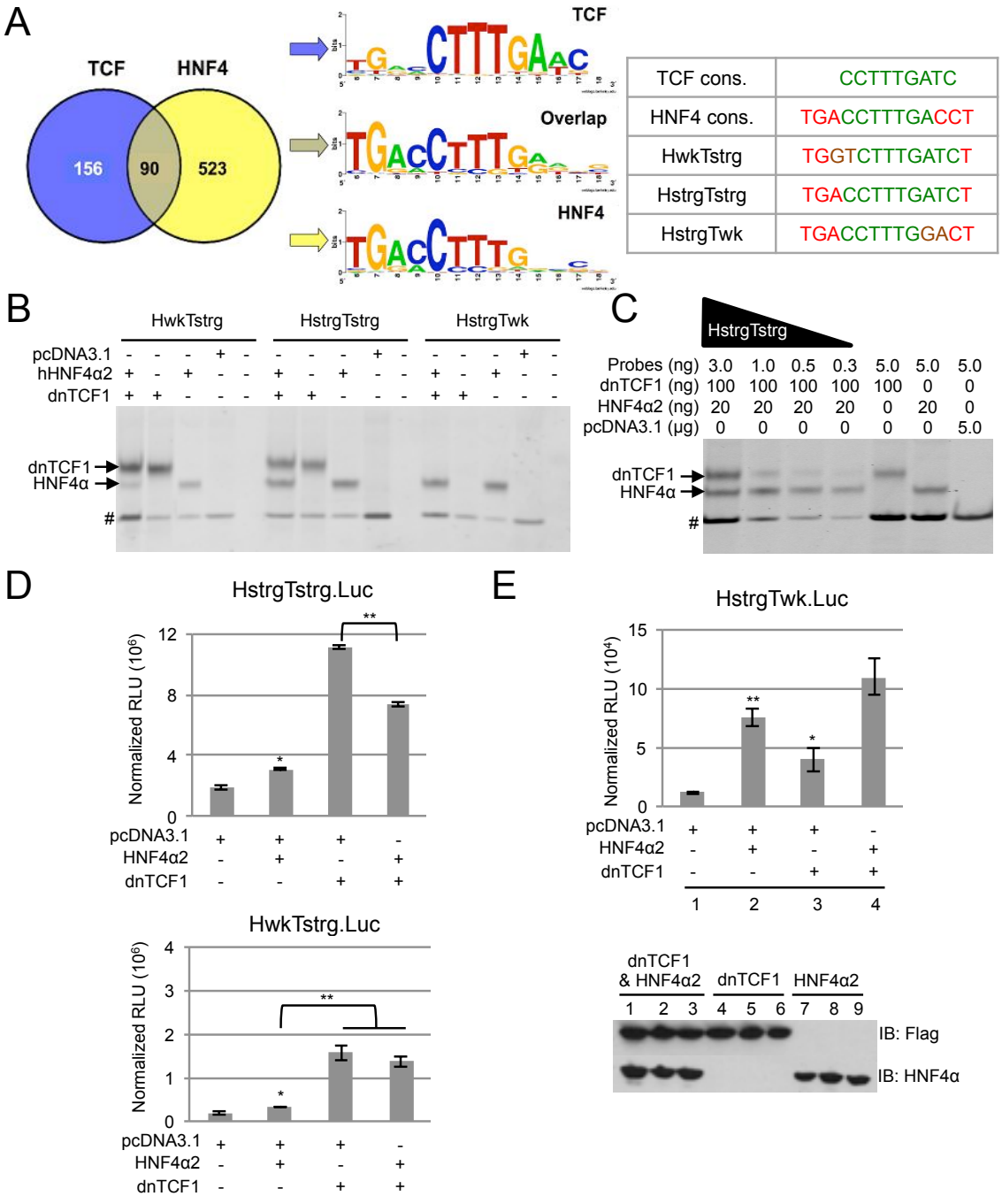
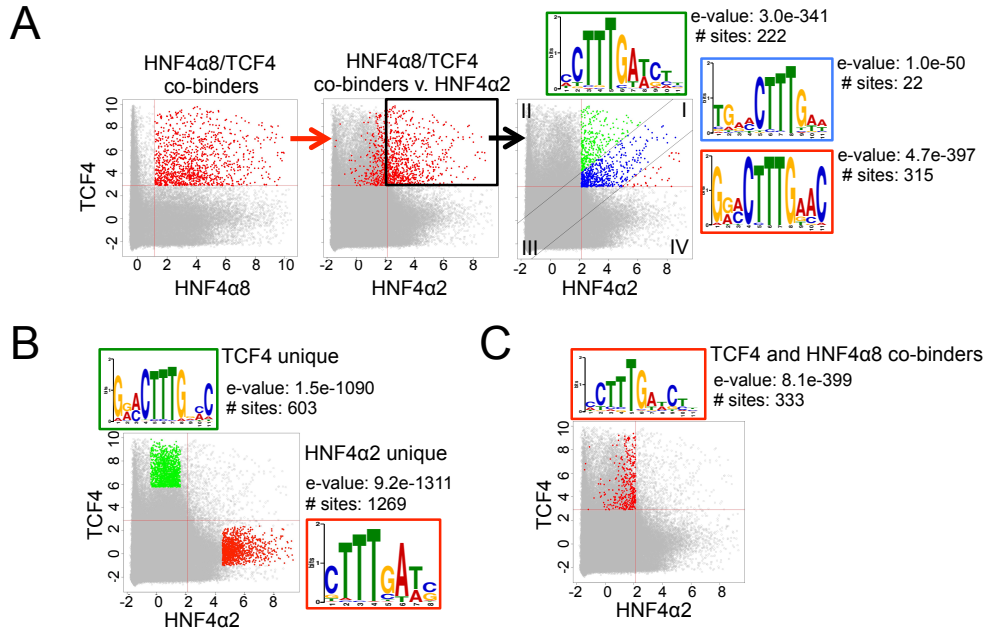

A
TCF
HNF4
156
90
523
TCF
Overlap
HNF4
TCF cons. CCTTTGATC
HNF4 cons. TGACCTTTGACCT
HwkTstrg TGGTCTTTGATCT
HstrgTstrg TGACCTTTGATCT
HstrgTwk TGACCTTTGGACT
B
HwkTstrg
HstrgTstrg
HstrgTwk
pcDNA3.1
hHNF4α2
dnTCF1
dnTCF1
HNF4α
#
C
HstrgTstrg
Probes (ng) 3.0 1.0 0.5 0.3 5.0 5.0 5.0
dnTCF1 (ng) 100 100 100 100 100 0 0
HNF4α2 (ng) 20 20 20 20 0 20 0
pcDNA3.1 (μg) 0 0 0 0 0 0 5.0
dnTCF1
HNF4α
#
D
HstrgTstrg.Luc
Normalized RLU (10⁶)
pcDNA3.1
HNF4α2
dnTCF1
HwkTstrg.Luc
Normalized RLU (10⁶)
pcDNA3.1
HNF4α2
dnTCF1
E
HstrgTwk.Luc
Normalized RLU (10⁴)
pcDNA3.1
HNF4α2
dnTCF1
1 2 3 4
dnTCF1 & HNF4α2
dnTCF1
HNF4α2
1 2 3 4 5 6 7 8 9
IB: Flag
IB: HNF4α

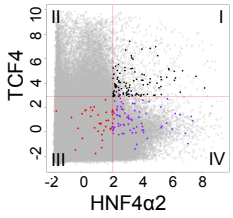
Figure 3.6 Overlapping DNA binding specificity of HNF4α and TCF1. (A) Venn diagram of the number of unique DNA sequences bound by TCF (dnTCF1) and HNF4α (HNF4α2) in PBM2. Center, position weight matrices (PWM) showing the motif derived from the sequences bound only by TCF, only by HNF4, or by both (overlap). Right, select DNA sequences with different binding strengths (strg, strong; wk, weak) for HNF4α2 (H) and dnTCF1 (T) identified in the PBM and aligned to the consensus sequence (cons.) of TCF and HNF4α. Nucleotides are color-coded to match the consensus sequence for TCF (green) and HNF4α (red), which is typically presented in the reverse complement, AGGTCAaAGGTCA. (B) Gel shift assay with probes containing the indicated motifs as defined in (A) and NE transfected COS-7 cells expressing human HNF4α2 or dnTCF1. #, non-specific band. (C) Gel shift as in (B) but with decreasing amount of probe containing HstrgTstrg. (D) Luciferase assay of transiently transfected HEK293T cells with the indicated expression vectors (80 ng) and reporter constructs (0.5 μg) containing a single binding motif driving expression of the luciferase gene (Luc) from a minimal promoter. Shown are the mean$\pm$SD of relative light units (RLU) normalized to β-gal activity of triplicate samples from one of four independent experiments. *$P < 0.003$ for HNF4α2 vs. pcDNA3.1 (top and bottom), **$P \leq 0.0004$ for other comparisons as indicated. (E) Luciferase assay as in (D) but with the HstrgTwk element. *$P \leq 0.02$ for condition 3 compared to 1, 2, and 4; **$P < 0.008$ for condition 2 compared to 1 and 4. Lower panel, IB of WCE (20 μg total protein) showing appropriate and consistent expression of dnTCF1 (Flag-tagged) and HNF4α2 in the transfected cells.

Dr. Eugene Bolotin performed and analyzed the PBM data (Fig. 3.6A). Dr. Nathan P. Hoverter provided extracts to put on the PBM.

Fig. 3.7



D



	HNF4α/TCF4 cobinders to:	
Altered NT in	**Both nonbinders**	**HNF4α only binders**
CTTTG	22	10
CTTTG flank	6	17
No CTTTG	7	45
Total:	35	72

Altered NT in:		**PBM score**		**Z-score**	
CTTTG	**sequence**	**TCF4**	**HNF4α**	**TCF4**	**HNF4α**
rs2297887	ggtaaacgtggactTtggtccggcatggc	3.1	3.7	5.8	5.8
CISD1	ggtaaacgtggactCtggtccggcatggc	0.4	1.7		
rs77304671	accagcatgtaaccTttgtaccaattgat	3.1	4.1	3.5	6.8
MS4A4A	accagcatgtaaccAttgtaccaattgat	2.0	0.7		
rs76390860	gtctctgtgaccttTgatcgctctgctcc	6.8	4.6	2.7	3.4
NPY	gtctctgtgaccttCgatcgctctgctcc	-1.3	1.4		
rs1273578	ggtagagctggactTtgaacccacatctt	4.6	7.6	6.0	1.3
COQ9	ggtagagctggactGtgaacccacatctt	-0.2	2.5		
rs77344780	acccccgcaatgctTtgttgtagttacac	2.9	2.2	4.9	0.2
C9orf9	acccccgcaatgctCtgttgtagttacac	0.1	2.1		
rs75862633	aggaagctgaacttTgatcttcttggttg	6.8	4.0	4.6	2.3
AQP4	aggaagctgaacttCgatcttcttggttg	-0.1	2.1		
CTTTG flank	**sequence**	**TCF4**	**HNF4α**	**TCF4**	**HNF4α**
rs10972567	gcctgagaactttgAtccctttccattg	5.0	2.4	10.1	3.2
CREB3	gcctgagaactttgCtccctttccattg	-0.2	1.4		
rs35436776	ccagtgctctgcagTtcaaaggtccctga	6.7	3.0	4.4	6.5
RHOU	ccagtgctctgcagCtcaaaggtccctga	1.3	0.7		
rs112643302	tagtgcctctttgaAcccgcaggtcatg	3.9	2.2	2.4	4.4
GRID1-AS1	tagtgcctctttgaCacccgcaggtcatg	0.3	0.9		
rs1666561	cctttgccctgtgaTctttgttggaccct	3.7	5.3	8.4	0.2
MIR548A3	cctttgccctgtgaCctttgttggaccct	0.7	5.7		
rs9501868	ctgttgacattattTcaaaggtcagaatt	4.2	4.0	7.6	0.1
C6orf195	ctgttgacattattCcaaaggtcagaatt	0.1	4.0		
rs28640963	gcaatttgactttgTactttgtctcacta	3.5	2.7	4.0	1.0
PCTP	gcaatttgactttgCactttgtctcacta	0.4	2.2		
No CTTTG	**sequence**	**TCF4**	**HNF4α**	**TCF4**	**HNF4α**
rs57637384	tttagtttgtctgcGgttgtttttttgtt	3.0	2.2	3.7	0.2
SPTBN1	tttagtttgtctgcAgttgtttttttgtt	1.6	2.1		
rs73007638	gggtcactggttgtTtgttcttggactgt	3.4	3.3	4.8	0.1
NCKAP5	gggtcactggttgtCtgttcttggactgt	1.0	3.4		

Figure 3.7 High throughput DNA binding analysis of TCF4 and HNF4α. (A) Scatter plots of DNA binding scores for TCF4 and the two HNF4α isoforms on 1,000,000 spots of DNA in the PBM. Each of the 250,000 spots represents the average score of four replicates on the array. Red lines indicate threshold of binding set at 2 SD above background. Sequences bound by both HNF4α8 and TCF4 are marked in red (left) and then identified in the TCF4/HNF4α2 plot (center). The red spots in quadrant I that bind all three TFs were divided into three groups with different binding affinities: TCF4 > HNF4α (green), TCF4 ≈ HNF4α (blue), HNF4α > TCF4 (red). PWMs were generated by MEME for each color group in quadrant I. Only the top scoring PWM is shown along with the e-value and number of sites/sequences used to generate the PWM. (B) PWMs generated by MEME as in (A, top right) but for the green (TCF4 unique) and red spots (HNF4α2 unique) highlighted in the plot. (C) as in (B) but for sequences bound by HNF4α8 and TCF4 but not HNF4α2. (D) DNA sequences in quadrant I for which the alternate SNP allele significantly reduced DNA binding (black spots, 107 total). Red spots, SNP resulted in non binders for both TCF4 and HNF4α2. Purple spots, SNP resulted in non binder for TCF4 only (HNF4α only binders). Bottom left, the 107 black spots are grouped into six categories depending on whether the SNP is in the CTTTG core, within a CTTTG flank or there is no CTTTG present in the sequence. Right, examples of SNPs in each category along with the PBM (binding) score and Z-score (difference between the two alleles). rs number of SNPs are color coded to match the spots in the scatter plot.

Dr. Bin Fang designed the PBM. Dr. Nina Titova performed the PBM. Dr. Nate P. Hoverter provided the extracts to put on the PBM. Jonathan Deans analyzed the PBM data and performed motif analysis.

Fig. 3.8

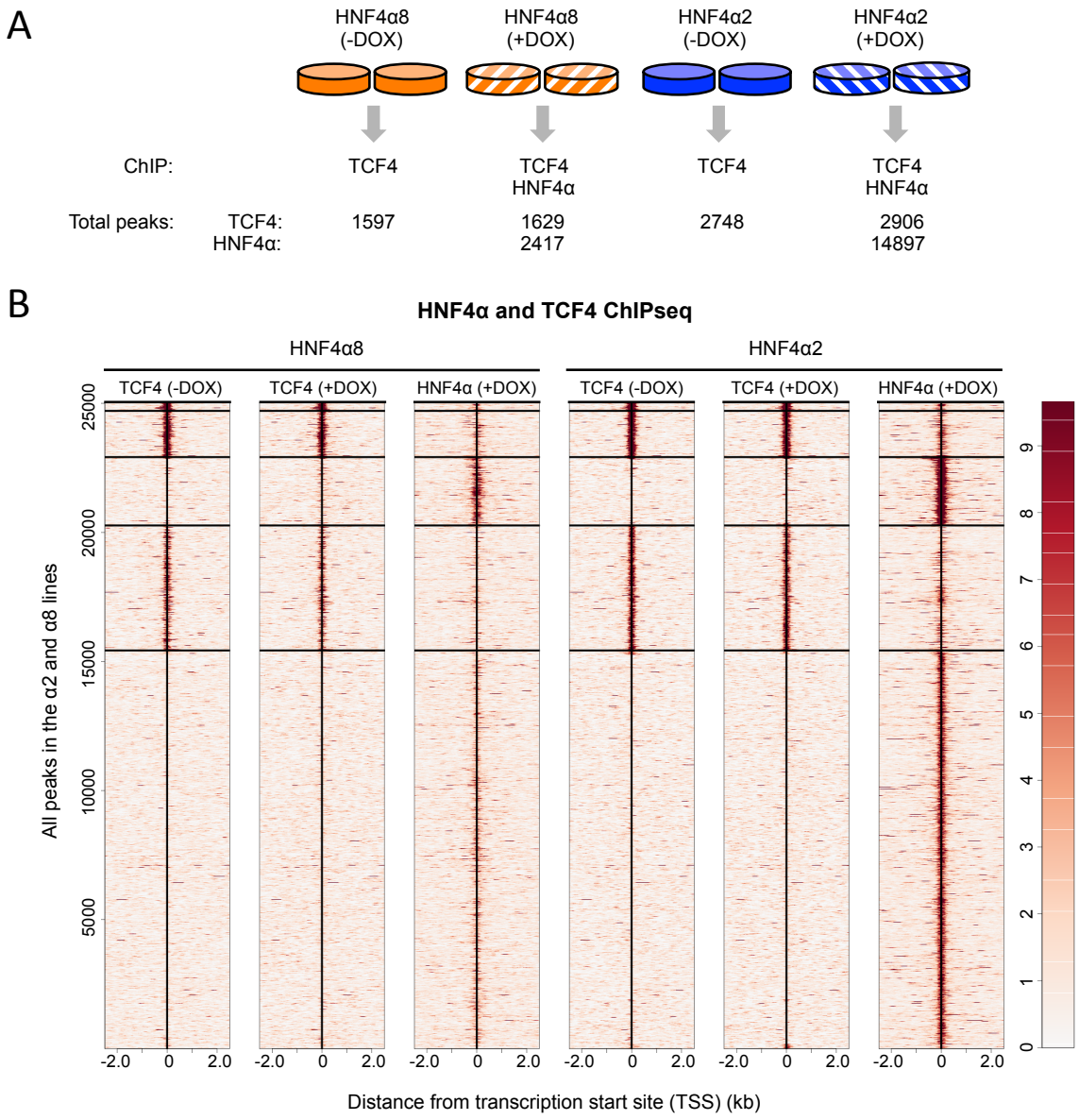
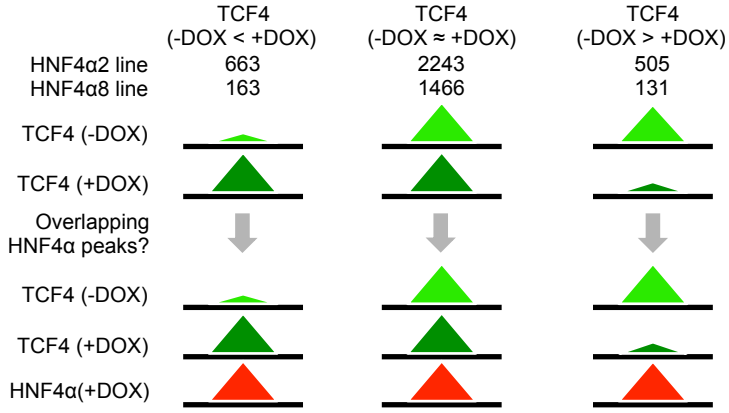

A
HNF4α8 (-DOX)
HNF4α8 (+DOX)
HNF4α2 (-DOX)
HNF4α2 (+DOX)
ChIP:
TCF4
TCF4 HNF4α
TCF4
TCF4 HNF4α
Total peaks:
TCF4:
HNF4α:
1597
1629
2417
2748
2906
14897
B
HNF4α and TCF4 ChIPseq
HNF4α8
HNF4α2
TCF4 (-DOX)
TCF4 (+DOX)
HNF4α (+DOX)
All peaks in the α2 and α8 lines
25000
20000
15000
10000
5000
-2.0
0
2.0
Distance from transcription start site (TSS) (kb)
0
1
2
3
4
5
6
7
8
9

Figure 3.8 ChIP-Seq analysis of TCF4 and HNF4α2/α8 in HCT116 cells. (A) Schematic of samples analyzed in ChIP-Seq with the total number of HNF4α (+DOX) and TCF4 (-/+DOX) peaks in the HNF4α2 and α8 lines. (B) Heatmap generated from Cistrome showing the global view of binding by HNF4α and TCF4 in the α2 and α8 lines (-/+ DOX, as indicated) relative to TSS (+1, "0"). Right, scale from 0 to 10, least to most enriched sequences.

Dr. Joseph M. Dhahbi generated the heatmap (Fig. 3.8B).

Fig. 3.9

A

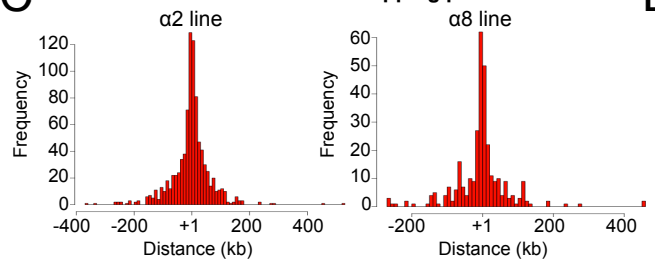


B

Overlapping HNF4α and TCF4 peaks	HNF4α recruits TCF4	HNF4α colocalizes with TCF4	HNF4α competes with TCF4	Total HNF4α peaks
HNF4α2	41 (0.18%)	485 (2.2%)	43 (0.19%)	22471
HNF4α8	75 (1.4%)	126 (2.4%)	0 (0%)	5281
HNF4α2 + HNF4α8	13 (0.25%)	68 (1.3%)	0 (0%)	5285

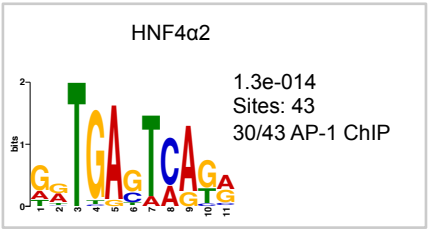
C

Distance of HNF4α/TCF4 overlapping peaks to TSS



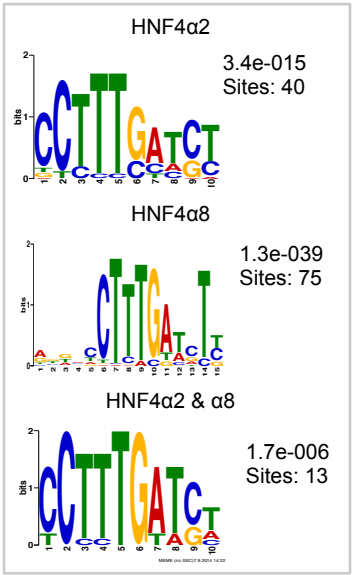
E

Competition of TCF4 by HNF4α:



D

Recruitment of TCF4 by HNF4α:



Co-localization of TCF4 with HNF4α:

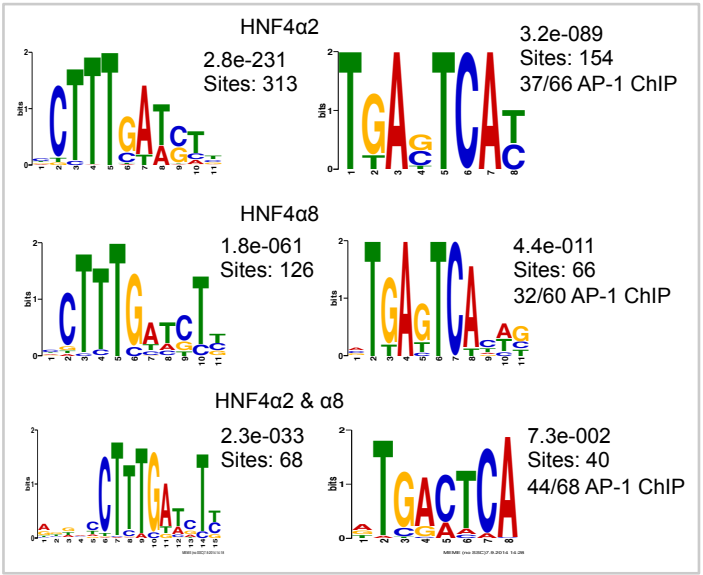


Figure 3.9 TCF4 and HNF4α exhibit different categories of overlapping ChIP-Seq peaks. (A) Schematic showing three distinct categories of TCF4 peaks +/- DOX and the number of peaks in the HNF4α2 and HNF4α8 lines in each category; each category was queried for an overlapping HNF4α peak. (B) Absolute number of HNF4α and TCF4 overlapping peaks in each category (recruited, colocalized, and competed) and percentage of overlapping peaks compared to total HNF4α peaks. HNF4α2+HNF4α8: HNF4α2 peaks overlap (13 NTs or more) with HNF4α8 peaks. (C) Histograms of all overlapping peaks (recruited, colocalized, and competed) for the HNF4α2 and HNF4α8 lines plotted relative to TSS (+1) using the R Bioconductor package, ChIPpeakAnno. PWM generated by MEME of HNF4α2/α8 recruited and co-localized peaks (D) and HNF4α2 competed peaks (E). Peaks in each category containing a TGAxTCA motif were manually cross-referenced to FosL1 and JunD (AP-1) ChIP-Seq in the HCT116 ENCODE database using Integrative Genomics Viewer (IGV). Given is the number of overlapping TCF4/HNF4α peaks that overlap with both FosL1 and JunD peaks divided by the total number of peaks examined.

Dr. Joseph M. Dhahbi analyzed the ChIP-Seq data to look for differential HNF4α and TCF4 binding peaks in the presence and absence of DOX (Fig. 3.9A) and generated the histogram graphs (Fig. 3.9C). Jonathan Deans performed motif analysis for the recruitment of TCF4 by HNF4α2 (Fig. 3.9D, left).

Fig. 3.10

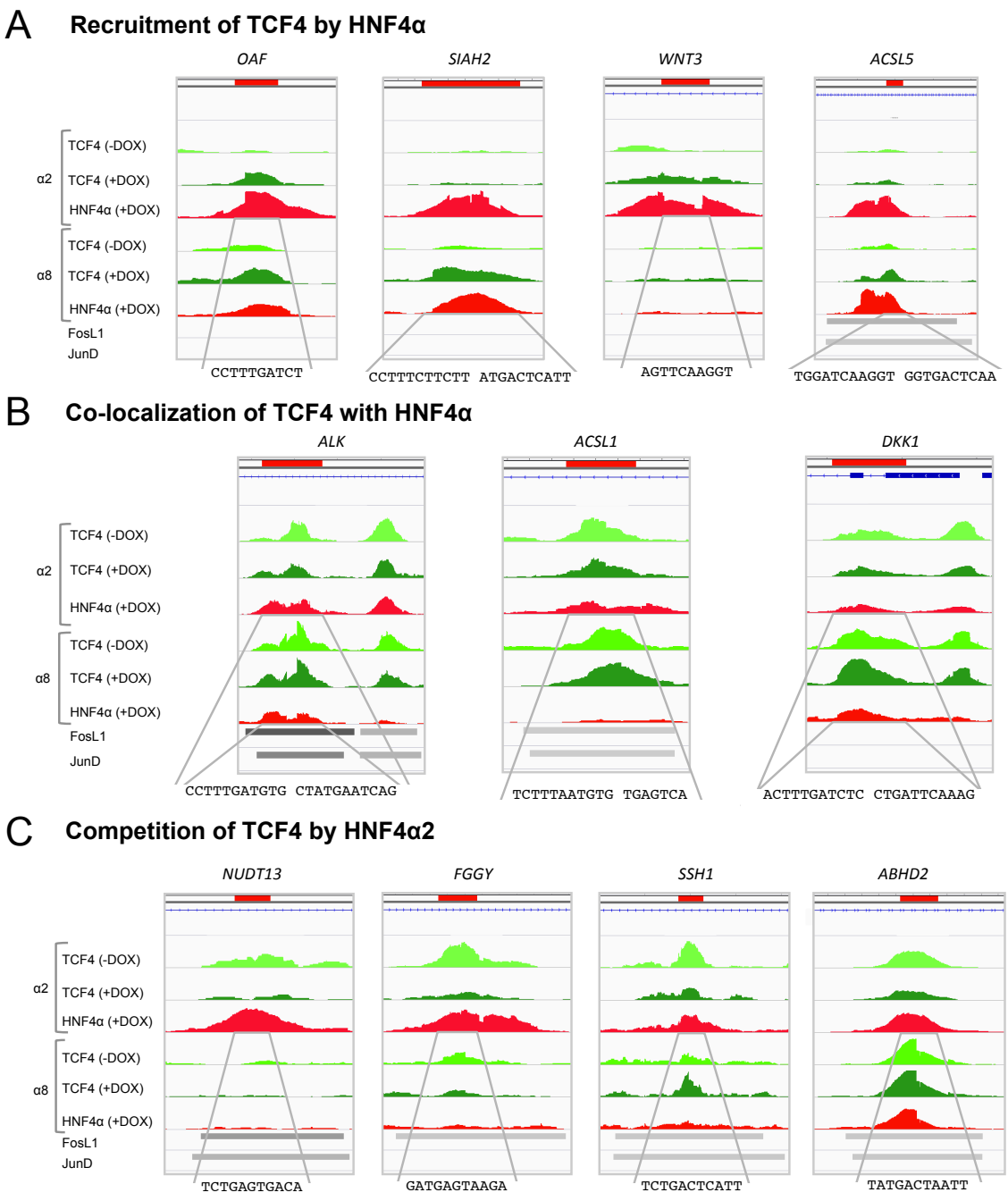

A
Recruitment of TCF4 by HNF4α
OAF
SIAH2
WNT3
ACSL5
α2
α8
TCF4 (-DOX)
TCF4 (+DOX)
HNF4α (+DOX)
FosL1
JunD
CCTTTGATCT
CCTTTCTTCTT ATGACTCATT
AGTTCAAGGT
TGGATCAAGGT GGTGACTCAA
B
Co-localization of TCF4 with HNF4α
ALK
ACSL1
DKK1
CCTTTGATGTG CTATGAATCAG
TCTTTAATGTG TGAGTCA
ACTTTGATCTC CTGATTCAAAG
C
Competition of TCF4 by HNF4α2
NUDT13
FGGY
SSH1
ABHD2
TCTGAGTGACA
GATGAGTAAGA
TCTGACTCATT
TATGACTAATT

Figure 3.10 Examples of HNF4α and TCF4 recruiting, co-localizing, and competing ChIP-Seq peaks. Snapshots from IGV of HNF4α (+DOX) and TCF4 (-/+DOX) ChIP-Seq peaks from HCT116 HNF4α2 and HNF4α8 inducible lines and FosL1 and JunD ChIP-Seq peaks in HCT116 from ENCODE with corresponding RefSeq genes. The CTTTG and/or TGAxTCA motifs with surrounding NT sequence from the peaks are given below. Examples of (A) recruiting peaks, (B) co-localizing peaks and (C) competing peaks.

Table 3.1 Genes with overlapping HNF4α and TCF4 peaks dysregulated in RNAseq[1]

Competition			**Co-localization**		
HNF4α2	**Up**	**Down**	**HNF4α2**	**Up**	**Down**
(33 genes total)	*FGGY*	*CPOX*	(324 genes total)	***ACSL1***	*ATP6AP1L*
	MICAL2	***LRRTM4***		*AMOTL1*	***DKK1****
		SSH1		***ASS1***	*ETV1*
		TIAM1		*CD59*	***IER3***
Recruitment				***COBL***	*KIAA1430*
HNF4α2	**Up**	**Down**		***CTSB***	*LRP6*
(33 genes total)	*ALDH3A2*			*EVPL*	***MAP2***
	EPDR1			*F2RL1*	*MET*
	FAM169A			*FAM129B*	*MID1*
	NFE2L3			***GLCE***	*PCDH7*
	OAF			***GRK5***	*PPP3CA*
	PFKP			*HMGCL*	*PRKACB*
	PRKAG2			***ITGB5***	*RAP2B*
	WNT3			*ITPKA*	*SASS6*
HNF4α8	**Up**	**Down**		*LDLR*	*SCG2*
(53 genes total)	***ACSL5***	***C16orf45***		*NEK6*	*STEAP1*
	*CBLB**	*CENPF*		***OSBP***	*SYTL3*
	CLEC16A	*EFHD2*		***PDE2A***	*TEAD1*
	EPDR1	*PHGDH*		*PFKP*	*TNIK*
	FAM169A			*PTPRH*	*TSPANS*
	FRMD6			***SERINC2***	*WDR62*
	GPD2			*SERPINB1*	
	HDGF			***SLC2A1***	
	IRF2BP2		**HNF4α8**	**UP**	**DOWN**
	KIAA1671		(84 genes total)	***ABHD2***	***DKK1****
	LRP5			*AMN1*	***IER3***
	NSMCE1			*AMOTL1*	*MOSPD1*
	PPAP2B			*CDC14B*	*MSRB3*
	PRKAG2			*ETFB*	*TEAD1*
	SIAH2			***GRK5***	***TIAM1***
	SLC2A1			***LPCAT3***	*ZNRF3*
	*SPTBN1**			*IRF2BP2*	
				MICAL2	
				MLEC	
				SERINC2	
				SLC35D2	

See Fig. 3.9 for different categories of colocalizing, recruiting, and competing HNF4α and TCF4 peaks. Total genes, with a TSS within 50 kb of the peak called by Cisgenome, in a given category are indicated.

Bold, genes with FC ≥ 1.5 or more. Non-bold, genes with FC ≥ 1.2-1.4

*Genes identified in Fig. 3.7 as containing a binding site bound by both HNF4α and TCF4.

Fig. 3.11

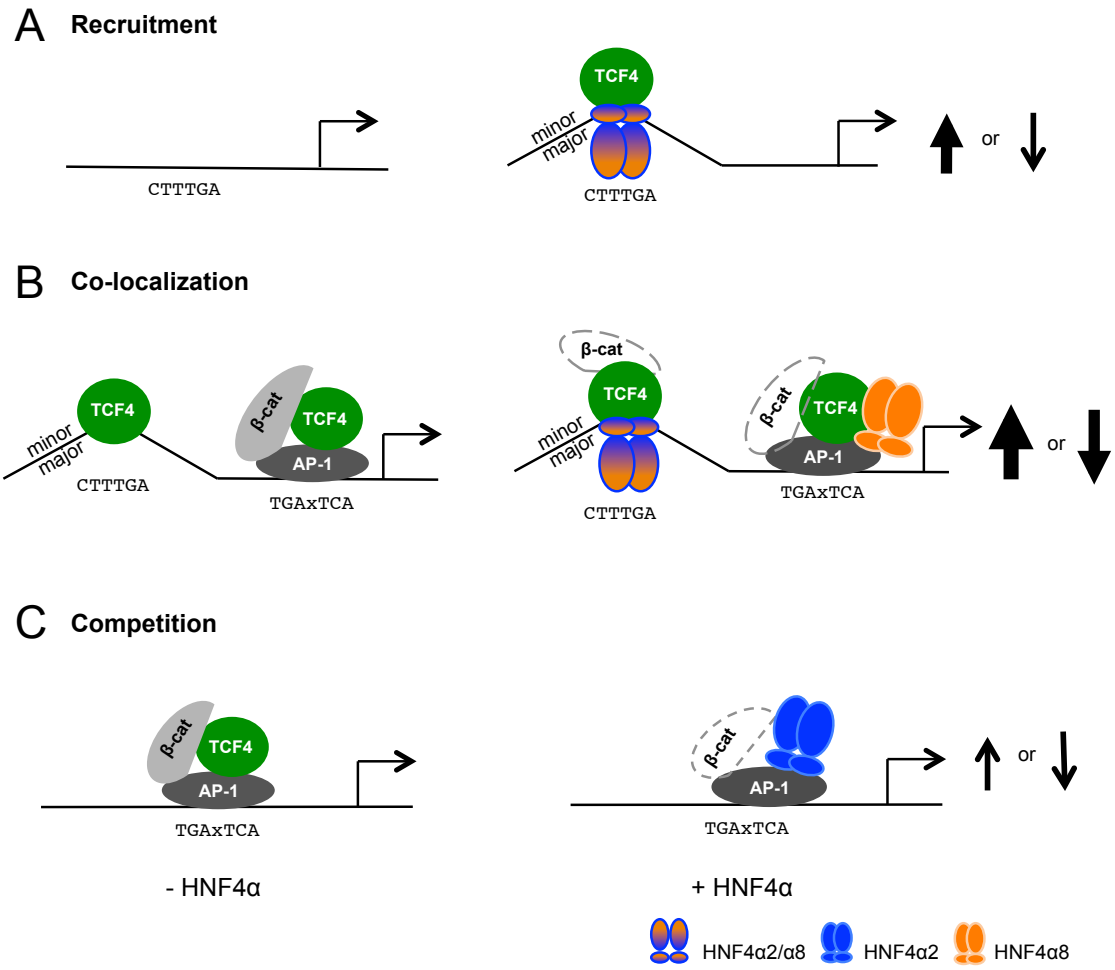

A
Recruitment
CTTTGA
TCF4
minor
major
CTTTGA
or
B
Co-localization
TCF4
minor
major
CTTTGA
β-cat
TCF4
AP-1
TGAxTCA
β-cat
TCF4
minor
major
CTTTGA
β-cat
TCF4
AP-1
TGAxTCA
or
C
Competition
β-cat
TCF4
AP-1
TGAxTCA
- HNF4α
β-cat
AP-1
TGAxTCA
or
+ HNF4α
HNF4α2/α8
HNF4α2
HNF4α8

Figure 3.11 Schematic model of HNF4α, TCF4 and AP-1 interplay in HCT116 human colon cancer cells. Three different categories of interplay (recruitment, co-localization, and competition) between HNF4α and TCF4 were identified in this study. Shown are potential interactions with AP-1 (Fos/Jun heterodimer) based on ChIP data and the literature. Dashed line, presence of β-catenin in the complex is not known. See text for details.

References

Ahn SH, Shah YM, Inoue J, Morimura K, Kim I, Yim S, Lambert G, Kurotani R, Nagashima K, Gonzalez FJ et al. 2008. Hepatocyte nuclear factor 4alpha in the intestinal epithelial cells protects against inflammatory bowel disease. *Inflammatory bowel diseases* **14**: 908-920.

Babeu JP, Darsigny M, Lussier CR, Boudreau F. 2009. Hepatocyte nuclear factor 4alpha contributes to an intestinal epithelial phenotype in vitro and plays a partial role in mouse intestinal epithelium differentiation. *American journal of physiology Gastrointestinal and liver physiology* **297**: G124-134.

Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings / International Conference on Intelligent Systems for Molecular Biology ; ISMB International Conference on Intelligent Systems for Molecular Biology* **2**: 28-36.

Berenguer-Daize C, Boudouresque F, Bastide C, Tounsi A, Benyahia Z, Acunzo J, Dussault N, Delfino C, Baeza N, Daniel L et al. 2013. Adrenomedullin blockade suppresses growth of human hormone-independent prostate tumor xenograft in mice. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**: 6138-6150.

Blahnik KR, Dou L, O'Geen H, McPhillips T, Xu X, Cao AR, Iyengar S, Nicolet CM, Ludascher B, Korf I et al. 2010. Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data. *Nucleic acids research* **38**: e13.

Bolotin E, Liao H, Ta TC, Yang C, Hwang-Verslues W, Evans JR, Jiang T, Sladek FM. 2010. Integrated approach for the identification of human hepatocyte nuclear factor 4alpha target genes using protein binding microarrays. *Hepatology* **51**: 642-653.

Bonzo JA, Ferry CH, Matsubara T, Kim JH, Gonzalez FJ. 2012. Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4alpha in adult mice. *The Journal of biological chemistry* **287**: 7345-7356.

Bottomly D, Kyler SL, McWeeney SK, Yochum GS. 2010. Identification of {beta}-catenin binding regions in colon cancer cells using ChIP-Seq. *Nucleic acids research* **38**: 5735-5745.

Briancon N, Weiss MC. 2006. In vivo role of the HNF4alpha AF-1 activation domain revealed by exon swapping. *The EMBO journal* **25**: 1253-1262.

Cancer Genome Atlas N. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**: 330-337.

Cattin AL, Le Beyec J, Barreau F, Saint-Just S, Houllier A, Gonzalez FJ, Robine S, Pincon-Raymond M, Cardot P, Lacasa M et al. 2009. Hepatocyte nuclear factor 4alpha, a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. *Molecular and cellular biology* **29**: 6294-6308.

Chahar S, Gandhi V, Yu S, Desai K, Cowper-Sal Lari R, Kim Y, Perekatt AO, Kumar N, Thackray JK, Musolf A et al. 2014. Chromatin Profiling Reveals Regulatory Network Shifts and a Protective Role for Hepatocyte Nuclear Factor 4alpha during Colitis. *Molecular and cellular biology* **34**: 3291-3304.

Chandra V, Huang P, Potluri N, Wu D, Kim Y, Rastinejad F. 2013. Multidomain integration in the structure of the HNF-4alpha nuclear receptor complex. *Nature* **495**: 394-398.

Chartier FL, Bossu JP, Laudet V, Fruchart JC, Laine B. 1994. Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. *Gene* **147**: 269-272.

Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CL, Chan C, Dent OF, Clarke SJ, Robertson GR et al. 2012. Src tyrosine kinase phosphorylation of nuclear receptor HNF4alpha correlates with isoform-specific loss of HNF4alpha in human colon cancer. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 2302-2307.

Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, Darnell JE, Jr. 1994. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes & development* **8**: 2466-2477.

Clevers H, Nusse R. 2012. Wnt/beta-catenin signaling and disease. *Cell* **149**: 1192-1205.

Colletti M, Cicchini C, Conigliaro A, Santangelo L, Alonzi T, Pasquini E, Tripodi M, Amicone L. 2009. Convergence of Wnt signaling on the HNF4alpha-driven transcription in controlling liver zonation. *Gastroenterology* **137**: 660-672.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome research* **14**: 1188-1190.

Darsigny M, Babeu JP, Dupuis AA, Furth EE, Seidman EG, Levy E, Verdu EF, Gendron FP, Boudreau F. 2009. Loss of hepatocyte-nuclear-factor-4alpha affects colonic ion transport and causes chronic inflammation resembling inflammatory bowel disease in mice. *PloS one* **4**: e7609.

Eeckhoute J, Moerman E, Bouckenooghe T, Lukoviak B, Pattou F, Formstecher P, Kerr-Conte J, Vandewalle B, Laine B. 2003. Hepatocyte nuclear factor 4 alpha isoforms originated from the P1 promoter are expressed in human pancreatic

beta-cells and exhibit stronger transcriptional potentials than P2 promoter-driven isoforms. *Endocrinology* **144**: 1686-1694.

Erdmann S, Senkel S, Arndt T, Lucas B, Lausen J, Klein-Hitpass L, Ryffel GU, Thomas H. 2007. Tissue-specific transcription factor HNF4alpha inhibits cell proliferation and induces apoptosis in the pancreatic INS-1 beta-cell line. *Biological chemistry* **388**: 91-106.

Fang B, Mane-Padros D, Bolotin E, Jiang T, Sladek FM. 2012. Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. *Nucleic acids research* **40**: 5343-5356.

Frietze S, Wang R, Yao L, Tak YG, Ye Z, Gaddis M, Witt H, Farnham PJ, Jin VX. 2012. Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. *Genome biology* **13**: R52.

Furuta H, Iwasaki N, Oda N, Hinokio Y, Horikawa Y, Yamagata K, Yano N, Sugahiro J, Ogata M, Ohgawara H et al. 1997. Organization and partial sequence of the hepatocyte nuclear factor-4 alpha/MODY1 gene and identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* **46**: 1652-1657.

Garrison WD, Battle MA, Yang C, Kaestner KH, Sladek FM, Duncan SA. 2006. Hepatocyte nuclear factor 4alpha is essential for embryonic development of the mouse colon. *Gastroenterology* **130**: 1207-1220.

Giese K, Cox J, Grosschedl R. 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69**: 185-195.

Gougelet A, Torre C, Veber P, Sartor C, Bachelot L, Denechaud PD, Godard C, Moldes M, Burnol AF, Dubuquoy C et al. 2014. T-cell factor 4 and beta-catenin chromatin occupancies pattern zonal liver metabolism in mice. *Hepatology* **59**: 2344-2357.

Hanse EA, Mashek DG, Becker JR, Solmonson AD, Mullany LK, Mashek MT, Towle HC, Chau AT, Albrecht JH. 2012. Cyclin D1 inhibits hepatic lipogenesis via repression of carbohydrate response element binding protein and hepatocyte nuclear factor 4alpha. *Cell cycle* **11**: 2681-2690.

Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, Steele A, Njolstad PR, Ellard S, Hattersley AT. 2008. The diabetic phenotype in HNF4A mutation carriers is moderated by the expression of HNF4A isoforms from the P1 promoter during fetal development. *Diabetes* **57**: 1745-1752.

Hatziapostolou M, Polytarchou C, Aggelidou E, Drakaki A, Poultsides GA, Jaeger SA, Ogata H, Karin M, Struhl K, Hadzopoulou-Cladaras M et al. 2011. An HNF4alpha-miRNA inflammatory feedback circuit regulates hepatocellular oncogenesis. *Cell* **147**: 1233-1247.

Hatzis P, van der Flier LG, van Driel MA, Guryev V, Nielsen F, Denissov S, Nijman IJ, Koster J, Santo EE, Welboren W et al. 2008. Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. *Molecular and cellular biology* **28**: 2732-2744.

Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ. 2001. Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Molecular and cellular biology* **21**: 1393-1403.

Hosono Y, Yamaguchi T, Mizutani E, Yanagisawa K, Arima C, Tomida S, Shimada Y, Hiraoka M, Kato S, Yokoi K et al. 2012. MYBPH, a transcriptional target of TTF-1, inhibits ROCK1, and reduces cell motility and metastasis. *The EMBO journal* **31**: 481-493.

Hoverter NP, Ting JH, Sundaresh S, Baldi P, Waterman ML. 2012. A WNT/p21 circuit directed by the C-clamp, a sequence-specific DNA binding domain in TCFs. *Molecular and cellular biology* **32**: 3648-3662.

Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**: 44-57.

Hwang-Verslues WW, Sladek FM. 2008. Nuclear receptor hepatocyte nuclear factor 4alpha1 competes with oncoprotein c-Myc for control of the p21/WAF1 promoter. *Molecular endocrinology* **22**: 78-90.

Ji H, Jiang H, Ma W, Wong WH. 2011. Using CisGenome to analyze ChIP-chip and ChIP-seq data. *Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]* **Chapter 2**: Unit2 13.

Jiang G, Lee U, Sladek FM. 1997. Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. *Molecular and cellular biology* **17**: 6546-6554.

Jiang G, Nepomuceno L, Hopkins K, Sladek FM. 1995. Exclusive homodimerization of the orphan receptor hepatocyte nuclear factor 4 defines a new subclass of nuclear receptors. *Molecular and cellular biology* **15**: 5131-5143.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome research* **12**: 996-1006.

Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B, Clevers H. 1997. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**: 1784-1787.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**: R25.

Li AG, Piluso LG, Cai X, Gadd BJ, Ladurner AG, Liu X. 2007. An acetylation switch in p53 mediates holo-TFIID recruitment. *Molecular cell* **28**: 408-421.

Liew CG, Draper JS, Walsh J, Moore H, Andrews PW. 2007. Transient and stable transgene expression in human embryonic stem cells. *Stem cells* **25**: 1521-1528.

Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, Shin H, Wong SS, Ma J, Lei Y et al. 2011. Cistrome: an integrative platform for transcriptional regulation studies. *Genome biology* **12**: R83.

Love JJ, Li X, Case DA, Giese K, Grosschedl R, Wright PE. 1995. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **376**: 791-795.

Lu P, Liu J, Melikishvili M, Fried MG, Chi YI. 2008. Crystallization of hepatocyte nuclear factor 4 alpha (HNF4 alpha) in complex with the HNF1 alpha promoter element. *Acta crystallographica Section F, Structural biology and crystallization communications* **64**: 313-317.

Maeda Y, Hwang-Verslues WW, Wei G, Fukazawa T, Durbin ML, Owen LB, Liu X, Sladek FM. 2006. Tumour suppressor p53 down-regulates the expression of the human hepatocyte nuclear factor 4alpha (HNF4alpha) gene. *The Biochemical journal* **400**: 303-313.

Maeda Y, Seidel SD, Wei G, Liu X, Sladek FM. 2002. Repression of hepatocyte nuclear factor 4alpha tumor suppressor p53: involvement of the ligand-binding domain and histone deacetylase activity. *Molecular endocrinology* **16**: 402-410.

Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, Kinzler KW. 1997. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* **275**: 1787-1790.

Nakhei H, Lingott A, Lemm I, Ryffel GU. 1998. An alternative splice variant of the tissue specific transcription factor HNF4alpha predominates in undifferentiated murine cell types. *Nucleic acids research* **26**: 497-504.

Nateri AS, Spencer-Dene B, Behrens A. 2005. Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. *Nature* **437**: 281-285.

Ning BF, Ding J, Yin C, Zhong W, Wu K, Zeng X, Yang W, Chen YX, Zhang JP, Zhang X et al. 2010. Hepatocyte nuclear factor 4 alpha suppresses the development of hepatocellular carcinoma. *Cancer research* **70**: 7640-7651.

Nishita M, Tomizawa C, Yamamoto M, Horita Y, Ohashi K, Mizuno K. 2005. Spatial and temporal regulation of cofilin activity by LIM kinase and Slingshot is critical for directional cell migration. *The Journal of cell biology* **171**: 349-359.

Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK et al. 2004. Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**: 1378-1381.

Oliveros JC. 2007. VENNY. An interactive tool for comparing lists with Venn Diagrams.

Oshima T, Kawasaki T, Ohashi R, Hasegawa G, Jiang S, Umezu H, Aoyagi Y, Iwanari H, Tanaka T, Hamakubo T et al. 2007. Downregulated P1 promoter-driven hepatocyte nuclear factor-4alpha expression in human colorectal carcinoma is a new prognostic factor against liver metastasis. *Pathology international* **57**: 82-90.

Paech K, Webb P, Kuiper GG, Nilsson S, Gustafsson J, Kushner PJ, Scanlan TS. 1997. Differential ligand activation of estrogen receptors ERalpha and ERbeta at AP1 sites. *Science* **277**: 1508-1510.

Pearce D, Matsui W, Miner JN, Yamamoto KR. 1998. Glucocorticoid receptor transcriptional activity determined by spacing of receptor and nonreceptor DNA sites. *The Journal of biological chemistry* **273**: 30081-30085.

Pendas-Franco N, Garcia JM, Pena C, Valle N, Palmer HG, Heinaniemi M, Carlberg C, Jimenez B, Bonilla F, Munoz A et al. 2008. DICKKOPF-4 is induced by TCF/beta-catenin and upregulated in human colon cancer, promotes tumour cell invasion and angiogenesis and is repressed by 1alpha,25-dihydroxyvitamin D3. *Oncogene* **27**: 4467-4477.

Ruse MD, Jr., Privalsky ML, Sladek FM. 2002. Competitive cofactor recruitment by orphan receptor hepatocyte nuclear factor 4alpha1: modulation by the F domain. *Molecular and cellular biology* **22**: 1626-1638.

Saha K. Supriya CAP, Krishna S. Ghanta, Julien Fitamant, Kenneth N. Ross, Mortada S. Najem, Sushma Gurumurthy, Esra A. Akbay, Daniela Sia, Helena Cornella, Oriana Miltiadous, Chad Walesky, Vikram Deshpande, Andrew X. Zhu, Aram F. Hezel, Katharine E. Yen, Kimberly S. Straley, Jeremy Travins, Janeta Popovici-Muller, Camilia Gliser, Cristina R. Ferrone, Udayan Apte, Josep M. Llovet, Kwok-Kin Wong, Sridhar Ramaswamy, and Nabeel Bardeesy. 2014. Mutation IDH inhibits HNF-4a to block hepatocyte differentiation and promote biliary cancer. *Nature*: 1-5.

Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic acids research* **18**: 6097-6100.

Siu A, Virtanen C, Jongstra J. 2011. PIM kinase isoform specific regulation of MIG6 expression and EGFR signaling in prostate cancer cells. *Oncotarget* **2**: 1134-1144.

Sladek FM. 2011. What are nuclear receptor ligands? *Molecular and cellular endocrinology* **334**: 3-13.

Sladek FM. 2012. The yin and yang of proliferation and differentiation: cyclin D1 inhibits differentiation factors ChREBP and HNF4alpha. *Cell cycle* **11**: 3156-3157.

Sladek FM, and Seidel, S.D. 2001. In Nuclear Receptors and Genetic Diseases. *TP Burris and ERB McCabe, eds (London: Academic Press)*: pp.309-361.

Sladek FM, Zhong WM, Lai E, Darnell JE, Jr. 1990. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes & development* **4**: 2353-2365.

Sollome JJ, Thavathiru E, Camenisch TD, Vaillancourt RR. 2014. HER2/HER3 regulates extracellular acidification and cell migration through MTK1 (MEKK4). *Cellular signalling* **26**: 70-82.

Takano K, Hasegawa G, Jiang S, Kurosaki I, Hatakeyama K, Iwanari H, Tanaka T, Hamakubo T, Kodama T, Naito M. 2009. Immunohistochemical staining for P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha may complement mucin phenotype of differentiated-type early gastric carcinoma. *Pathology international* **59**: 462-470.

Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C et al. 2006. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. *The Journal of pathology* **208**: 662-672.

Taraviras S, Monaghan AP, Schutz G, Kelsey G. 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mechanisms of development* **48**: 67-79.

Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* **14**: 178-192.

Tokuda E, Fujita N, Oh-hara T, Sato S, Kurata A, Katayama R, Itoh T, Takenawa T, Miyazono K, Tsuruo T. 2007. Casein kinase 2-interacting protein-1, a novel Akt pleckstrin homology domain-interacting protein, down-regulates PI3K/Akt signaling and suppresses tumor growth in vivo. *Cancer research* **67**: 9666-9676.

Torres-Padilla ME, Sladek FM, Weiss MC. 2002. Developmentally regulated N-terminal variants of the nuclear receptor hepatocyte nuclear factor 4alpha mediate multiple interactions through coactivator and corepressor-histone deacetylase complexes. *The Journal of biological chemistry* **277**: 44677-44687.

Torres-Padilla ME, Weiss MC. 2003. Effects of interactions of hepatocyte nuclear factor 4alpha isoforms with coactivators and corepressors are promoter-specific. *FEBS letters* **539**: 19-23.

Toyama T, Sasaki Y, Horimoto M, Iyoda K, Yakushijin T, Ohkawa K, Takehara T, Kasahara A, Araki T, Hori M et al. 2004. Ninjurin1 increases p21 expression and induces cellular senescence in human hepatoma cells. *Journal of hepatology* **41**: 637-643.

Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105-1111.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**: 511-515.

Van de Wetering M, Castrop J, Korinek V, Clevers H. 1996. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. *Molecular and cellular biology* **16**: 745-752.

Verzi MP, Shin H, He HH, Sulahian R, Meyer CA, Montgomery RK, Fleet JC, Brown M, Liu XS, Shivdasani RA. 2010. Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Developmental cell* **19**: 713-726.

Walesky C, Gunewardena S, Terwilliger EF, Edwards G, Borude P, Apte U. 2013. Hepatocyte-specific deletion of hepatocyte nuclear factor-4alpha in adult mice results in increased hepatocyte proliferation. *American journal of physiology Gastrointestinal and liver physiology* **304**: G26-37.

Wallerman O, Motallebipour M, Enroth S, Patra K, Bysani MS, Komorowski J, Wadelius C. 2009. Molecular interactions between HNF4a, FOXA2 and GABP identified at regulatory DNA elements through ChIP-sequencing. *Nucleic acids research* **37**: 7498-7508.

Weirauch U, Beckmann N, Thomas M, Grunweller A, Huber K, Bracher F, Hartmann RK, Aigner A. 2013. Functional role and therapeutic potential of the pim-1 kinase in colon carcinoma. *Neoplasia* **15**: 783-794.

Xia X, Ayala M, Thiede BR, Zhang SC. 2008. In vitro- and in vivo-induced transgene expression in human embryonic stem cells and derivatives. *Stem cells* **26**: 525-533.

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI. 1996. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* **384**: 458-460.

Yang J, Kong X, Martins-Santos ME, Aleman G, Chaco E, Liu GE, Wu SY, Samols D, Hakimi P, Chiang CM et al. 2009. Activation of SIRT1 by resveratrol represses transcription of the gene for the cytosolic form of phosphoenolpyruvate

carboxykinase (GTP) by deacetylating hepatic nuclear factor 4alpha. *The Journal of biological chemistry* **284**: 27042-27053.

Yang M, Li SN, Anjum KM, Gui LX, Zhu SS, Liu J, Chen JK, Liu QF, Ye GD, Wang WJ et al. 2013. A double-negative feedback loop between Wnt-beta-catenin signaling and HNF4alpha regulates epithelial-mesenchymal transition in hepatocellular carcinoma. *Journal of cell science* **126**: 5692-5703.

Yao L, Tak YG, Berman BP, Farnham PJ. 2014. Functional annotation of colon cancer risk SNPs. *Nature communications* **5**: 5114.

Yeung TM, Gandhi SC, Wilding JL, Muschel R, Bodmer WF. 2010. Cancer stem cells from colorectal cancer-derived cell lines. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 3722-3727.

Yuan X, Ta TC, Lin M, Evans JR, Dong Y, Bolotin E, Sherman MA, Forman BM, Sladek FM. 2009. Identification of an endogenous ligand bound to a native orphan nuclear receptor. *PloS one* **4**: e5609.

Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S et al. 2014. Proteogenomic characterization of human colon and rectal cancer. *Nature*.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**: R137.

Zhang Y, Zagnitko O, Rodionova I, Osterman A, Godzik A. 2011. The FGGY carbohydrate kinase family: insights into the evolution of functional specificities. *PLoS computational biology* **7**: e1002318.

Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, Green MR. 2010. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC bioinformatics* **11**: 237.

Supplemental Materials

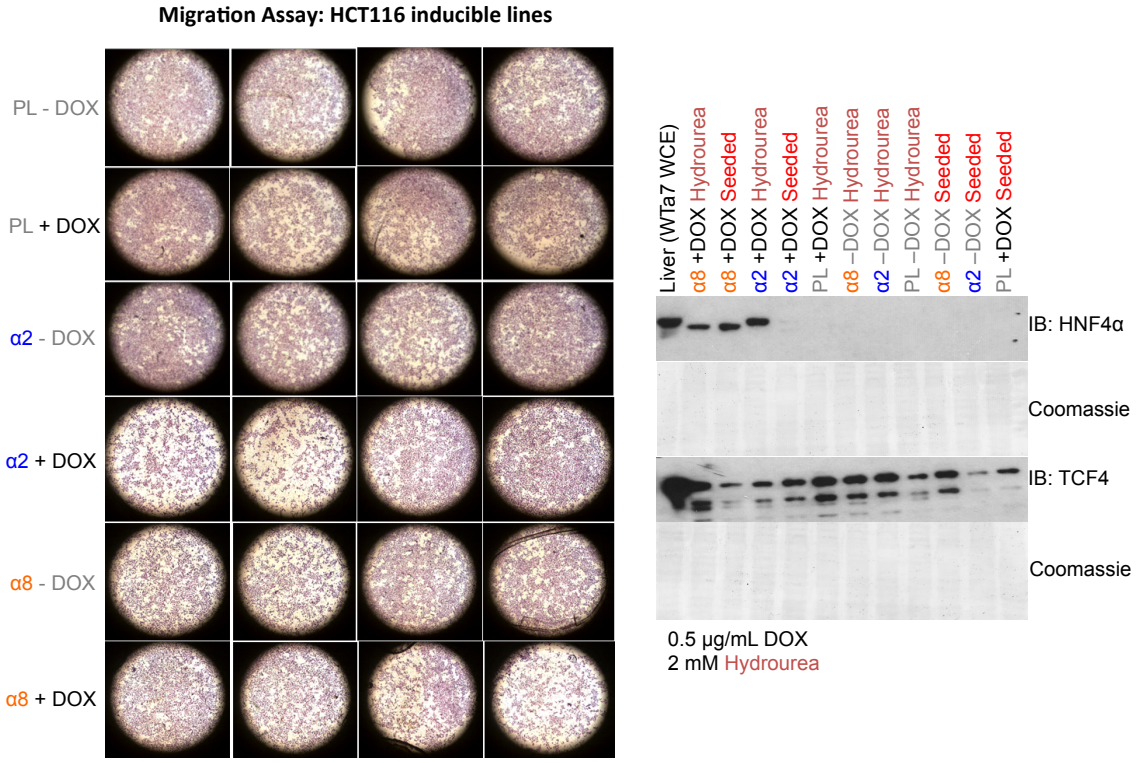
A repeat of the migration assay

I recently repeated a migration assay to try and reproduce the result shown in figure 3.2E. I found that the HNF4α2 line showed a decrease in migration in the presence of DOX while the PL (all 4 replicates) and HNF4α8 (3 out of 4 replicates) lines showed, although not quantitative, a relatively slight increased in migration (Fig.3.S1). The difference between the previous experiment (Fig. 3.2E) and this experiment is that I had seeded the cells in 1% Bovine Serum Albumin (BSA) instead of 0.1%. The IB analysis showed that HNF4α2 protein increased while TCF4 slightly decreased after 2 mM hydrourea treatment in the HNF4α2 line.

Experimental designs and sequencing read results from RNA-Seq and ChIP-Seq

Additional figures and tables of RNA-Seq and ChIP-Seq experimental designs and sequencing reads are included in this section.

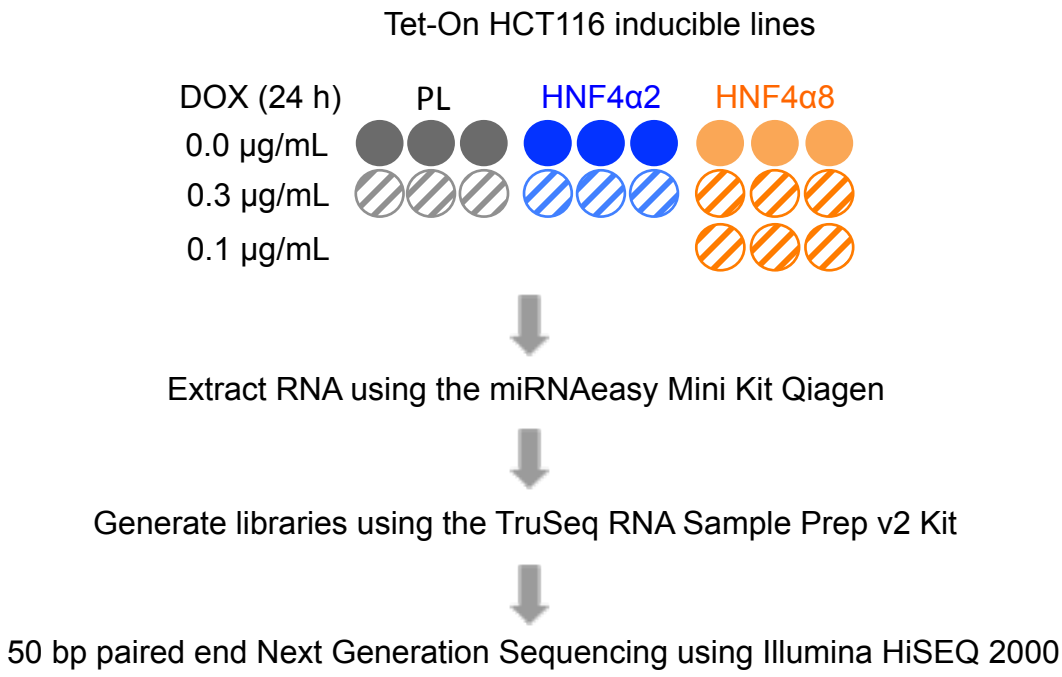
Fig. 3.S1


Migration Assay: HCT116 inducible lines
PL - DOX
PL + DOX
α2 - DOX
α2 + DOX
α8 - DOX
α8 + DOX
Liver (WTa7 WCE)
α8 +DOX Hydrourea
α8 +DOX Seeded
α2 +DOX Hydrourea
α2 +DOX Seeded
PL +DOX Hydrourea
α8 –DOX Hydrourea
α2 –DOX Hydrourea
PL –DOX Hydrourea
α8 –DOX Seeded
α2 –DOX Seeded
PL +DOX Seeded
IB: HNF4α
Coomassie
IB: TCF4
Coomassie
0.5 µg/mL DOX
2 mM Hydrourea

Supplemental Figure 3.1 Repeat of the migration assay.

Cells were seeded at 2.25 x 10^6 cells per 100-mm plate. The following day, cells were treated with 0.5 μg/mL DOX. Forty-eight hours later, cells were trypsinized, counted, and plated on the top of 24-transwell plates [that were incubated in serum free medium (SFM) for 15 min at 37^{o}C before used]. Cells were incubated in ~500 μL of SFM supplemented with 0.5 μg/mL, 2 mM hydrourea, and 1% BSA; the bottom wells were filled with ~750 μL 20% fetal bovine serum (FBS) medium. Forty-eight hours after treatment, cells were stained with crystal violet solution. Cells were fixed in 100% methanol, stained with crystal violet, and washed several times with 95% ethanol. Porous membranes from the transwells were cut and placed on coverslips, mounted, and viewed under a microscope. 4x magnification. Left, IB analysis of HNF4α and TCF4 from (20 μg/mL) WCE of the same batched of cells used in the migration assay, using the HNF4α anti-P1/P2 antibody and ant-TCF7L2 antibody. These cells were treated in the same way as those for the migration assay, but were plated on regular 24-well plates.

Fig. 3.S2



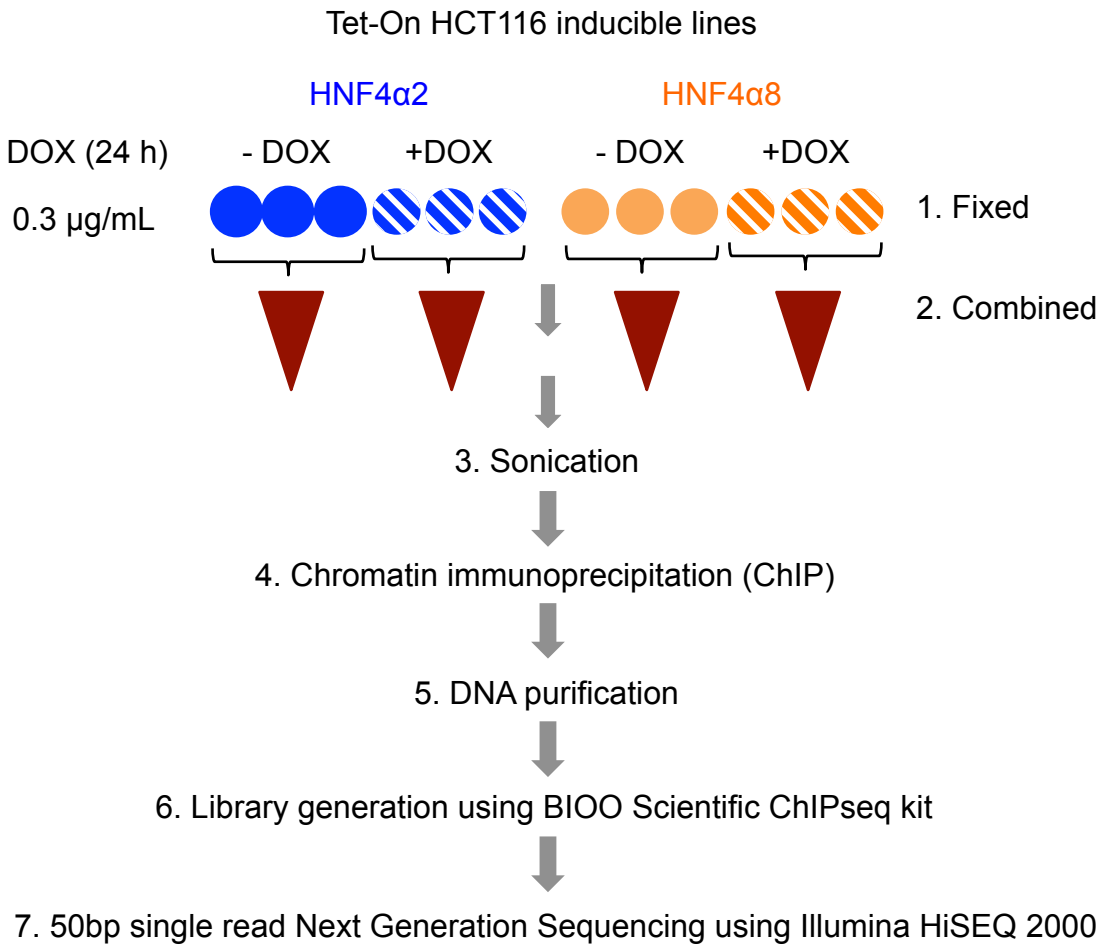
Supplemental Figure 3.2 Schematic of the RNA-Seq experimental design.

Cells were seeded at 5.5-6.5 x 10^5 cells per well of 6-well plates and (0.1 or 0.3 µg/mL) DOX treated 6 h later. RNA extracts were collected at 24 h after induction. RNA extracts were passed through an RNA column to collect total RNA. Poly A+ RNAs were selected and libraries were generated. The libraries were sent for 50 bp paired end Next Generation Sequencing (See M&M above for details). IB is shown in figure 3.3B.

Table 3.S1 RNA sequencing reads (HCT116 inducible lines)

Sample name	Index	Sequencing read
α8_0.1_DOX_rep1	CCGTCC	47,398,850
α8_0.3_DOX_rep1	GTCCGC	35,779,470
PL_0.0_DOX_rep1	CGATGT	30,352,082
PL_0.0_DOX_rep2	TGACCA	42,487,508
PL_0.3_DOX_rep1	ACAGTG	35,029,720
PL_0.3_DOX_rep2	GCCAAT	47,774,080
α2_0.0_DOX_rep1	CAGATC	36,073,304
α2_0.0_DOX_rep2	CTTGTA	38,892,684
α2_0.3_DOX_rep1	AGTCAA	45,032,626
α2_0.3_DOX_rep2	AGTTCC	38,421,126
α8_0.0_DOX_rep1	ATGTCA	42,251,180
α8_0.3_DOX_rep3	CCGTCC	39,579,642
PL_0.0_DOX_rep3	CGATGT	32,865,452
PL_0.3_DOX_rep3	TGACCA	68,409,410
α2_0.0_DOX_rep3	ACAGTG	49,553,690
α2_0.3_DOX_rep3	GCCAAT	30,898,168
α8_0.0_DOX_rep2	CAGATC	41,650,022
α8_0.0_DOX_rep3	CTTGTA	41,644,742
α8_0.1_DOX_rep2	AGTCAA	41,047,378
α8_0.1_DOX_rep3	AGTTCC	42,923,790
α8_0.3_DOX_rep2	ATGTCA	56,451,374

Fig. 3.S3



Supplemental Figure 3.3 Schematic of the ChIP-Seq experimental design.

Cells were seeded at 5-6 x 10^6 cells per 100-mm or 7-9 x 10^6 cells per 150-mm and (0.3 μg/mL) DOX treated 6 h later. The following ChIP steps were done: cells were fixed 24 h after treatment; three plates were combined into one 1.5 mL eppendorf tube; chromatins were sonicated, protein-chromatin interaction was immunopreicipated; and DNA fragments were purified. Libraries were generated and sent for 50 bp single end Next Generation Sequencing (See M&M above for details). IB analysis is shown in figure 3.5A.

Table 3.S2 ChIP sequencing reads (HCT116 inducible lines)

Flowcell 203 and 215	Sample name	Index	Sequencing reads
flowcell203_lane8_TAGCTT.fastq.gz	HNF4a2_DOX_IP_TCF4_rep1	TAGCTT	13,895,829
flowcell203_lane8_GGCTAC.fastq.gz	HNF4a2_DOX_IP_TCF4_rep2	GGCTAC	14,332,498
flowcell203_lane8_AGTCAA.fastq.gz	HNF4a8_DOX_IP_TCF4_rep1	AGTCAA	14,804,452
flowcell203_lane8_AGTTCC.fastq.gz	HNF4a8_DOX_IP_TCF4_rep2	AGTTCC	12,462,310
flowcell203_lane8_CGATGT.fastq.gz	HNF4a2_DOX_IP_a445_rep1	CGATGT	12,714,002
flowcell203_lane8_TGACCA.fastq.gz	HNF4a8_DOX_IP_a445_rep1	TGACCA	11,297,584
flowcell203_lane8_ACAGTG.fastq.gz	HNF4a8_input	ACAGTG	14,614,561
flowcell203_lane8_GCCAAT.fastq.gz	HNF4a2_input	GCCAAT	19,447,815
flowcell203_lane8_CTTGTA.fastq.gz	HNF4a8_DOX_IP_a445_rep2	CTTGTA	23,409,571
flowcell203_lane8_ATCACG.fastq.gz	HNF4a2_IP_TCF4_rep1	ATCACG	15,506,763
flowcell203_lane8_TTAGGC.fastq.gz	HNF4a2_IP_TCF4_rep2	TTAGGC	10,599,300
flowcell203_lane8_ACTTGA.fastq.gz	HNF4a8_IP_TCF4_rep1	ACTTGA	14,538,651
flowcell203_lane8_GATCAG.fastq.gz	HNF4a8_IP_TCF4_rep2	GATCAG	11,118,888
flowcell215_lane8_GGCTAC.fastq.gz	HNF4a2_DOX_IP_TCF4_rep2	GGCTAC	14,280,576
flowcell215_lane8_AGTCAA.fastq.gz	HNF4a8_DOX_IP_TCF4_rep1	AGTCAA	13,333,264
flowcell215_lane8_AGTTCC.fastq.gz	HNF4a8_DOX_IP_TCF4_rep2	AGTTCC	11,182,128
flowcell215_lane8_CGATGT.fastq.gz	HNF4a2_DOX_IP_a445_rep1	CGATGT	14,956,309
flowcell215_lane8_TGACCA.fastq.gz	HNF4a8_DOX_IP_a445_rep1	TGACCA	13,137,349
flowcell215_lane8_ACAGTG.fastq.gz	HNF4a2_DOX_IP_a445 rep2	ACAGTG	28,125,905
flowcell215_lane8_GCCAAT.fastq.gz	HNF4a2_DOX_IP_a445 rep2.1	GCCAAT	15,735,236
flowcell215_lane8_ATCACG.fastq.gz	HNF4a2_IP_TCF4_rep1	ATCACG	15,993,253
flowcell215_lane8_TTAGGC.fastq.gz	HNF4a2_IP_TCF4_rep2	TTAGGC	13,955,583
flowcell215_lane8_ACTTGA.fastq.gz	HNF4a8_IP_TCF4_rep1	ACTTGA	17,661,156
flowcell215_lane8_GATCAG.fastq.gz	HNF4a8_IP_TCF4_rep2	GATCAG	12,543,690
flowcell215_lane8_TAGCTT.fastq.gz	HNF4a2_DOX_IP_TCF4_rep1	TAGCTT	14,303,181

1. HNF4α2_DOX_IP_α445 (blue) condition has 24 million more reads than HNF4α8_DOX_IP_α445 (orange) condition; HNF4α2_DOX_IP_TCF4 has 5 million more reads than HNF4α8_DOX_IP_TCF4.
2. Samples were sequenced twice, two samples were replaced with other samples in the second sequencing round.
    a. First sequencing round: Samples multiplex in one lane.
        HNF4α2_input (1); HNF4α2_IP_TCF4 (2); HNF4α2_IP_DOX_TCF4 (2);
        HNF4α2_IP_DOX_a445 (1); HNF4α8_input (1); HNF4α8_IP_TCF4 (2);
        HNF4α8_IP_DOX_TCF4 (2); HNF4α8_IP_DOX_a445 (2)
    b. Second sequencing round: Two inputs were replaced with two samples of HNF4α2_IP_DOX_a445. Original samples plus two new samples were sequenced in one lane.
        HNF4α2_IP_TCF4 (2); HNF4α2_IP_DOX_TCF4 (2); HNF4α2_IP_DOX_a445 (3);
        HNF4α8_IP_TCF4 (2); HNF4α8_IP_DOX_TCF4 (2); HNF4α8_IP_DOX_a445 (2)
3. Analysis from both sequencing datasets were combined.

Table 3.S3 Genes dysregulated in HCT116 inducible α2 and α8 lines (+ DOX verses -DOX)

HNF4α2 0.3 / HNF4α2 0.0	
Gene	**FC**
HNF4A	80.0
C11orf86	21.2
CREB3L3	12.8
APOBEC1	8.4
EPS8L3	7.2
IL32	6.3
HPD	6.3
AQP3	6.1
RBP2	5.8
VTN	5.8
MYO15B	5.0
SBSN	4.9
SEPP1	4.6
ACSL5	4.0
KLK6	4.0
CORO2B	3.6
DMBX1	3.4
NMB	3.4
DHRS3	3.3
SUSD2	3.1
SEMA3B	3.0
TRIM15	2.9
SERPINE1	2.8
ABCG2	2.7
PDE2A	2.7
ALS2CL	2.6
RBKS	2.6
41521	2.5
MLXIPL	2.5
MGST2	2.5
KCNQ4	2.5
VPS9D1	2.5
POLD4	2.4
ASS1	2.4
PLEKHO1	2.4
ISG20	2.3
SLC9A3R1	2.3
CIDEC,EMC3	2.3
OSGIN1	2.3
SH3D21	2.3
METTL7B	2.3
FAM214B	2.3
NAT1,NAT2	2.2
USH1C	2.2
CLCF1	2.2
ANGPTL4	2.2
CORO2A	2.2

HNF4α2 0.3 / HNF4α2 0.0	
Gene	**FC**
C10orf114	2.1
OAF	2.1
BLVRB	2.1
KRT7	2.1
KLC4	2.1
NAGS	2.1
LAMB2	2.1
CLIC3	2.1
CCDC120	2.1
LRP10	2.1
PPAP2B	2.1
FAM109A	2.0
FBXO2	2.0
ABCC3	2.0
PC	2.0
CLDN7	2.0
CRB3	2.0
DHRS2	-2.3
RPS17	-38.2
RPS17	-42.9

HNF4α8 0.1 / HNF4α8 0.0	
Gene	**FC**
HNF4A	57.3
C11orf86	13.3
CREB3L3	8.5
ACSL5	5.6
EPS8L3	5.5
IL32	5.0
MYO15B	4.8
TRIM15	3.4
DHRS3	3.2
S100A14	3.0
SH3D21	3.0
RHOU	2.7
SEPP1	2.7
C8G	2.5
TRIM10	2.5
SEMA3B	2.4
FRMD6	2.3
VIPR1	2.3
FIBCD1	2.3
SUSD2	2.2
CD58	2.1
KLF4	2.1
CLCF1	2.0
ATRN	2.0
SLC16A6	-2.0
HAS3	-2.0
CYP24A1	-2.3

HNF4α8 0.3 / HNF4α8 0.0	
Gene	**FC**
HNF4A	79.7
C11orf86	20.1
CREB3L3	13.7
ACSL5	8.1
IL32	7.2
EPS8L3	6.3
PIPOX	6.1
MYO15B	5.6
TUBAL3	5.5
HPD	4.8
OAS1	4.7
PDE2A	4.6
DHRS3	4.4
TRIM15	4.1
SH3D21	3.7
41521	3.4
RHOU	3.4
S100A14	3.3
C8G	3.2

HNF4α8 0.3 / HNF4α8 0.0	
Gene	**FC**
FRMD6	3.1
ATRN	2.9
TSPAN18	2.9
SEPP1	2.9
ABCG2	2.9
SEMA3B	2.9
CORO2B	2.7
GP1BB,SEPT5	2.7
VIPR1	2.6
TRIM10	2.6
FIBCD1	2.6
KLK6	2.5
SUSD2	2.5
CLCF1	2.5
PIM1	2.5
CIDEC,EMC3	2.4
RBKS	2.4
PC	2.4
MLXIPL	2.3
MGST2	2.3
ABCC3	2.3
TRIM34,TRIM6, TRIM6-TRIM34	2.3
KLF4	2.2
AGPAT9	2.2
NAGS	2.2
PGPEP1	2.2
ZDHHC9	2.2
FCHO2	2.1
S100A16	2.1
SLC35D2	2.1
MPP1	2.1
JUN	2.1
SERINC2	2.1
DKK4	2.1
METTL7B	2.1
IRAK2	2.1
CD58	2.0
NPAS1	2.0
FAM214B	2.0
LRP10	2.0
ART5	2.0
CPPED1	2.0
ADM	2.0
SDC1	2.0
EPS8L2	2.0
C2CD2L	2.0
PRADC1	2.0
BLVRB	2.0
MPZ	2.0
SP6	2.0
ERBB3	2.0

HNF4α8 0.3 / HNF4α8 0.0	
Gene	**FC**
PDE4B	-2.0
SOX4	-2.1
MYBPH	-2.1
ERRFI1	-2.1
HTR7	-2.3
BDNF	-2.3
DHRS2	-2.4
HAS3	-2.5
SLC16A6	-2.5
NRG1	-2.6
TNFSF18	-2.7
KRTAP3-1	-2.9
ANXA10	-2.9
CYP24A1	-2.9
RPS17	-3.6

Note:
Fold change (FC) non-log
HNF4α2 (α2) and HNF4α8 (α8) lines
0.0 μg/mL – no doxycycline (DOX)
0.1 μg/mL – DOX
0.3 μg/mL – DOX
Negative FC – genes downregulated
Positive FC – genes upregulated

Appendix to Chapter 3

In this Appendix I will discuss some of the preliminary results on the c-Myc, HNF4α, and p21 project that led us to shift our focus and examine the HNF4α and TCF4 interaction instead. I will also go over the co-IP experiments on HNF4α, β-catenin, TCF4, and dnTCF1. In addition, I will explore the dnTCF1 transactivation in the HEK293T cell line and look closely at the HstrgTwk and HwkTstrg elements in a gel shift. Finally, I will present the IB analyses on some of the samples used in the gel shifts that are shown in Chapter 3.

Materials and Methods

Plasmid constructs

pcDNA3.1.HNF4α2 and pcDNA3.1.HNF4α8 plasmids are described in the Material and Methods (M&M) of Chapter 3; the pcDNA3.1.HNF4α2 (Y63E) mutant plasmid was made by Dr. Karthikeyani Chellappa in the Sladek's lab. The luciferase reporter constructs of human full-length p21 promoter (a gift from Dr. Xiao-Fan Wang at the Duke University of Medical Center, Durhman, NC) and ApoB.-85-47.E4.Luc were previously described (Datto MB et al. 1995; Maeda Y et al. 2002; Hwang-Verslues and Sladek 2008). The full-length mouse pCBS.Flag.c-Myc (N-terminus of c-Myc fused to the Flag epitope of pCBS) was a gift from Dr. Ernest Martinez (University of California, Riverside) (Faiola et al. 2005). pcDNA4/TO.Flag.TCF4E2 (Hoverter et al. 2012) was provided by Dr. Marian Waterman at the University of California, Irvine. The TopFlash reporter construct (M50 Super 8x TopFlash; Addgene) was a gift from Dr. Nicole zur Nieden at the University of California, Riverside.

Cell Culture

Cell lines (HEK293, HEK293T, HCT116, and HCT116 inducible cell lines [HNF4α2 (α2), HNF4α8 (α8), and Parental line (PL)] contain the rtTA) were maintained as described in the M&M of Chapter 3.

Cell count assay

Both transient and stable cell lines were used to evaluate cell number. In transient transfection, HEK293T or HCT116 cells were seeded at 4.5 x 10^6 or 6.5 x 10^5 cells per well, respectively, in a 6-well plate. The following day, cells were transfected with 1-2 μg of pcDNA3.1.HNF4α2, pcDNA3.1.HNF4α8, or pcDNA3.1 plasmid using HiFect (LONZA). Six to 24 h after transfection, cells were trypinized and transferred to a 12-well plate at a density of 1 x 10^5 cells per well. For the Tet-On inducible stable lines (α2, α8, PL), cells were seeded at 1.0 x 10^4 cells per well of 12-well and, 24 h later, treated with 0.5 μg/mL of DOX. Media [plus 0.5 μg/mL DOX] were changed every 48 h and the cells were counted at the indicated time points [in hours or days], using either a hemocytometer or Coulter Counter (Beckman).

MTT assay

Tet-On inducible HCT116 clones (PL, α2, or α8) were seeded at 1.5-2.25 x 10^6 cells in a 100-mm plate and, 24 h later, treated with or without 0.5 μg/mL of Doxycycline (DOX, Clontech). Then 48 h after induction, cells were trypsinized, counted, and plated on a 96-well plate for the MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide) assay, following the manufacturer's protocol (Vybrant).

Co-Immunoprecipitation (co-IP)

HNF4α and c-Myc

CaCo2 or HEK293 cells were seeded at 6-8 or 9 x 10^6 cells in 150-mm plates. HEK293 cells were transfected the next day with 12 µg each of Flag.c-Myc and pcDNA3.1.HNF4α2/α8 via $CaPO_4$ precipitation as described previously (Jiang et al. 1995). Nuclear extracts (NE) of CaCo2 and whole cell extracts (WCE) of HEK293 were performed as previously described (Hwang-Verslues and Sladek 2008). For immunoprecipitation (IP) of c-Myc or HNF4α, 200 µg to 1 mg of crude WCE (HEK293 cells) or NE (CaCo2 cells) diluted in a 1:1 Ab/Ag buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.2 mM EGTA, 10% glycerol, and 1 mM dithiothreitol (DTT)) were incubated (at 4^{o}C for 2 h) with 10 µg of anti-Flag M2 (Sigma) or 5 µg of an affinity purified anti-HNF4α antibody (α445) that recognizes the C-terminus of P1- and P2-HNF4α isoforms, respectively. Normal rabbit IgG (Santa Cruz) was used as a negative control. Twenty-five microliters of Protein G beads (1:1 slurry in PBS) (Pierce) were added and incubated at 4^{o}C overnight. The following day, IP samples were washed 3 times (inverted 6-10 times) with 0.05% NP-40 lysis buffer (50 mM Tris-HCl pH8.0 and 120 mM NaCl) and 1-2 times with 1 x PBS and eluted with 5 x Sodium Dodecyl Sulfate (SDS) buffer.

HNF4α, TCF4, dnTCF1, and β-catenin

COS-7 cells were seeded at 3.5-4.0 x 10^6 cells per 150-mm plate of and, 24 h later, transfected with 12-12.5 µg each of Flag.TCF4, Flag.TCF3, Flag.dnTCF1, pcDNA3.1.HNF4α2, or HNF4α8 via $CaPO_4$ precipitation. The following day, NE were prepared and 300-500 µg total protein diluted in a 1:1 Ab/Ag buffer was incubated with 5 µg of anti-Flag M2 or α445, at 4^{o}C for 2 h. Thirty microliters of Protein G beads (1:1 slurry) were added and incubated at 4^{o}C overnight. The following day, IP samples were

washed 3 times (inverted 6-10 times) with 0.05% NP-40 lysis buffer and 1-2 times with 1 x PBS and eluted with 5 x SDS buffer.

Luciferase assay

p21 and ApoB promoters

COS-7 cells were seeded at 9.4 x 10^4 cells per well of a 12-well plate. The following day, cells were transfected with 0.5 µg of expression plasmid (pcDNA3.1.HNF4α2, pcDNA3.1.HNF4α8, or pcDNA3.1), 0.5 µg of reporter construct (p21FL.Luc (-2.4 kb) or ApoB.-85-47.E4.Luc), and 0.1 µg CMV.β-gal using HiFect (LONZA) reagent and harvested 24 h later with lysis buffer (25 mM Glycylglycine pH 7.8, 15 mM $MgSO_4$, 4 mM EGTA, 1% (v/v) Triton X-100). Luciferase and β-galactosidase activities were measured as previously described (Hwang-Verslues and Sladek 2008). The reporter construct ApoB.-85-47.E4.Luc was used as a control. Error bars were reported as mean±SD. Immunoblot (IB) analysis was performed to verify protein expression.

HNF4α and TCF reporter constructs

Luciferase and β-galactosidase assays were performed as previously described in the M&M of Chapter 3. The reporter construct that contains the HNF4 strong and TCF4 strong site (HstrgTstrg), the TopFlash and ApoB.-85-47.E4.Luc reporter constructs, and the G4.23[*luc2/minP*] minimal reporter plasmid were used. Additional amount (60, 120, and 400 ng) of HNF4α and dnTCF1 expression vectors, other expression vectors (HNF4α2 (Y63E) mutant, pcDNA4/TO.Flag.TCF4, and TCF3) were tested. Bar graphs are mean±SD of triplicate samples.

Immunoblot (IB) assay

IB analysis was performed as described in M&M of Chapter 3 using the anti-TCF7L1 antibody (Santa Cruz, sc166411) (1:1000), anti-HNF4α P1/P2 antibody (R&D) that recognizes the C-terminal end of all HNF4α isoforms, and anti-Flag M2 (Sigma). Ectopic HNF4α and flag tagged TCFs expressions in NE of COS-7 were quantified using known amounts of a standard recombinant HNF4α protein that contains the Ligand Binding Domain plus the F domain (LBD/F) (Ruse et al. 2002) and the carboxy-terminal Flag fusion protein (Flag-BAP from Sigma).

Fluorescent Gel Shift

The gel shift assay was performed as described in M&M of Chapter 3, except fluorescently labeled probes containing the HNF4α weak and TCF4 strong (HwkTstrg) and HNF4α strong and TCF4 weak (HstrgTwk) elements were used.

Results

HNF4α isoforms and c-Myc in cellular proliferation

Proliferation and differentiation are two fundamental events in developmental biology. The mechanism(s) that transition cells from proliferation to differentiation and reverse are yet clear. One factor that is involved in regulating these two events is the cyclin-dependent kinase inhibitor gene p21/WAFT1 (*CDKN1A*). It is known to block cell cycle progression especially in response to cellular stress. Hwang-Verslues and Sladek (Hwang-Verslues and Sladek, 2008) identified a mechanism that involves the interaction between HNF4α1 (the P1-HNF4α isoform) and the proto-oncogene c-Myc on promoter-bound Sp1 for transcriptional control of the p21/WAF1 (*CDKN1A*) promoter. The authors

showed that, in a low c-Myc milieu when cells are not proliferating, HNF4α1 is recruited to Sp1-bound sites both proximal and distal to the transcriptional start site (TSS) of the *CDNK1A* promoter and thereby positively regulates the activity of the *CDNK1A* gene. However, when cells are proliferating, the concomitant high level of c-Myc protein inhibits this activity by either competing with HNF4α1 for Sp1-bound sites, sequestering HNF4α1 off of the promoter, or interacting with HNF4α1 on the promoter. While this work showed how a differentiated factor like HNF4α can inhibit cell proliferation, only one HNF4α isoform (P1-driven HNF4α1) was examined; P2-driven HNF4α (HNF4α7/8) was not examined. Therefore, I performed experiments similar to those in Hwang-Verslues and Sladek 2008 to determine whether the P2-HNF4α isoform exhibits a similar interaction with c-Myc and activation of the p21 promoter. Since P2-HNF4α has been shown to be upregulated in certain liver cancers (Tanaka et al. 2006), our hypothesis was that it would not upregulate p21 expression.

The difference between P1- and P2-HNF4α is 16-38 aa in the N-terminal region; the remainder of the protein – DBD, LBD, F – are identical. In a co-IP assay I observed an interaction between HNF4α8 (the P2-HNF4α isoform) and c-Myc in HEK293 cells ectopically expressing Flag.c-Myc and human HNF4α8 (Fig. 3.1.1A, left). Similarly, I also observed interaction between endogenous HNF4α8 and c-Myc in human epithelial colorectal adenocarcinoma cells (CaCo2) (Fig. 3.1.1A, right). These results are similar to those observed by Hwang-Verslues and Sladek with P1-HNF4α.

To determine whether HNF4α8 can activate the p21 promoter, we transfected in the full-length p21 promoter (-2.4 kb) and human HNF4α8 [or HNF4α2] in COS-7 cells and found that HNF4α8 (P2-HNF4α) increases the p21 promoter activity better than HNF4α2 (P1-HNF4α) (Fig. 3.1.1B, left); the ApoB promoter (-85-47E4) was used as a

positive control (Fig. 3.1.1B, right) and showed that on that promoter, HNF4α8 activates transcription less well than HNF4α2. The IB analysis verified roughly equal level of HNF4α2 and HNF4α8 expression (Fig. 3.1.1B, below), although HNF4α8 was expressed at a somewhat higher level, especially in the ApoB luc experiment.

To compare the effect of the HNF4α isoforms on cell proliferation, I performed cell count assay in HEK293T and HCT116 lines ectopically expressing HNF4α2 or HNF4α8. Cells transiently transfected with either human HNF4α2 or HNF4α8 showed a decrease in cell number in both lines, with HNF4α8 slowing proliferation (Fig. 3.1.1C) somewhat more than HNF4α2, although the difference was not statistically significant. This result was reproducible using the Tet-On inducible HCT116 lines (clone 11 and clone 17) at days 4 and 6 in the cell count and MTT assays (Fig. 3.1.2A,B), although cell numbers did not change significantly in clone 17 lines.

Co-IPs of HNF4α and TCF4/dnTCF1/β-catenin

Several papers used co-IPs to show an interaction between HNF4α and LEF1/TCF4/β-catenin in colon and liver (Colletti et al. 2009; Cattin et al. 2009; Yang et al. 2013; Gougelet et al. 2014). Colletti et al. and Cattin et al. did not see interaction between HNF4α and β-catenin. Similarly, we did not observe interaction between HNF4α and β-catenin in CaCo2, HepG2, and COS7 cells ectopically expressing human HNF4α2 in a co-IP assay (Fig. 3.1.3A,B), although Yang et al and Gougelet et al. were able to see the interaction in their co-IP (Yang et al. 2013; Gougelet et al. 2014). Cattin et al, Yang et al, and Gougelet et al showed interaction between HNF4α and TCF4 (Cattin et al. 2009; Yang et al. 2013; Gougelet et al. 2014) in intestinal epithelial cells, hepatocellular carcinoma, and hepatocytes. We did observed in one experiment an

interaction between HNF4α and TCF4 in COS7 cells overexpressing HNF4α2 and Flag.TCF4 (Fig. 3.1.3C). However, we could not reproduce the interaction in the HCT116 inducible lines expressing HNF4α2 or HNF4α8 (Fig. 3.1.3D). Yang et al mapped the interaction domains of HNF4α to its N-terminal region [includes the activation function (AF-1), DNA binding domain (DBD), and the hinge (H)] and TCF4 to its N-terminal β-catenin binding domain (Yang et al. 2013), and hence did not observed an interaction between HNF4α and the dominant negative TCF4 (dnTCF4) which lacks the β-catenin binding domain. This result is similar to our finding in which we did not observe an interaction with the dnTCF1 (lacks β-catenin binding domain) (Fig. 3.1.3B). However, under our co-IP conditions we did not observe an interaction with full length TCF4 in HCT116 cells either (Fig. 3.1.3D).

Flag.dnTCF1 acting as an activator

We were surprised to find that Flag.dnTCF1 activated transcription on the reporter construct containing the HstrgTstrg in HEK293T cell line even though the dnTCF1 does not contain the β-catenin binding domain (Fig.s 3.1.4A and 3.1.5A,B; and Fig. 3.6D, Chapter 3); β-catenin is known to function as a co-activator. To confirm dnTCF1 activity on the HstrgTstrg element, we updated these experiments using the construct containing the HstrgTstrg as well as the TopFlash reporter construct, luciferase reporter that contains several Wnt responsive (WRE) elements next to a minimal promoter, and compared the activity of dnTCF1 alongside that of full length TCF4 and TCF3. Reproducibly, the results show that dnTCF1 activates transcription (Fig. 3.1.5A,B). Although the dnTCF1 activity also was observed on the pG4.23(*luc2/minP*) vector lacking the HNF4α and TCF4 binding element (Fig. 3.1.4B), the RLU value is

much lower (10^4 compared to 10^5). One group examined closely the functional differences between the short and long TCF1 isoforms and found that the short N termini isoforms can in fact activate transcription. (van de Wetering et al. 1996), They identified eight human TCF1 isoforms ranging from 25 to 50 kDa in human thymocytes and Jurkat T cells and found that the four isoforms with short N termini and the four with the long N termini transactivated transcription to similar extents through the T-cell receptor (TCR)-α enhancer (van de Wetering et al. 1996). However, when the ApoB.-85-47.E4.Luc reporter construct was tested, activity was observed only for HNF4α2 but not dnTCF1 or HNF4α2 mutant (Y63E) – mutations in the DBD domain that disrupt binding to the DNA (Fig. 3.1.4C). This would suggest that transactivity by dnTCF1 was observed specifically on the HstrgTstrg element.

In addition to dnTCF1 activating transcription, the other surprising result in our luciferase assay was that neither TCF3 nor TCF4 activated transcription well, despite updated attempts and verification of protein expression (Fig. 3.1.5A,B). This could be due to the long E-tail of TCF4 (sequence in turquoise; Fig. 3.1.6) and TCF3. There are two isoforms of the TCF4 protein, the long (E) and short (B) C-termini tails. The E-tail contains a Ctbp motif that is known to repress transcription; the B-tail does not (Tang et al. 2008; Weise et al. 2009). TCF3 only has the long E-tail (Arce et al. 2006). Uniquely, TCF3 can repress transcription, independent of the extended C-terminal E-tail or Ctbp binding motifs (Merill et al. 2001; Pereira et al. 2006). Also of note is that closer inspection of the amino acid sequence of Flag.dnTCF1 and the full length TCF1, the sequences in orange are identical between the two isoforms (Fig. 3.1.6). It is possible that this sequence is what drives transcription, but we would need to test this hypothesis to confirm it.

HwkTstrg and HstrgTwk in gel shift assay

The three sequences, discussed in Chapter 3, that had a weak or strong affinity for HNF4α and TCF (HstrgTstrg, HstrgTstrg, HstrgTwk) as determined by the PBM, were examined in a gel shift and confirmed the relative affinity for HNF4α and TCF that was determined by the PBM (Fig. 3.6A, Chapter 3). In the gel shift, both HNF4α and dnTCF1 bound the HstrgTstrg element, although more dnTCF1 protein was required to produce a shift band of the same binding intensity as HNF4α (Fig. 3.6B,C, Chapter 3) despite the fact that dnTCF1 gave a good binding score in the PBM. As predicted by the PBM data, dnTCF1 gave a low to almost no intensity shift band on the HstrgTwk (Fig. 3.1.7, lane 10, 13, & 16) or HNF4α gave a low intensity band on the HwkTstrg, even with increasing amount of protein (Fig. 3.1.7, lane 3 & 6). On the HstrgTwk, we observed better binding by HNF4α at 30 ng of protein (lane 12), and interestingly, when 130 ng of dnTCF1 protein was added to the reaction, the binding intensity of HNF4α was reduced by half (lane 11 & 12). Likewise, on the HwkTstrg sequence, dnTCF1 binding gave a better signal than HNF4α; and this time, when HNF4α was added, dnTCF1 binding intensity was reduced slightly (lane 2 & 4). This reduced intensity suggests that the probe is occupied by the two factors independently.

Proteins used in the gel shifts were analyzed on IBs prior to performing the shift reactions (Fig. 3.1.8). Total nuclear proteins were extracted from COS-7 cells that had ectopically expressed either the human HNF4α2 or Flag tagged TCF4, TCF3, or dnTCF1. The Flag protein and LBD/F peptide of HNF4α were used as standards to quantify the amount of TCFs and HNF4α proteins, respectively. Concentrations of the proteins used in the gel shift ranged from as low as 3 ng/μL to as high as 130 ng/μL.

Discussion

HNF4α and c-Myc

Taken together, the results indicate that, contrary to our original hypothesis, P2-driven HNF4α8 activates p21 gene expression and inhibits cell proliferation just as well as P1-driven HNF4α1/2. While we do not know whether HNF4α8 does this through the same mechanism as HNF4α2 (i.e., via interaction with c-Myc and Sp1 at least on *CDNK1A*), it is certainly possible as we also found that HNF4α8 interacts with c-Myc in a co-IP, although we did not compare HNF4α8 interaction a long side HNF4α2 interaction.

HNF4α and TCF

While others were able to show interaction between HNF4α and TCF4 or LEF1 (Colletti et al. 2009; Cattin et al. 2009; Yang et al. 2013; Gougelet et al. 2014), we were not able to recapitulate the co-IP experiment between HNF4α and TCF4 in the inducible HCT116 lines. We were, however, able to show, in one experiment, an interaction between the two factors in COS-7 cells, ectopically expressing HNF4α2 and Flag.TCF4. And even though others and we did not see interaction between HNF4α and β-catenin, Yang et al. and Gougelet et al. showed HNF4α interacts with TCF4 and β-catenin (Yang et al. 2013; Gougelet et al. 2014) in hepatocytes.

The dogma that the dominant negative TCF1 (dnTCF1) lacking the β-catenin domain, which could not interact with β-catenin, represses activity on the Wnt responsive element (WRE: 5'-CTTTGWWW-3') is changing. One group showed that various TCF1/LEF1 lengths activate transcription with similar degree on the TCRα enhancer (van de Wetering et al. 1996). A review mentioned a recent report on TCF1 and LEF1 recruitment and interaction with a new partner, ATFs – part of the AP-1 complex, to

activate β-catenin-independent activity (Sprowl and Waterman 2013; Grumolato et al. 2013). Grumolato et al showed that the full length as well as the dominant negative TCF1 and LEF1 but not TCF4 and TCF3 recruit ATFs to the TopFlash multiple WRE elements to activate the TopFlash reporter gene. This is consistent with our findings for the TopFlash and the HstrgTstrg reporter constructs. However, dnTCF1 was not able to activate the ApoB.-85-47.E4.Luc reporter construct, suggesting that the binding sequence is HNF4α but not HNF4α and TCF/LEF specific binding site. This would also suggest that our HstrgTstrg element, like the TopFlash (Gougelet et al. 2014), is both HNF4α and TCF/LEF specific, as determined by the PBM.

Although, in vivo ChIP-Seq assay, we found peaks of recruitment of TCF4 by HNF4α and peaks of co-localization with TCF4 by HNF4α, as discussed in Chapter 3, we were not able to see corporative binding between HNF4α and TCF in vitro. Instead, we observed independent binding by the two factors on the three elements with varying degree of affinity for HNF4α and TCF; this could be because such an interaction in a gel shift is not stable.

Fig. 3.1.1

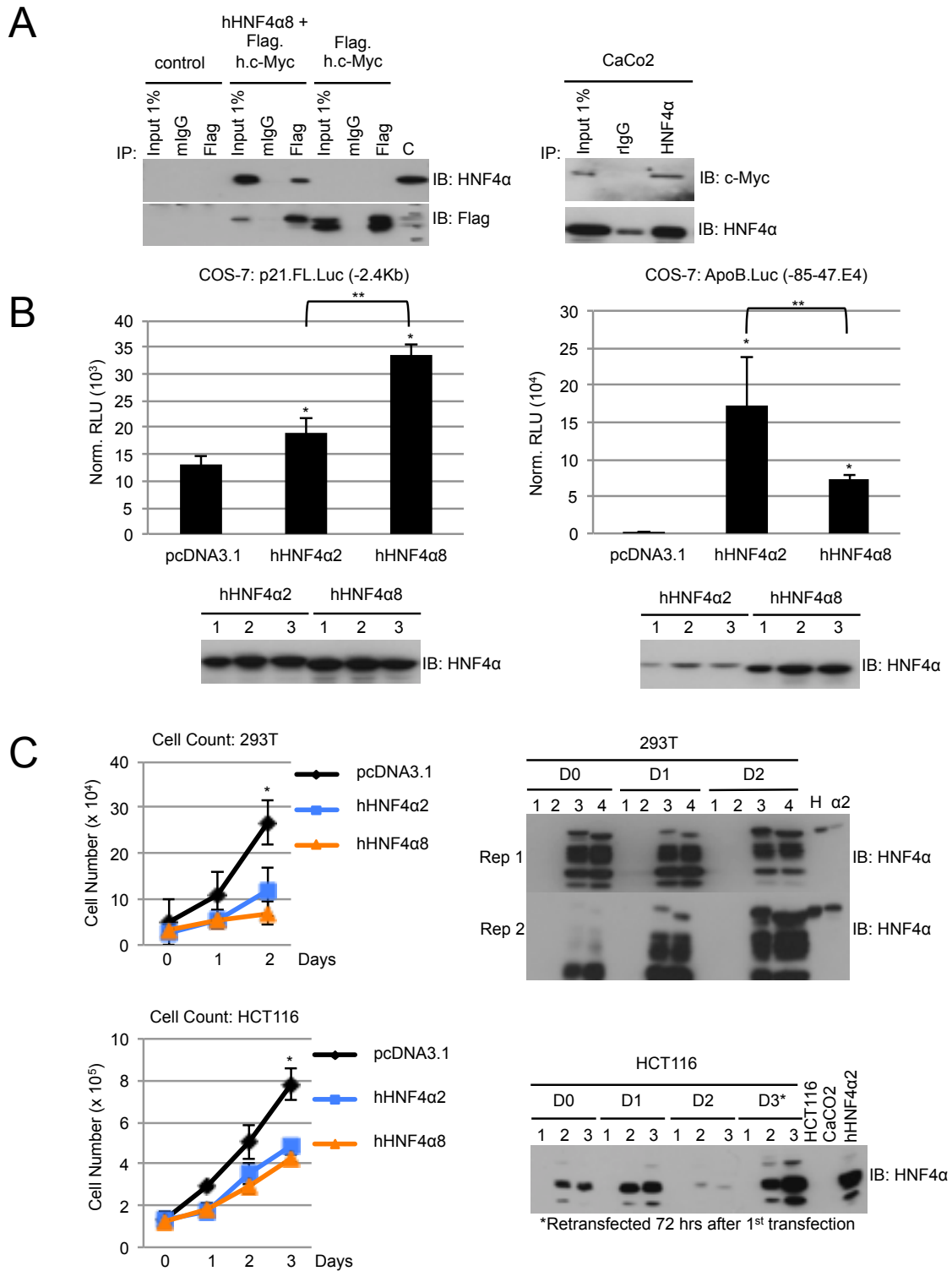

A
hHNF4α8 + Flag. h.c-Myc
Flag. h.c-Myc
control
Input 1%
mIgG
Flag
C
IP:
IB: HNF4α
IB: Flag
CaCo2
rIgG
HNF4α
IB: c-Myc
B
COS-7: p21.FL.Luc (-2.4Kb)
Norm. RLU (10³)
pcDNA3.1
hHNF4α2
hHNF4α8
COS-7: ApoB.Luc (-85-47.E4)
Norm. RLU (10⁴)
C
Cell Count: 293T
Cell Number (x 10⁴)
Days
293T
D0
D1
D2
H α2
Rep 1
Rep 2
Cell Count: HCT116
Cell Number (x 10⁵)
HCT116
D3*
CaCO2
*Retransfected 72 hrs after 1st transfection

Figure 3.1.1 Interplay between HNF4α and c-Myc in cellular proliferation.

(A) Co-IPs of (200 µg to 1 mg) NE of exogenous (left) and endogenous (right) c-Myc and HNF4α8 (P2-HNF4α isoform) in HEK293 (left) and CaCo2 cells (right). IB of the co-IP samples using anti-HNF4α (P1/P2) and anti-Flag antibodies as indicated. (B) Transfection of COS-7 cells with the full-length (-2.4 kb) p21 or ApoB (-85-47.E4) promoter construct (0.5 µg) and HNF4α2 or HNF4α8 expression vector (0.2 µg left, 0.5 µg right). Bar graphs represent mean$\pm$SD of triplicate samples from one independent experiment. IB analysis of HNF4α isoform expression in COS-7 cells. * HNF4α2/8 vector compared to empty vector, ** HNF4α2 vector compared to HNF4α8 vector, *P-value* $\leq$ 0.05. (C) Cell proliferation assay of transient transfected HEK293T and HCT116 cells as indicated. After the second day, HCT116 cells were retransfected with the same corresponding expression vectors and counted on the third day (bottom). Bar graphs represent mean$\pm$SD of six samples from two independent experiments (HEK293T) or triplicate samples from one experiment (HCT116). * Empty vector vs. HNF4α2 or HNF4α8 vector, *P-value* $\leq$ 0.05. IB analysis of HNF4α in WCE (20 µg) of HEK293T (top left) and HCT116 cells (below left). (C, Top left) WCE from two independent experiments (Rep 1, Rep 2) are shown (lane 1, no vector; lane 2, empty vector; lane 3, hHNF4α2 vector; lane 4, hHNF4α8 vector). Controls: H, HepG2; HNF4α2, from transfected COS-7 cells with the HNF4α2 construct; CaCO2; and HCT116; cells were counted (Day 0 (D0)) 7-10, (D1) 24-30, (D2) 48-54, and/or (D3) 72-76 hours after transfection.

Fig. 3.1.2

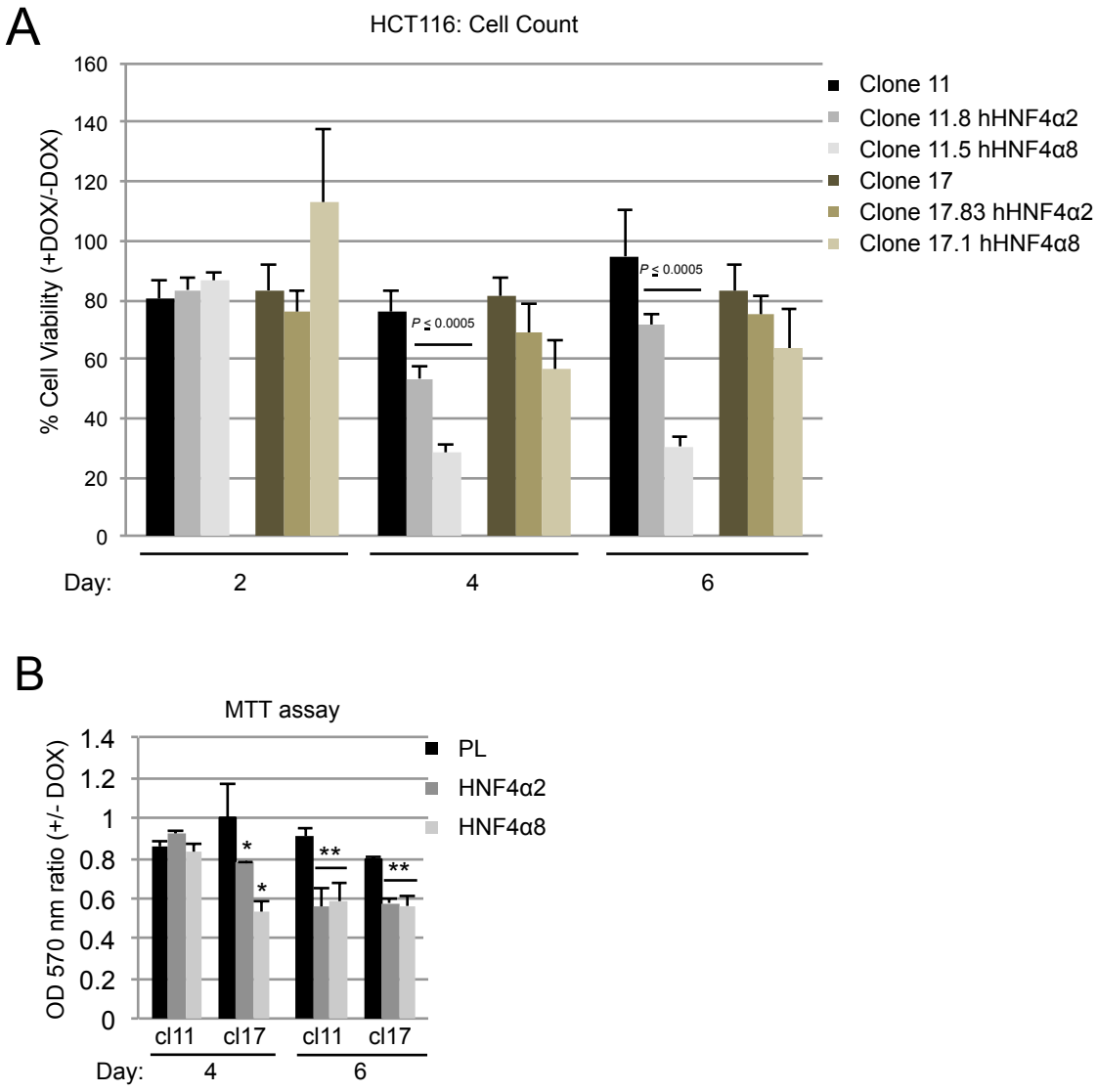

A
HCT116: Cell Count
% Cell Viability (+DOX/-DOX)
160
140
120
100
80
60
40
20
0
P ≤ 0.0005
P ≤ 0.0005
Day:
2
4
6
Clone 11
Clone 11.8 hHNF4α2
Clone 11.5 hHNF4α8
Clone 17
Clone 17.83 hHNF4α2
Clone 17.1 hHNF4α8
B
MTT assay
OD 570 nm ratio (+/- DOX)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
*
*
**
**
cl11
cl17
cl11
cl17
Day:
4
6
PL
HNF4α2
HNF4α8

Figure 3.1.2 HNF4α2 and HNF4α8 both decrease cellular proliferation in HCT116 lines.

(A) Stable inducible HCT116 cell lines (Parental, clones 11 and 17; HNF4α2, clones 11.8 and 17.83; HNF4α8, clones 11.5 and 17.1) seeded at 1.0 x 10^4 cells per well of a 12-well plate were treated with 0.5 µg/mL DOX 24 hours after plating (Day 0). Every 48 h cells were counted using a Beckman Coulter Counter. Media with fresh DOX was changed every 48 hours. Bar graph represents mean$\pm$SD of triplicate samples from one independent experiment. (B) HCT116 cell lines were seeded at 1.5-2.25 x 10^6 cells per 100-mm plate and, 24 hours later, treated with 0.5 µg/mL DOX. Forty-eight hours after induction, cells were transferred to a 96-well plate and subjected to the MTT assay. Bar graph represents mean$\pm$SD of triplicate samples from one experiment. * *P-value* $\leq$ 0.0007; ** *P-value* $\leq$ 0.00007.

Dr. Karthikeyani Chellappa performed the cell count and MTT assays.

Fig. 3.1.3

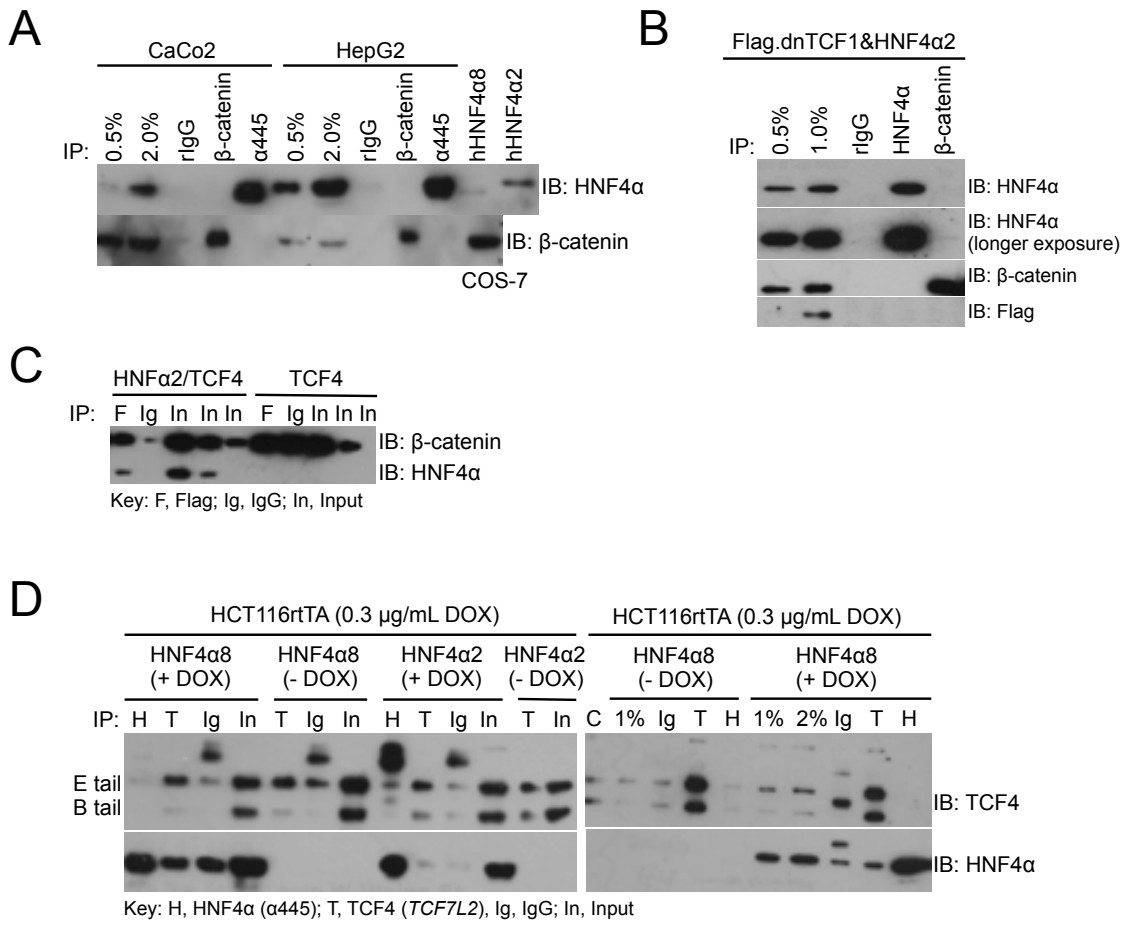

A
CaCo2
HepG2
IP:
0.5%
2.0%
rIgG
β-catenin
α445
0.5%
2.0%
rIgG
β-catenin
α445
hHNF4α8
hHNF4α2
IB: HNF4α
IB: β-catenin
COS-7
B
Flag.dnTCF1&HNF4α2
IP:
0.5%
1.0%
rIgG
HNF4α
β-catenin
IB: HNF4α
IB: HNF4α
(longer exposure)
IB: β-catenin
IB: Flag
C
HNFα2/TCF4
TCF4
IP:
F
Ig
In
In
In
F
Ig
In
In
In
IB: β-catenin
IB: HNF4α
Key: F, Flag; Ig, IgG; In, Input
D
HCT116rtTA (0.3 μg/mL DOX)
HCT116rtTA (0.3 μg/mL DOX)
HNF4α8
(+ DOX)
HNF4α8
(- DOX)
HNF4α2
(+ DOX)
HNF4α2
(- DOX)
HNF4α8
(- DOX)
HNF4α8
(+ DOX)
IP:
H
T
Ig
In
T
Ig
In
H
T
Ig
In
T
In
C
1%
Ig
T
H
1%
2%
Ig
T
H
E tail
B tail
IB: TCF4
IB: HNF4α
Key: H, HNF4α (α445); T, TCF4 (TCF7L2), Ig, IgG; In, Input

Figure 3.1.3 Co-IP of HNF4α and TCF/β-catenin.

(A) Co-IP analysis of HNF4α, β-catenin, and TCF4 in soluble nuclear extracts. Shown are IBs of input (0.5% and 2.0%) and co-IPs of endogenous HNF4α and β-catenin from CaCo2 and HepG2 NE with rIgG, anti-β-catenin, and α445 (HNF4α) antibodies; probed as indicated. (B) Co-IP as in (A) but with NE from COS-7 cells ectopically expressing HNF4α2 and Flag.dnTCF1 or endogenous β-catenin expression. IB analysis of HNF4α, Flag.dnTCF1, and β-catenin. (C) Co-IP of NE from COS-7 cells ectopically expressing HNF4α2 and Flag.TCF4; as indicated F, Flag antibody; In, input; and Ig, normal IgG control. (D) Co-IP of HNF4α and TCF4 from Tet-On inducible HCT116 cell lines expressing either HNF4α2 (clone 11.8) or HNF4α8 (clone 11.5), and as indicated, 24 hours induction with 0.3 μg/mL DOX. IB analyses of HNF4α, β-catenin, dnTCF1, and TCF4 from NEs of Co-IP samples. Antibodies for: H, HNF4α (α445), T, TCF4 (TCF7L2), F, Flag, Ig, IgG. In, input (0.1-0.5%).

Fig. 3.1.4

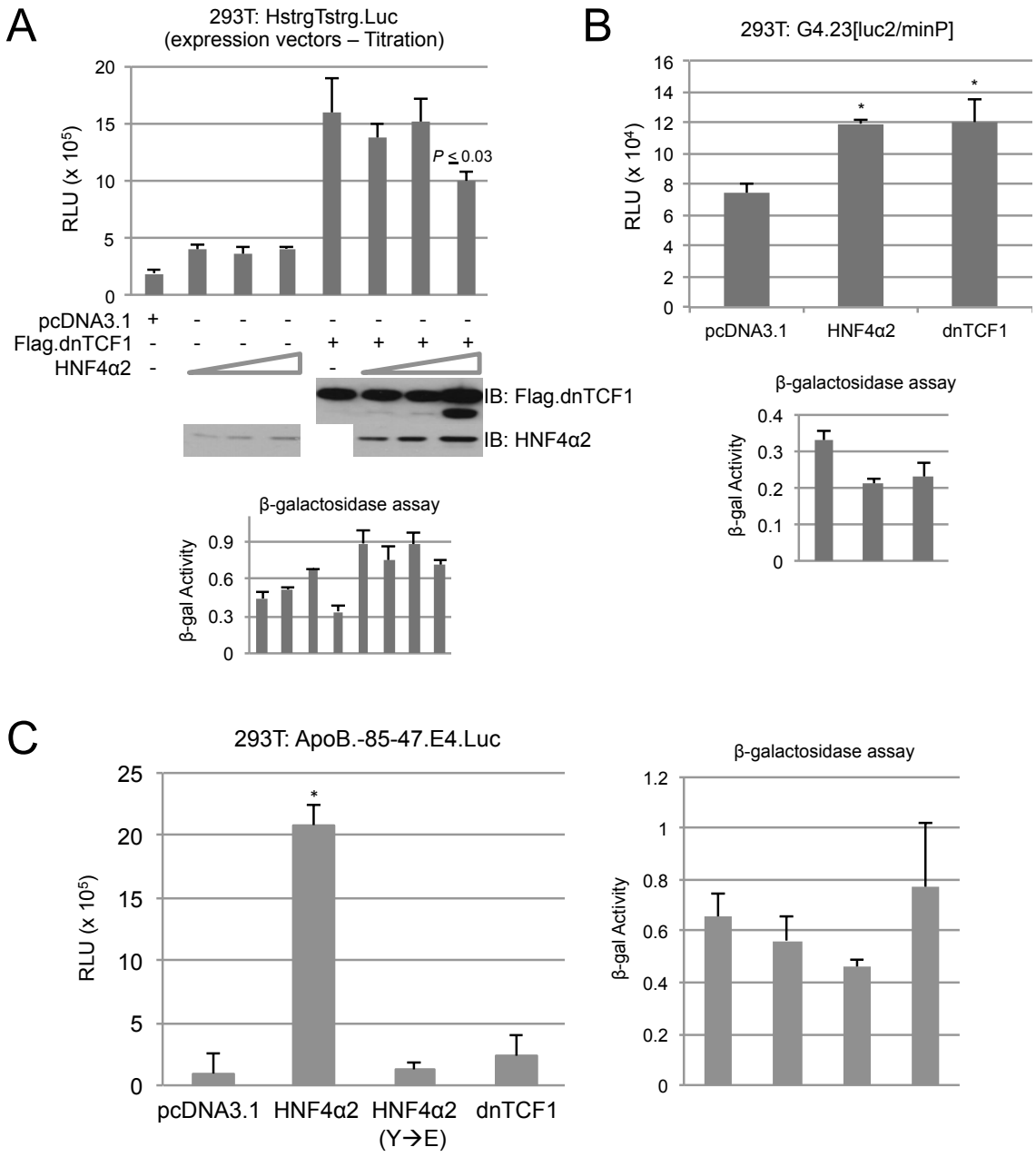

A
293T: HstrgTstrg.Luc
(expression vectors – Titration)
RLU (x 10^5)
P ≤ 0.03
pcDNA3.1
Flag.dnTCF1
HNF4α2
IB: Flag.dnTCF1
IB: HNF4α2
β-galactosidase assay
β-gal Activity
B
293T: G4.23[luc2/minP]
RLU (x 10^4)
pcDNA3.1
HNF4α2
dnTCF1
β-galactosidase assay
β-gal Activity
C
293T: ApoB.-85-47.E4.Luc
RLU (x 10^5)
pcDNA3.1
HNF4α2
HNF4α2 (Y→E)
dnTCF1
β-galactosidase assay
β-gal Activity

Figure 3.1.4 dnTCF1 activates transcription on the HstrgTstrg but not the ApoB element in HEK293T cells.

(A and B) HEK293T cells transfected with 0.5 μg of reporter construct containing the HstrgTstrg element (pHstrgTstrg.Luc) or no element (G4.23[luc2/minP]) and different concentrations (HNF4α2 – 60, 80, and 120 ng; dnTCF1 – 400 ng) or one concentration (80ng) of expression vectors. IB showing HNF4α2 and Flag.dnTCF1 expressions using the anti-HNF4α (P1/P2) and anti-Flag antibodies of WCE (20 μg) from HEK293T. Bar graphs represent mean±SD of triplicate samples from one experiment. (C and D) HEK293T cells transfected with 0.5 μg of the ApoB.-85-47.E4.Luc reporter plasmid and 80 ng of each expression vector (HNF4α2, HNF4α2 (Y63E) mutant, and Flag.dnTCF1). Bar graph represent mean±SD of triplicate samples from one experiment. β-galactosidase activity shown in a separate graph with each luciferase activity.

Fig. 3.1.5

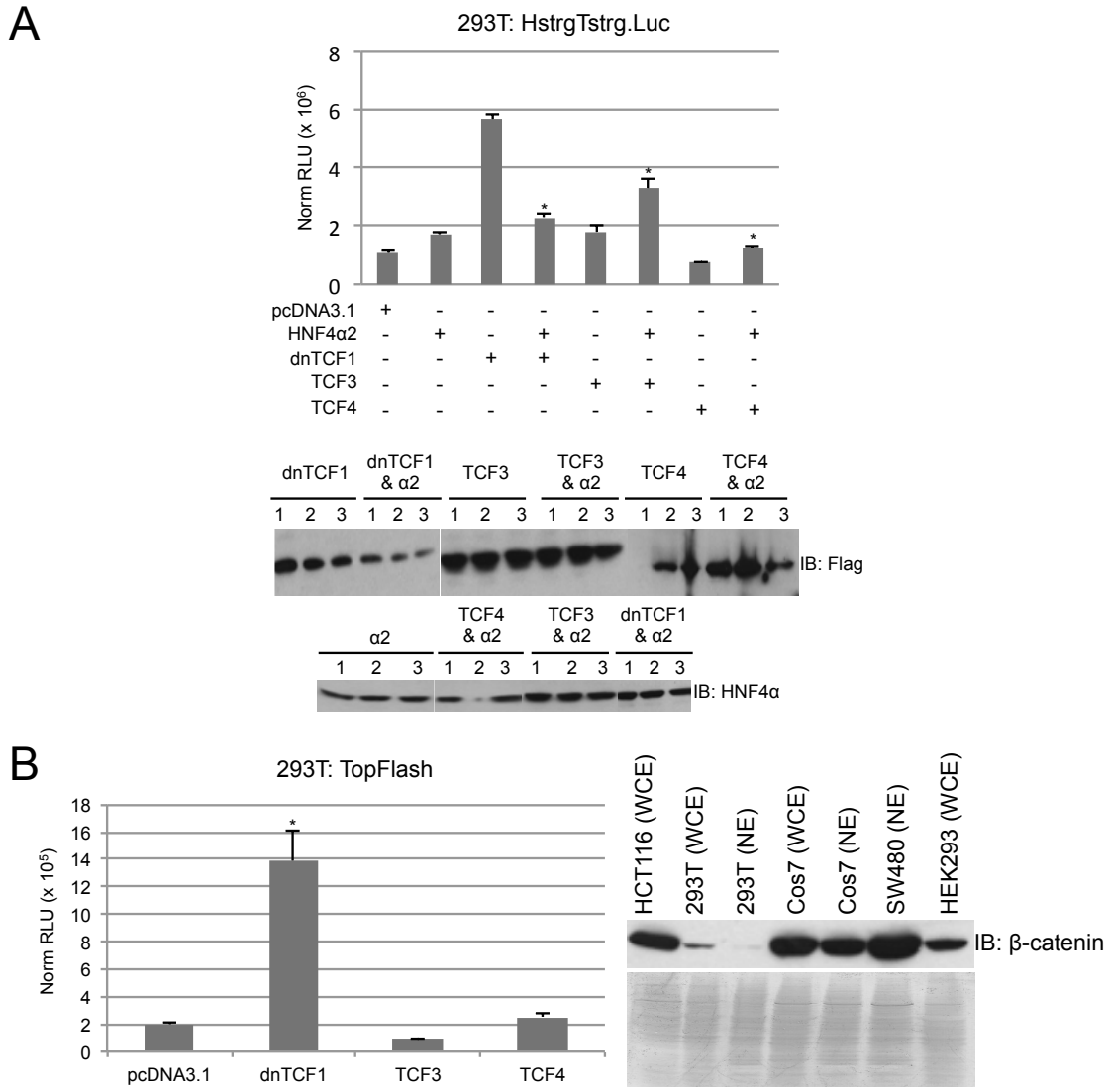

A
293T: HstrgTstrg.Luc
Norm RLU (x 10^6)
pcDNA3.1
HNF4α2
dnTCF1
TCF3
TCF4
dnTCF1
dnTCF1 & α2
TCF3
TCF3 & α2
TCF4
TCF4 & α2
IB: Flag
α2
TCF4 & α2
TCF3 & α2
dnTCF1 & α2
IB: HNF4α
B
293T: TopFlash
Norm RLU (x 10^5)
pcDNA3.1
dnTCF1
TCF3
TCF4
HCT116 (WCE)
293T (WCE)
293T (NE)
Cos7 (WCE)
Cos7 (NE)
SW480 (NE)
HEK293 (WCE)
IB: β-catenin

Figure 3.1.5 dnTCF1 but not TCF3 and TCF4 activates HstrgTstrg and TopFlash reporter constructs.

(A and B) HEK293T cells transfected with 0.5 µg of HstrgTstrg or TopFlash reporter construct and 80 ng of expression vector(s) (HNF4α2, Flag.dnTCF1, Flag.TCF4, and/or Flag.TCF3). IB analysis of HNF4α and the three different Flag tagged TCFs from WCE of transfected HEK293T cells in (A). IB analysis of β-catenin from WCE or NE of HCT116, HEK293T, COS-7, SW480, HEK293 cells in (B). Bar graphs represent mean±SD of triplicate samples from one experiment. * *P-value* ≤ 0.05; A: HNF4α2+dnTCF1 vs. dnTCF1; HNF4α2+TCF3 vs. TCF3; HNF4α2+TCF4 vs. TCF4; B: dnTCF1 vs. pcDNA3.1, TCF4, or TCF3.

Fig. 3.1.6

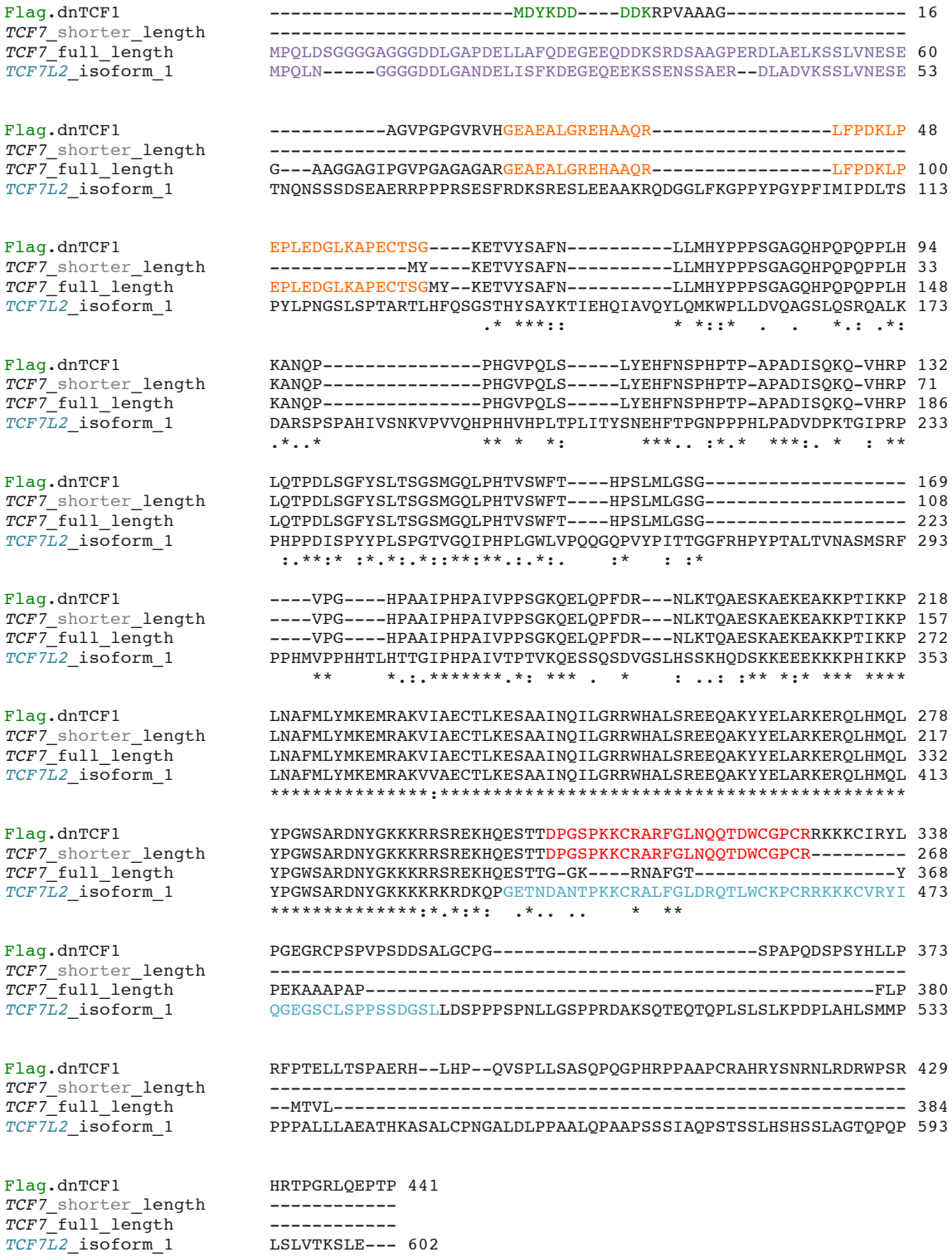

Flag.dnTCF1 ----------MDYKDD----DDKRPVAAAG---------- 16
TCF7_shorter_length ----------
TCF7_full_length MPQLDSGGGGAGGGDDLGAPDELLAFQDEGEEQDDKSRDSAAGPERDLAELKSSLVNESE 60
TCF7L2_isoform_1 MPQLN-----GGGGDDLGANDELISFKDEGEQEEKSSENSSAER--DLADVKSSLVNESE 53
Flag.dnTCF1 ----------AGVPGPGVRVHGEAEALGREHAAQR----------LFPDKLP 48
TCF7_shorter_length ----------
TCF7_full_length G---AAGGAGIPGVPGAGAGARGEAEALGREHAAQR----------LFPDKLP 100
TCF7L2_isoform_1 TNQNSSSDSEAERRPPPRSESFRDKSRESLEEAAKRQDGGLFKGPPYPGYPFIMIPDLTS 113
Flag.dnTCF1 EPLEDGLKAPECTSG----KETVYSAFN----------LLMHYPPPSGAGQHPQPQPPLH 94
TCF7_shorter_length ----------MY----KETVYSAFN----------LLMHYPPPSGAGQHPQPQPPLH 33
TCF7_full_length EPLEDGLKAPECTSGMY--KETVYSAFN----------LLMHYPPPSGAGQHPQPQPPLH 148
TCF7L2_isoform_1 PYLPNGSLSPTARTLHFQSGSTHYSAYKTIEHQIAVQYLQMKWPLLDVQAGSLQSRQALK 173
Flag.dnTCF1 KANQP----------PHGVPQLS-----LYEHFNSPHPTP-APADISQKQ-VHRP 132
TCF7_shorter_length KANQP----------PHGVPQLS-----LYEHFNSPHPTP-APADISQKQ-VHRP 71
TCF7_full_length KANQP----------PHGVPQLS-----LYEHFNSPHPTP-APADISQKQ-VHRP 186
TCF7L2_isoform_1 DARSPSPAHIVSNKVPVVQHPHHVHPLTPLITYSNEHFTPGNPPPHLPADVDPKTGIPRP 233
Flag.dnTCF1 LQTPDLSGFYSLTSGSMGQLPHTVSWFT----HPSLMLGSG---------- 169
TCF7_shorter_length LQTPDLSGFYSLTSGSMGQLPHTVSWFT----HPSLMLGSG---------- 108
TCF7_full_length LQTPDLSGFYSLTSGSMGQLPHTVSWFT----HPSLMLGSG---------- 223
TCF7L2_isoform_1 PHPPDISPYYPLSPGTVGQIPHPLGWLVPQQGQPVYPITTGGFRHPYPTALTVNASMSRF 293
Flag.dnTCF1 ----VPG----HPAAIPHPAIVPPSGKQELQPFDR---NLKTQAESKAEKEAKKPTIKKP 218
TCF7_shorter_length ----VPG----HPAAIPHPAIVPPSGKQELQPFDR---NLKTQAESKAEKEAKKPTIKKP 157
TCF7_full_length ----VPG----HPAAIPHPAIVPPSGKQELQPFDR---NLKTQAESKAEKEAKKPTIKKP 272
TCF7L2_isoform_1 PPHMVPPHHTLHTTGIPHPAIVTPTVKQESSQSDVGSLHSSKHQDSKKEEEKKKPHIKKP 353
Flag.dnTCF1 LNAFMLYMKEMRAKVIAECTLKESAAINQILGRRWHALSREEQAKYYELARKERQLHMQL 278
TCF7_shorter_length LNAFMLYMKEMRAKVIAECTLKESAAINQILGRRWHALSREEQAKYYELARKERQLHMQL 217
TCF7_full_length LNAFMLYMKEMRAKVIAECTLKESAAINQILGRRWHALSREEQAKYYELARKERQLHMQL 332
TCF7L2_isoform_1 LNAFMLYMKEMRAKVVAECTLKESAAINQILGRRWHALSREEQAKYYELARKERQLHMQL 413
Flag.dnTCF1 YPGWSARDNYGKKKRRSREKHQESTTDPGSPKKCRARFGLNQQTDWCGPCRKKKCIRYL 338
TCF7_shorter_length YPGWSARDNYGKKKRRSREKHQESTTDPGSPKKCRARFGLNQQTDWCGPCR---------- 268
TCF7_full_length YPGWSARDNYGKKKRRSREKHQESTTG-GK----RNAFGT----------Y 368
TCF7L2_isoform_1 YPGWSARDNYGKKKKRKRDKQPGETNDANTPKKCRALFGLDRQTLWCKPCRRKKKCVRYI 473
Flag.dnTCF1 PGEGRCPSPVPSDDSALGCPG----------SPAPQDSPSYHLLP 373
TCF7_shorter_length ----------
TCF7_full_length PEKAAAPAP----------FLP 380
TCF7L2_isoform_1 QGEGSCLSPPSSDGSLLDSPPPSPNLLGSPPRDAKSQTEQTQPLSLSLKPDPLAHLSMMP 533
Flag.dnTCF1 RFPTELLTSPAERH--LHP--QVSPLLSASQPQGPHRPPAAPCRAHRYSNRNLRDRWPSR 429
TCF7_shorter_length ----------
TCF7_full_length --MTVL---------- 384
TCF7L2_isoform_1 PPPALLLAEATHKASALCPNGALDLPPAALQPAAPSSSIAQPSTSSLHSHSSLAGTQPQP 593
Flag.dnTCF1 HRTPGRLQEPTP 441
TCF7_shorter_length ----------
TCF7_full_length ----------
TCF7L2_isoform_1 LSLVTKSLE--- 602

Figure 3.1.6 Similarities and differences between the Flag.dnTCF1 and other TCF1 isoforms.

Amino Acid sequence alignment from ClustalW2* (EMBL-EBI) of the Flag.dnTCF1 (dominant negative form of *TCF7*), shorter length TCF1 (*TCF7)* isoform 4, full-length TCF1 (*TCF7*), and full-length TCF4 (*TCF7L2*) isoform 1. All sequences except the Flag.dnTCF1 were obtained from the NCBI website (http://www.ncbi.nlm.nih.gov). The Waterman's lab at UCI provided the Flag.dnTCF1 amino acid sequences. Color-code: green, Flag epitope; orange, sequences shared between full-length TCF1 and Flag.dnTCF1; purple, β-catenin binding domain; turquoise, E tail; red, sequences shared between TCF1 shorter length and Flag.dnTCF1.

*ClustalW2. European Bioinformatics Institute. EMBL-EBI 2014. http://www.ebi.ac.uk/Tools/msa/clustalw2/. Accessed 17 November 2014.

Fig. 3.1.7

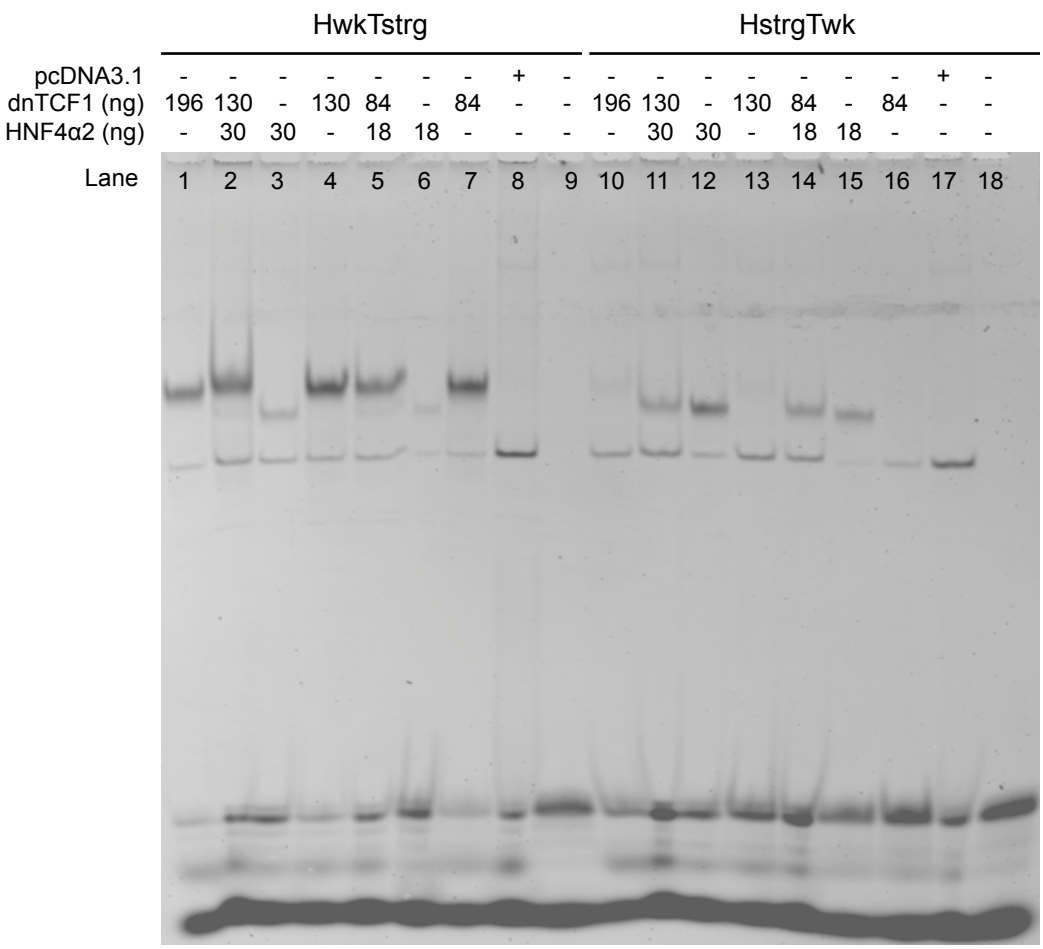


Figure 3.1.7 HwkTstrg and HstrgTwk binding sequences.

NEs from COS-7 cells ectopically expressing either the human HNF4α2 only, the human Flag tagged dnTCF-1 only, or both were used in the gel shift reaction. Protein concentration was determined before performing the shift reaction. Indicated amount of protein(s) (HNF4α2, 18 and 30 ng; dnTCF1, 84, 130, and 196 ng) was used in the shift reaction with either HstrgTwk or HwkTstrg fluorescence probe.

Fig. 3.1.8

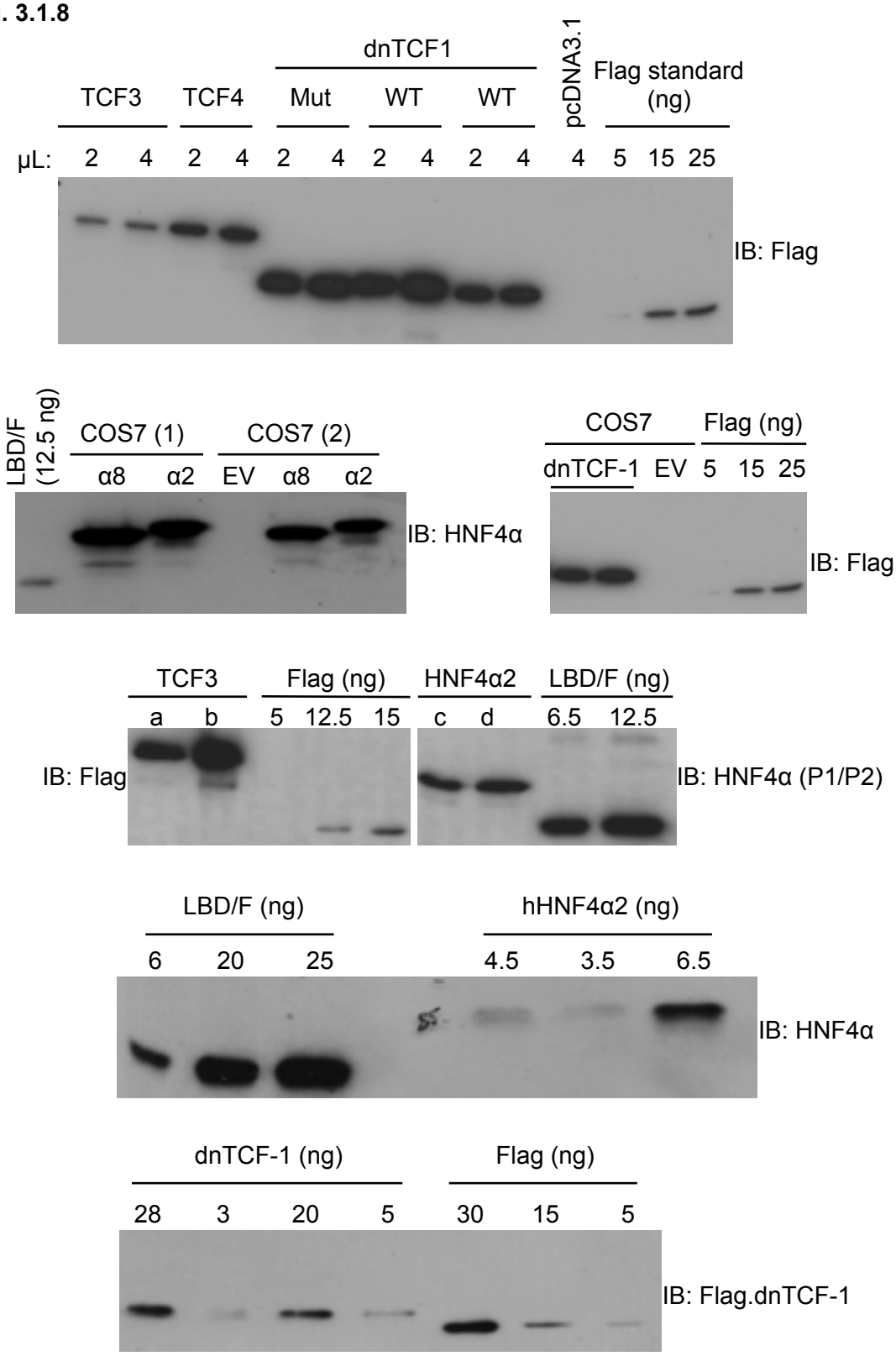

dnTCF1
pcDNA3.1
Flag standard (ng)
TCF3
TCF4
Mut
WT
WT
µL: 2 4 2 4 2 4 2 4 2 4 4 5 15 25
IB: Flag
LBD/F (12.5 ng)
COS7 (1)
COS7 (2)
α8 α2 EV α8 α2
IB: HNF4α
COS7
Flag (ng)
dnTCF-1 EV 5 15 25
IB: Flag
TCF3
Flag (ng)
HNF4α2
LBD/F (ng)
a b 5 12.5 15 c d 6.5 12.5
IB: Flag
IB: HNF4α (P1/P2)
LBD/F (ng)
hHNF4α2 (ng)
6 20 25 4.5 3.5 6.5
IB: HNF4α
dnTCF-1 (ng)
Flag (ng)
28 3 20 5 30 15 5
IB: Flag.dnTCF-1

Figure 3.1.8 IBs of HNF4α and dnTCF1 from samples used in the fluorescence gel shifts.

Nuclear extracts from COS7 cells that expressed either the human HNF4α2, Flag.dnTCF1, Flag.TCF4, or Flag.TCF3. The HNF4α LBD/F peptide and the Flag-BAP fusion protein were used to quantify the amount of protein expressed in the NE samples. EV – empty vector. Top IB, samples prepared by the Waterman's lab at UCI (TCF3, TCF4, mutant dnTCF1, and WT dnTCF1, center). Center IB, a, ~1.5 ng/μL (30 ng in 20 μL of NE loaded in lane); c, ~5.5 ng/μL (1 μL of NE loaded in lane).

Reference

Arce L, Yokoyama NN, Waterman ML. 2006. Diversity of LEF/TCF action in development and disease. *Oncogene* **25**: 7492-7504.

Cattin AL, Le Beyec J, Barreau F, Saint-Just S, Houllier A, Gonzalez FJ, Robine S, Pincon-Raymond M, Cardot P, Lacasa M et al. 2009. Hepatocyte nuclear factor 4alpha, a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. *Molecular and cellular biology* **29**: 6294-6308.

Colletti M, Cicchini C, Conigliaro A, Santangelo L, Alonzi T, Pasquini E, Tripodi M, Amicone L. 2009. Convergence of Wnt signaling on the HNF4alpha-driven transcription in controlling liver zonation. *Gastroenterology* **137**: 660-672.

Datto MB, Yu Y, Wang XF. 1995. Functional analysis of the transforming growth factor beta responsive elements in the WAF1/Cip1/p21 promoter. *The Journal of biological chemistry* **270**: 28623-28628.

Faiola F, Liu X, Lo S, Pan S, Zhang K, Lymar E, Farina A, Martinez E. 2005. Dual regulation of c-Myc by p300 via acetylation-dependent control of Myc protein turnover and coactivation of Myc-induced transcription. *Molecular and cellular biology* **25**: 10220-10234.

Gougelet A, Torre C, Veber P, Sartor C, Bachelot L, Denechaud PD, Godard C, Moldes M, Burnol AF, Dubuquoy C et al. 2014. T-cell factor 4 and beta-catenin chromatin occupancies pattern zonal liver metabolism in mice. *Hepatology* **59**: 2344-2357.

Grumolato L, Liu G, Haremaki T, Mungamuri SK, Mong P, Akiri G, Lopez-Bergami P, Arita A, Anouar Y, Mlodzik M et al. 2013. beta-Catenin-independent activation of TCF1/LEF1 in human hematopoietic tumor cells through interaction with ATF2 transcription factors. *PLoS genetics* **9**: e1003603.

Hoverter NP, Ting JH, Sundaresh S, Baldi P, Waterman ML. 2012. A WNT/p21 circuit directed by the C-clamp, a sequence-specific DNA binding domain in TCFs. *Molecular and cellular biology* **32**: 3648-3662.

Hwang-Verslues WW, Sladek FM. 2008. Nuclear receptor hepatocyte nuclear factor 4alpha1 competes with oncoprotein c-Myc for control of the p21/WAF1 promoter. *Molecular endocrinology* **22**: 78-90.

Maeda Y, Seidel SD, Wei G, Liu X, Sladek FM. 2002. Repression of hepatocyte nuclear factor 4alpha tumor suppressor p53: involvement of the ligand-binding domain and histone deacetylase activity. *Molecular endocrinology* **16**: 402-410.

Merrill BJ, Gat U, DasGupta R, Fuchs E. 2001. Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. *Genes & development* **15**: 1688-1705.

Pereira L, Yi F, Merrill BJ. 2006. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Molecular and cellular biology* **26**: 7479-7491.

Ruse MD, Jr., Privalsky ML, Sladek FM. 2002. Competitive cofactor recruitment by orphan receptor hepatocyte nuclear factor 4alpha1: modulation by the F domain. *Molecular and cellular biology* **22**: 1626-1638.

Sprowl S, Waterman ML. 2013. Past visits present: TCF/LEFs partner with ATFs for beta-catenin-independent activity. *PLoS genetics* **9**: e1003745.

Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C et al. 2006. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. *The Journal of pathology* **208**: 662-672.

Tang W, Dodge M, Gundapaneni D, Michnoff C, Roth M, Lum L. 2008. A genome-wide RNAi screen for Wnt/beta-catenin pathway components identifies unexpected roles for TCF transcription factors in cancer. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 9697-9702.

Van de Wetering M, Castrop J, Korinek V, Clevers H. 1996. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. *Molecular and cellular biology* **16**: 745-752.

Weise A, Bruser K, Elfert S, Wallmen B, Wittel Y, Wohrle S, Hecht A. 2010. Alternative splicing of Tcf7l2 transcripts generates protein variants with differential promoter-binding and transcriptional activation properties at Wnt/beta-catenin targets. *Nucleic acids research* **38**: 1964-1981.

Yang M, Li SN, Anjum KM, Gui LX, Zhu SS, Liu J, Chen JK, Liu QF, Ye GD, Wang WJ et al. 2013. A double-negative feedback loop between Wnt-beta-catenin signaling and HNF4alpha regulates epithelial-mesenchymal transition in hepatocellular carcinoma. *Journal of cell science* **126**: 5692-5703.

Chapter 4

Ectopic Expression of a Lineage-Specific Transcription Factor HNF4α in mES Cells: *Potential Differences between the Promoter-driven (P1- and P2-) HNF4α in Directing mES Differentiation.*

Abstract

In order to study the role of a transcription factor (TF) in embryogenesis a decade ago, one would either generate a knockout mouse or differentiate embryonic stem cells (ESCs) into a specific lineage with defined growth factors. Finally, the ability to introduce four TFs into somatic cells to re-program them back to a pluripotent state made it possible to study the role of any TFs in de- and re-differentiation. Others began to ectopically express one or more TFs in fibroblast or ES cells to transform or differentiate those cells into a specific lineage that expresses the TFs of interest and study their role in the process. Here, we introduced HNF4α into mouse ESCs and showed that ectopic expression of HNF4α drives the cells to differentiate without the use of defined medium. HNF4α is a member of the nuclear receptor superfamily that plays a critical role in embryogenesis. There are two promoters to the *HNF4A* gene (P1 and P2) that drive expression in different tissues and can give rise to splice variants that differ in their N-terminal region. P1-HNF4α is expressed in the fetal and adult liver, whereas the P2-HNF4α is expressed in the fetal but not the adult liver and the pancreas. Both the hepatic and pancreatic lineages arise from the same progenitor cells. To determine the role of P1- and P2-HNF4α in mES-derived endodermal differentiation, we generated Tet-On inducible mES cell lines that express either HNF4α2 (P1) or HNF4α8 (P2) under control of doxycycline (DOX). When the HNF4α2 or HNF4α8 line was treated with DOX for 72 hours, the cells differentiate and express HNF3α/β (Foxa1/2), another liver-enriched TF essential for liver development. Interestingly, the expression of HNF4α is lost over time in the cells, but the morphological changes remain, suggesting that HNF4α can initiate the differentiation but that the downstream effect is independent of the factor. In order to determine what type of cells were induced, we performed RNA-Seq at 24 and

72 hours DOX induction for both the HNF4α2 and HNF4α8 lines. These HNF4α-induced ESCs may provide a better understanding into the early stages of endodermal differentiation.

Introduction

The discovery and successful derivation of embryonic stem cells (ESCs) from mouse to subsequently human blastocysts made it possible to recapitulate and study the events of early development (Martin 1981; Thomson et al. 1998). To elucidate the role of transcription factors (TFs) in development, one could either generate a knockout mouse model or direct ESC differentiation into the specific lineage that expresses the TFs of interest using defined growth factors (GFs) (Watt et al. 2003; Keller 2005). Forced expression of a gene of interest in ESCs was not considered; in this case Oct3/4, Sox2, c-Myc, and Klf4 into normal fibroblasts to convert the cells into induced pluripotent stem cells (iPSC). The development of iPSCs by Takahashi and Yamanaka in 2006 led to the idea that cellular transformation from one cell type to another could be made possible through the introduction of one or more TFs into the cells (Takahashi and Yamanaka 2006). This led to several studies that managed to differentiate hepatocytes from ES/iPS and fibroblast cells (reviewed in Montserrat et al. 2013).

Two groups introduced two or more liver-specific transcription factors (LSTFs) into fibroblasts to transform these cells into induced hepatocyte-like (iHep) cells (Huang et al. 2011; Sekiya and Suzuki et al. 2011). A year later, two more groups overexpressed LSTFs in ESCs/iPSCs via lentivirus to promote hepatocyte differentiation (Liu et al. 2013; Takayama et al. 2012). These groups showed that multiple factors, with the addition of potent GFs, are required to drive differentiation. Another group took a

different approach. They generated tetracycline-regulatable ES lines that overexpressed 137 TFs, singularly, into each line and identified TFs that could direct cell differentiation. They found that 63 of 137 TFs could disturb the pluripotent state and drive differentiation toward certain lineages (Yamamizu et al. 2013).

Hepatocyte nuclear factor (HNF) 4α is a member of the nuclear receptor superfamily that plays an essential role in gut, pancreatic, and hepatic development (Maestro et al. 2007; Duncan et al. 1994 and 1997; Garrison et al. 2006; Watt et al. 2007). Knockout of mouse *Hnf4a* gene is embryonic lethal at embryonic day nine (E9) of gestation (Chen et al. 1994). Importantly, HNF4α is required for the proliferation capacity of β-cells in the pancreas and the architectural structure of the embryonic and adult liver (Hansen et al. 2002; Hayhurst et al. 2001; Garrison et al. 2006; Gupta et al. 2005). In addition, *Hnf4a* expression is found in other tissues including the kidney and stomach, although less is known about its role in these tissues. HNF4α is responsible for regulating genes involved in insulin secretion, gluconeogenesis, bile acid synthesis, lipid metabolism, and xenobiotic and drug metabolism (Sladek and Seidel 2001; Gupta et al. 2005; Rhee et al. 2003; Inoue et al. 2006; Chen et al. 2003). Through the utilization of alternative promoters (P1 and P2) and 3' splicing events at exon 8 and 9, there are nine potential isoforms of the *HNF4A* gene (Taraviras et al. 1994; Furuta et al. 1997). The P1 and P2 promoters drive the expression of the isoforms HNF4α1-HNF4α6 and HNF4α7-HNF4α9, respectively. Three additional isoforms HNF4α10-12 transcribed by the P2-promoter have been identified as well (Huang J et al. 2009). Isoforms HNF4α2, HNF4α5, and HNF4α8 have an extra 10 amino acid insert in exon 9 that is absent in HNF4α1, HNF4α4, and HNF4α7. Finally, HNF4α3, HNF4α6, and HNF4α9 have a truncated C-terminal and an alterative splicing event in exon 8 (Harries et al. 2008). The promoter-

driven HNF4α isoforms exhibit temporal and spatial expression and have differential transcriptional properties (Torres-Padilla et al. 2001, 2002; Torres-Padilla and Weiss 2003; Briancon et al. 2004; Sladek et al. 1999).

The development of the liver and pancreas arise from the same precursor cells in the anterior endoderm. Both the pancreatic and hepatic cells bud outward from the ventral foregut. It is still not clear, though, how the ventral foregut endoderm is specified to bifurcate into the pancreas and liver. One group elegantly showed that the default fate of the ventral foregut endoderm cells is a pancreatic program and the cells that later come in contact with the cardiac mesoderm develop into the hepatic lineage (Deutsch et al. 2001). Signals released from the cardiac mesoderm include fibroblast growth factors (FGFs) and sonic hedgehog (Shh); these factors inhibit pancreatic formation (Deutsch et al. 2001).

During development, both P1 and P2-HNF4α are expressed in the fetal liver and pancreas (Harries et al. 2008; Torres-Padilla et al. 2001; Briancon et al. 2004), while it is not clear whether both isoforms are present in the ventral foregut endoderm cells. Little is known about how one promoter-driven isoform is lost while the other remains expressed during the postnatal development of the hepatic and pancreatic organs nor how that program remains set in place (Briancon et al. 2004). In the adult stage, the liver expresses only the P1-HNF4α1/α2 isoforms, whereas the β-cells in the pancreas express only the P2-HNF4α7/α8 isoforms (Briancon and Weiss 2006; Harris et al. 2008). Two questions arise from what we still do not know: (1) how the expression of P1 or P2-HNF4α is lost in the adult pancreas or liver, respectively; and (2) how HNF4α isoforms drive hepatic and pancreatic differentiation in ES cells. To address these questions, we established a Tet-On system in a mouse embryonic stem (mES) D3 line to express

either P1-HNF4a2 or P2-HNF4a8 under the control of doxcycline (DOX). We show that these cells, upon HNF4α induction, exhibit morphological changes that reflect changes in gene expression. Although another group has overexpressed HNF4α in mouse ES cells in an inducible manner, they have not looked at the isoforms (Yamamizu et al. 2013). Utilizing a similar approach, we introduced promoter-driven HNF4α variants into mouse D3 ESCs and identify their individual role in early development.

Materials and Methods

Cell Culture

Mouse embryonic stem (mES) D3 cells (a gift from Dr. Prudence Talbot at the University of California, Riverside) were initially cultured on irradiated mouse embryonic fibroblast (MEF) in mES medium that had the following composition: DMEM (Dulbecco's Modification of Eagle's Medium with 4.5 g/L glucose, L-glutamine, and pyruvate) supplemented with 15% Fetal Bovine Serum (FBS), 1% nonessential amino acids (NEAA), 1% penicillin/streptomycin (P/S), 1% sodium pyruvate, 1% L-glutamine, and 0.2% β-mercaptoethanol (β-mercap). For subsequent passages, the cells were cultured in MEF conditioned medium (CM), supplemented with recombinant mouse leukemia inhibitory factor (mLIF) (100-1000 U/mL) (Millipore; cat#ESG1107). The CM consisted of mES medium that had been cultured on irradiated MEFs for 24 h, collected and frozen for later use. Cells were maintained at 37^{o}C and 5% CO_2.

MEF Derivation

The MEFs, used for co-culture with mES cells and CM generation, were derived from a E13.5 embryos from mixed background mice. MEF derivations were done using

previously described protocols (Gardfield et al. 2010; WiCell Protocol 2003) with modifications: A pregnant dam at 13.5-14 days of gestation was sacrificed via CO_2 asphyxiation. The abdomen was sterilized with 95% ethanol and cut open. The uterine horns were removed and washed with 1 x phosphate buffered saline (PBS). Embryonic sacs were separated with forceps and scissors. The embryos were then washed with PBS. Heads and limbs were removed from the embryos. The remainder of the embryos were minced with a razor blade in 2 mL of trypsin solution (0.5%). Another 5 mL of trypsin was added to the minced tissue, which were incubated at 37^{o}C for 20-30 min; the cells were dislodged by pipetting up and down. The MEFs were cultured in MEF media (DMEM supplemented with 10% FBS, 1% nonessential amino acids, 1% P/S, 1% sodium pyruvate, 1% L-glutamine) in 100-mm plates or T75 flask for expansion and frozen for later use. Some MEFs were irradiated in the Faxitron machine at 90kVp for one hour to become feeders (see supplemental material for detail). Aliquots of irradiated MEFs were frozen and kept at -80^{o}C for later use. The University of California, Riverside, Institutional Animal Care and Use Committee (IACUC) guidelines were followed when handling mice.

Plasmid constructs

The full-length human HNF4α2 (NM_000457) and HNF4α8 (NM_175914.3) in pcDNA3.1 (pcDNA3.1.HNF4α2 & pcDNA3.1.HNF4α8) were gifts from Dr. Christophe Rachez at Pasteur Institute, Paris, France (Chartier et al. 1994; Eeckhoute et al. 2001). pTRE-HNF4α2 and pTRE-HNF4α8 were made by amplifying the human HNF4α2 and HNF4α8 constructs in the pcDNA3.1 (+) vector, using primers that contained EcoRI [plus a kozak sequence] (5'-HNF4α2_Koz.EcoRI: 5'-GGAATTCCCCACCATGGATATGGCC-3'; 5'-

HNF4α8_Koz.EcoRI: 5'-GGAATTCCCCACCATGGTCAGCGTG-3') or BamHI sites (3'-N1C465.BamHI: 5'-GCGGGATCCCGCTAGATAACTTCCTGCTT-3') that recognized the 5' and 3' ends of the HNF4α cDNA, and cloning the PCR fragments into the pTRE-Tight vector (Clontech) at the respective (EcoRI/BamHI) sites. The pCAG-rtTA (pCAG-rtTA-IR-PURO) and pTRE-RFP (pTRE-Red Max C) plasmids were gifts from Dr. Chee-Gee Liew (Liew et al. 2007; Xia et al. 2008).

Generations of the Tet-On HNF4α Inducible Cell Lines in mES cells

mES cells were seeded at 5 x 10^4 cells per well of 6-well plate on irradiated MEFs in mES medium and, 24 h later, transfected with 1 μg of linearized pCAG-rtTA-IR-PURO DNA via HiFect (LONZA). Cells were selected in CM with 1.5 μg/mL puromycin. Colonies were picked and transferred to gelatin-coated (0.1%) 12-well plates. Confluent colonies were passaged into two wells of gelatin-coated 12-well plates; one plate was expanded, while the second plate was tested for positive integration of the rtTA vector. pTRE-RFP (1 μg per 12-well) was transfected into each clone and, 24 h later, the transfected cells were treated with 1 μg/mL of doxycycline (DOX). The following day, red fluorescent protein (RFP) was observed under a Nikon fluorescent microscope at the University of California Riverside Stem Cell Core Facility. Clone 3.1 was selected for subsequent step (see Chapter 2, Fig. 2.4A). Cells were seeded at 2.5-3.5 x 10^5 cells per well of gelatin-coated 6-well plate. The following day, linearized pTRE-HNF4α2 or pTRE-HNF4α8 plus a DNA fragmented containing the Neo^R gene cut from the Tet-On vector (Clontech) were transfected at a 10:1 ratio respectively into the cells via lipofectamine 2000 (Invitrogen). Feeders (8 x 10^6) were plated on top of the cells for support. Forty-eight hours later, the cells were selected with 50 and 1.5 μg/mL of G418 and puromycin,

respectively. Colonies were transferred to a MEF-coated 24-well plate and tested for positive HNF4α2 or HNF4α8 expression following induction with 1 μg/mL of DOX using immunoblot (IB) analysis. The HNF4α2 (α2), HNF4α8 (α8), and parental (PL) lines were maintained in CM and at 37^{o}C incubator with 5% CO_2.

Immunoblot (IB) analysis

Protein extracts were separated by 10% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to immobilon (Millipore) as is described in Chapter 3. The bradford assay was performed to determine protein concentration and whole-cell lysates (WCE) or nuclear extracts (NE) (2-20 μg of total protein) were loaded per lane as indicated. Coomassie staining of the blot was performed after IB to ensure equal loading of total protein across lanes. Blot was stained with 0.25% Coomassie in 45% methanol (MeOH) and 10% glacial acetic acid for 30 seconds and destained with 1 to 2 washes, 5 – 10 min each of 50% MeOH and 10% glacial acetic acid. The following antibodies were used: primary antibodies were mouse monoclonal anti-HNF4α P1/P2 (R&D, cat#PP-H1415-00) at 1:30,000 and 1:60,000; OCT4 (Santa Cruz, sc-9081), and HNF3α (Santa Cruz, sc6553) at 1:5000 and 1:1000, respectively. Secondary antibodies were horseradish peroxidase (HRP)-conjuated goat anti-rabbit (GαR-HRP) or goat anti-mouse (GαM-HRP) from the Jackson ImmunoResearch Laboratories (1:5000 to 1:10,000).

Immunofluorescent (IF) analysis

Cells were seeded at 6 x 10^5 cells per well of a gelatin-coated 24-well and treated with 0.3 μg/mL DOX for 24 h. The following day, cells were fixed with ~3.7 formaldehyde for

10 to 15 min at RT followed by 0.1 M glycine for 5 to 10 min. Cells were blocked in 10% goat, 10% donkey serum (GS & DS), and 0.2% Triton X-100 in PBS for 30 min at RT, and then with 1% GS and DS in PBS overnight at 4^{o}C. Primary antibodies used were: anti-HNF4α P1/P2 (R&D, PP-H1415-00) 1:1200 and OCT4 (Santa Cruz, sc9081) 1:200. Secondary antibodies were Alexa Fluor goat anti-mouse (1:500; Invitrogen); 4'6-diamidino-2-phenylindole (DAPI) was used at 1:1000 to visualize nuclei.

RNA-Seq analysis

Tet-On inducible mES clones (PL, HNF4α2, and HNF4α8) were seeded at 1.5 x 10^5 cells per well of (0.1%) gelatin-coated 12-well plates. Twenty-four hours later, cells were treated with 0.3 μg/mL of DOX. Cells were harvested 24 and 72 h after induction by adding 700 μL QIAzol Lysis Reagent (Qiagen) to the adherent cells. The miRNeasy Mini kit (Qiagen) was used to extract and purify total RNA; 4 μg of each RNA sample was used to generate a polyA+ RNA library using the Illumina TruSeq RNA Sample Prep v2 Kit (Cat# RS-122-2001). Libraries were submitted for 50 bp single end next generation sequencing by Illumina HiSEQ 2000 at the Genomics Core in the UCR Institute of Integrated Genome Biology (IIGB). A total of 24 samples (eight different conditions, each condition in triplicates) were multiplexed and sequenced in two lanes, each of which yielded ~192 M (million) reads or ~16 M reads per sample.

Results

Establishment of inducible HNF4α isoforms in mESCs

To investigate the role of the P1 and P2-HNF4α isoforms in pancreatic and hepatic development, we ectopically expressed the isoforms in mESCs under the control of a reverse tetracycline-controlled transcriptional activator (rtTA) that is constitutively expressed (Fig. 4.1A). Twenty-four hours after DOX treatment, at least ninety percent of the cells expressed HNF4α2 or HNF4α8 (Fig. 4.1A; Fig. 2.6C, Chapter 2). Similar levels of expressions for HNF4α2 and HNF4α8 were verified by IB and showed to be comparable to levels of expression in the adult mouse liver (Fig. 4.1B). The isoforms, driven by two alternative promoters (P1 and P2) that are ~45 kb apart, differ by ~16-38 aa in the A/B domain at the N-terminus, and migrate at different molecular weights (Fig. 4.1B).

Ectopic expression of HNF4α drives mESCs differentiation toward an epithelial-like phenotype

Ectopic expression of HNF4α induces mESC differentiation into an epithelial-like morphology. We observed distinguishable morphological change 72 hours after induction (Fig. 4.2A). We also observed HNF4α expression decreases after 60 hours (Fig. 4.2B). Interestingly, cells remained differentiated even after HNF4α expression is gone (Fig. 4.2A,B). This would suggest that the initial differentiation by HNF4α is irreversible thereafter. These differentiated cells expressed Foxa1/2 (*Hnf3a/b*), more in α2 than α8 line (Fig. 4.2C). Foxa1/2 is an endodermal marker that is pivotal in liver, pancreas, and gut development (Kaestner 2000) and is considered a “pioneer” transcription factor (TF) that opens the chromatin for other TFs, including HNF4α, to

access and regulate transcription in early murine development (Zaret 1999; Cirillo et al. 2002; Sérandour et al. 2011). During development, Foxa1/2 acts upstream of HNF4α; however, our cell lines (α2 and α8) induced with DOX for 72 hours showed the opposite. This would suggest that perhaps HNF4α could potentially be a pioneer factor as well (Kir et al. 2012).

(RNA-Seq)

To determine the identity of these differentiated cells, we performed RNA-Seq at 24 and 72 hours and looked for any dysregulated genes in a global scale. We are anticipating for the analysis of the RNA-Seq data.

Discussion

Although Yamamizu et al demonstrated that a single HNF4α alone can drive differentiation in mouse ES cells, they did not look at the isoform specific differences in ES differentiation (Yamamizu K et al. 2013). Here we showed that both HNF4α2 and HNF4α8 drive mES differentiation and decrease cell number, although HNF4α2 is more effective at inhibiting cell number than HNF4α8. At this point, we do not know what cell types the HNF4α2 and HNF4α8 lines differentiate into. We are waiting for the analysis of the RNA-Seq data.

Fig. 4.1

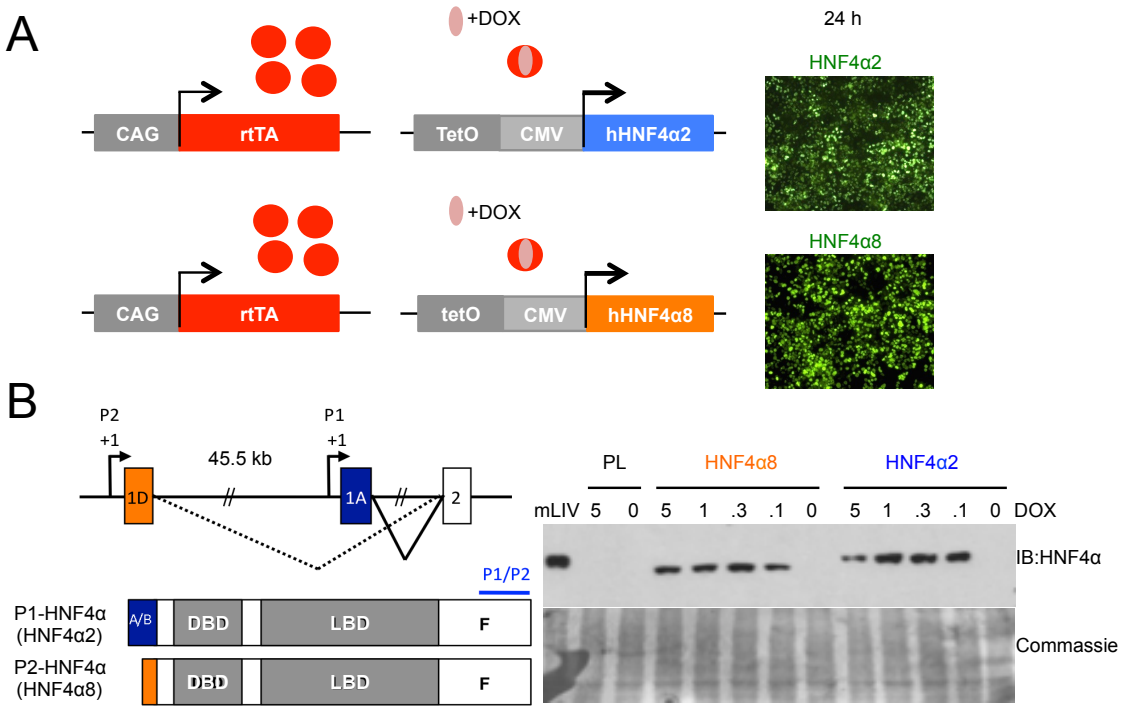

A
+DOX
24 h
CAG
rtTA
TetO
CMV
hHNF4α2
HNF4α2
+DOX
CAG
rtTA
tetO
CMV
hHNF4α8
HNF4α8
B
P2
+1
45.5 kb
P1
+1
1D
1A
2
P1/P2
P1-HNF4α
(HNF4α2)
A/B
DBD
LBD
F
P2-HNF4α
(HNF4α8)
DBD
LBD
F
PL
HNF4α8
HNF4α2
mLIV 5 0 5 1 .3 .1 0 5 1 .3 .1 0 DOX
IB:HNF4α
Commassie

Figure 4.1 Establishment of the DOX-inducible HNF4α isoforms in mESCs. (A) Schematic of the Tet-On inducible stable lines. The rtTA gene is under the control of a constitutive active promoter CAG in which the HCMV-MIEP coupled with the chicken β-actin promoter and the rabbit β-globin gene (CAGG) is linked to the polyoma virus mutant enhancer PyF101 (Liew et al. 2012). In the presence of DOX, the rtTA protein binds to the Tet operon promoter (TetO) and drives the expression of the human HNF4α2 or HNF4α8 gene from a core CMV promoter. A, right, shows IFs of HNF4α expression in the α2 and α8 lines, 24 h after DOX induction (10 x Magnification). (B) Schematic of the *HNF4A* gene and the P1 and P2 promoters that drive the expression of the P1- and P2-HNF4α isoforms. These isoforms differ at the N-terminal region of the protein. Right, IB analysis of HNF4α using the P1/P2 antibody from 20 μg of WCE. Cell lines (PL, HNF4α2, HNF4α8) were treated for 24 h with different amounts of DOX (0, 0.1, 0.3, 1, 5 μg/mL) as indicated. mLIV, WCE of mouse liver.

Fig. 4.2

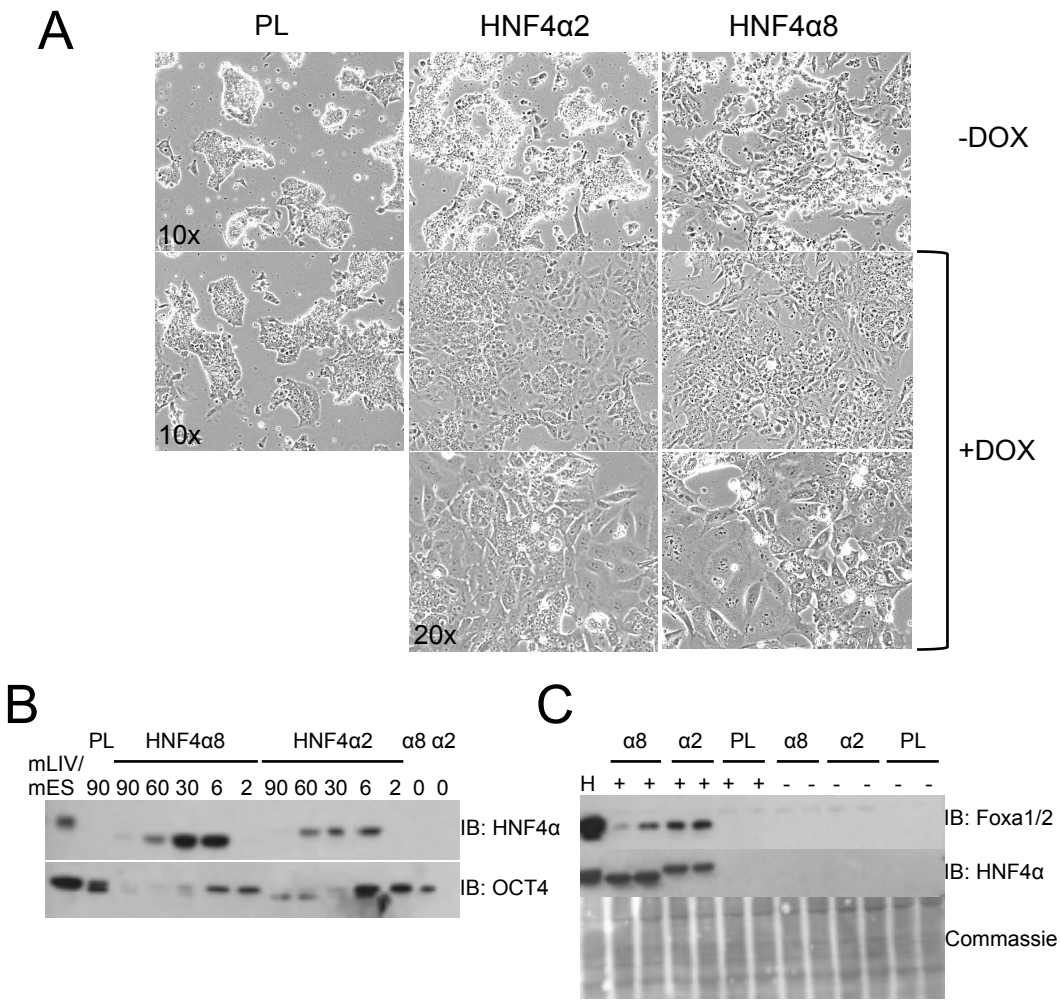
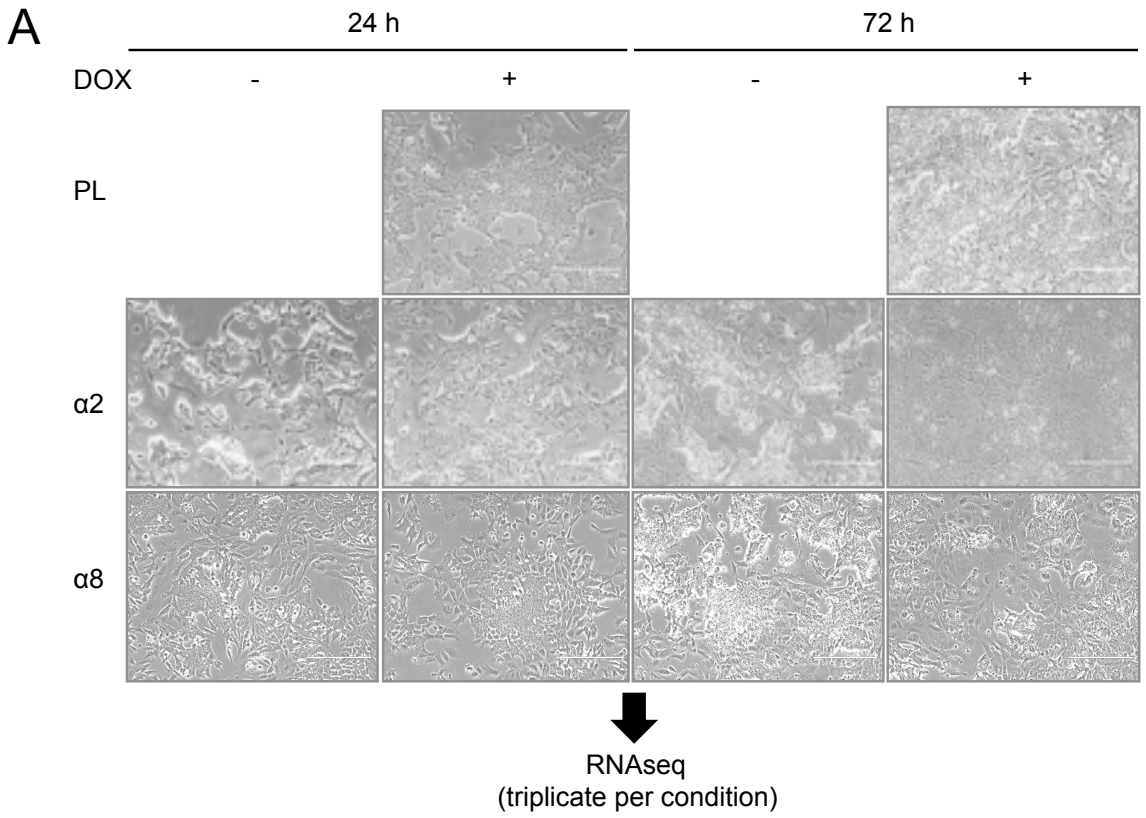

A
PL
HNF4α2
HNF4α8
-DOX
+DOX
10x
10x
20x
B
PL
HNF4α8
HNF4α2
α8 α2
mLIV/
mES 90 90 60 30 6 2 90 60 30 6 2 0 0
IB: HNF4α
IB: OCT4
C
α8 α2 PL α8 α2 PL
H + + + + + - - - - - -
IB: Foxa1/2
IB: HNF4α
Commassie

Figure 4.2 Ectopic expression of HNF4α drives mESC differentiation. (A) Phase contrast images of mESC lines treated with or without DOX (0.3 ug/mL) for 72 h at the indicated magnifications. IB analysis of HNF4α, OCT4, and Foxa1/2 (Hnf3a/b) from 20 μg of WCE (B) and NE (C). Cell lines (PL, α2, α8) were treated with DOX and harvested at different time points (0, 2, 6, 30, 60, 90 h) (B) or 72 h (C). Control: H, HepG2; mLIV, NE of mouse liver from WTα7.

Fig. 4.3


A
24 h
72 h
DOX
-
+
-
+
PL
α2
α8
RNAseq
(triplicate per condition)

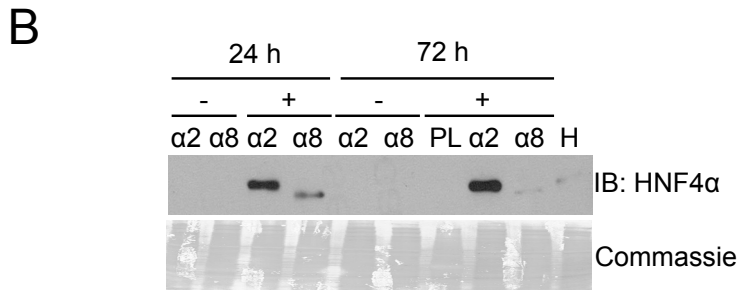

B
24 h
72 h
-
+
-
+
α2 α8 α2 α8 α2 α8 PL α2 α8 H
IB: HNF4α
Commassie

Figure 4.3 Global profile of differential expressions at 72 h using RNA-Seq. (A) Schematic set-up of the RNA-Seq analysis. Cells were induced with 0.3 µg/mL of DOX for 24 and 72 h and harvested at those time points. Phase contrast images of cells at 24 and 72 h (Mag. 10 x) (B) IB analysis of HNF4α from 20 µg of NE to verify isoform and expression. Control lanes: H, HepG2 (NE); PL, parental line (WCE).

Dr. Joseph M. Dhahbi is analyzing the mES inducible RNA-Seq data.

References

Briancon N, Bailly A, Clotman F, Jacquemin P, Lemaigre FP, Weiss MC. 2004. Expression of the alpha7 isoform of hepatocyte nuclear factor (HNF) 4 is activated by HNF6/OC-2 and HNF1 and repressed by HNF4alpha1 in the liver. *The Journal of biological chemistry* **279**: 33398-33408.

Briancon N, Weiss MC. 2006. In vivo role of the HNF4alpha AF-1 activation domain revealed by exon swapping. *The EMBO journal* **25**: 1253-1262.

Chartier FL, Bossu JP, Laudet V, Fruchart JC, Laine B. 1994. Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. *Gene* **147**: 269-272.

Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, Bachvarova RF, Darnell JE, Jr. 1994. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes & development* **8**: 2466-2477.

Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Molecular cell* **9**: 279-289.

Deutsch G, Jung J, Zheng M, Lora J, Zaret KS. 2001. A bipotential precursor population for pancreas and liver within the embryonic endoderm. *Development* **128**: 871-881.

Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, Darnell JE, Jr. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 7598-7602.

Duncan SA, Nagy A, Chan W. 1997. Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of Hnf-4(-/-) embryos. *Development* **124**: 279-287.

Eeckhoute J, Formstecher P, Laine B. 2001. Maturity-onset diabetes of the young Type 1 (MODY1)-associated mutations R154X and E276Q in hepatocyte nuclear factor 4alpha (HNF4alpha) gene impair recruitment of p300, a key transcriptional co-activator. *Molecular endocrinology* **15**: 1200-1210.

Furuta H, Iwasaki N, Oda N, Hinokio Y, Horikawa Y, Yamagata K, Yano N, Sugahiro J, Ogata M, Ohgawara H et al. 1997. Organization and partial sequence of the hepatocyte nuclear factor-4 alpha/MODY1 gene and identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* **46**: 1652-1657.

Garrison WD, Battle MA, Yang C, Kaestner KH, Sladek FM, Duncan SA. 2006. Hepatocyte nuclear factor 4alpha is essential for embryonic development of the mouse colon. *Gastroenterology* **130**: 1207-1220.

Green VJ, Kokkotou E, Ladias JA. 1998. Critical structural elements and multitarget protein interactions of the transcriptional activator AF-1 of hepatocyte nuclear factor 4. *The Journal of biological chemistry* **273**: 29950-29957.

Gupta RK, Vatamaniuk MZ, Lee CS, Flaschen RC, Fulmer JT, Matschinsky FM, Duncan SA, Kaestner KH. 2005. The MODY1 gene HNF-4alpha regulates selected genes involved in insulin secretion. *The Journal of clinical investigation* **115**: 1006-1015.

Hansen SK, Parrizas M, Jensen ML, Pruhova S, Ek J, Boj SF, Johansen A, Maestro MA, Rivera F, Eiberg H et al. 2002. Genetic evidence that HNF-1alpha-dependent transcriptional control of HNF-4alpha is essential for human pancreatic beta cell function. *The Journal of clinical investigation* **110**: 827-833.

Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, Steele A, Njolstad PR, Ellard S, Hattersley AT. 2008. The diabetic phenotype in HNF4A mutation carriers is moderated by the expression of HNF4A isoforms from the P1 promoter during fetal development. *Diabetes* **57**: 1745-1752.

Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ. 2001. Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Molecular and cellular biology* **21**: 1393-1403.

Huang J, Levitsky LL, Rhoads DB. 2009. Novel P2 promoter-derived HNF4alpha isoforms with different N-terminus generated by alternate exon insertion. *Experimental cell research* **315**: 1200-1211.

Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L. 2011. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **475**: 386-389.

Inoue Y, Peters LL, Yim SH, Inoue J, Gonzalez FJ. 2006a. Role of hepatocyte nuclear factor 4alpha in control of blood coagulation factor gene expression. *Journal of molecular medicine* **84**: 334-344.

Inoue Y, Yu AM, Yim SH, Ma X, Krausz KW, Inoue J, Xiang CC, Brownstein MJ, Eggertsen G, Bjorkhem I et al. 2006b. Regulation of bile acid biosynthesis by hepatocyte nuclear factor 4alpha. *Journal of lipid research* **47**: 215-227.

Kaestner KH. 2000. The hepatocyte nuclear factor 3 (HNF3 or FOXA) family in metabolism. *Trends in endocrinology and metabolism: TEM* **11**: 281-285.

Keller G. 2005. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes & development* **19**: 1129-1155.

Kir S, Zhang Y, Gerard RD, Kliewer SA, Mangelsdorf DJ. 2012. Nuclear receptors HNF4alpha and LRH-1 cooperate in regulating Cyp7a1 in vivo. *The Journal of biological chemistry* **287**: 41334-41341.

Liew CG, Draper JS, Walsh J, Moore H, Andrews PW. 2007. Transient and stable transgene expression in human embryonic stem cells. *Stem cells* **25**: 1521-1528.

Liu T, Zhang S, Xiang D, Wang Y. 2013. Induction of hepatocyte-like cells from mouse embryonic stem cells by lentivirus-mediated constitutive expression of Foxa2/Hnf4a. *Journal of cellular biochemistry* **114**: 2531-2541.

Maestro MA, Cardalda C, Boj SF, Luco RF, Servitja JM, Ferrer J. 2007. Distinct roles of HNF1beta, HNF1alpha, and HNF4alpha in regulating pancreas development, beta-cell function and growth. *Endocrine development* **12**: 33-45.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 7634-7638.

Montserrat N, Nivet E, Sancho-Martinez I, Hishida T, Kumar S, Miquel L, Cortina C, Hishida Y, Xia Y, Esteban CR et al. 2013. Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell stem cell* **13**: 341-350.

Rhee J, Inoue Y, Yoon JC, Puigserver P, Fan M, Gonzalez FJ, Spiegelman BM. 2003. Regulation of hepatic fasting response by PPARgamma coactivator-1alpha (PGC-1): requirement for hepatocyte nuclear factor 4alpha in gluconeogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 4012-4017.

Sekiya S, Suzuki A. 2011. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475**: 390-393.

Serandour AA, Avner S, Percevault F, Demay F, Bizot M, Lucchetti-Miganeh C, Barloy-Hubler F, Brown M, Lupien M, Metivier R et al. 2011. Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome research* **21**: 555-565.

Sladek FM, and Seidel, S.D. 2001. In Nuclear Receptors and Genetic Diseases. *TP Burris and ERB McCabe, eds (London: Academic Press)*: pp.309-361.

Sladek FM, Ruse MD, Jr., Nepomuceno L, Huang SM, Stallcup MR. 1999. Modulation of transcriptional activation and coactivator interaction by a splicing variation in the F domain of nuclear receptor hepatocyte nuclear factor 4alpha1. *Molecular and cellular biology* **19**: 6509-6522.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Takayama K, Inamura M, Kawabata K, Katayama K, Higuchi M, Tashiro K, Nonaka A, Sakurai F, Hayakawa T, Furue MK et al. 2012. Efficient generation of functional hepatocytes from human embryonic stem cells and induced pluripotent stem cells by HNF4alpha transduction. *Molecular therapy : the journal of the American Society of Gene Therapy* **20**: 127-137.

Taraviras S, Monaghan AP, Schutz G, Kelsey G. 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mechanisms of development* **48**: 67-79.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. 1998. Embryonic stem cell lines derived from human blastocysts. *Science* **282**: 1145-1147.

Torres-Padilla ME, Fougere-Deschatrette C, Weiss MC. 2001. Expression of HNF4alpha isoforms in mouse liver development is regulated by sequential promoter usage and constitutive 3' end splicing. *Mechanisms of development* **109**: 183-193.

Torres-Padilla ME, Sladek FM, Weiss MC. 2002. Developmentally regulated N-terminal variants of the nuclear receptor hepatocyte nuclear factor 4alpha mediate multiple interactions through coactivator and corepressor-histone deacetylase complexes. *The Journal of biological chemistry* **277**: 44677-44687.

Torres-Padilla ME, Weiss MC. 2003. Effects of interactions of hepatocyte nuclear factor 4alpha isoforms with coactivators and corepressors are promoter-specific. *FEBS letters* **539**: 19-23.

Watt AJ, Garrison WD, Duncan SA. 2003. HNF4: a central regulator of hepatocyte differentiation and function. *Hepatology* **37**: 1249-1253.

Watt AJ, Zhao R, Li J, Duncan SA. 2007. Development of the mammalian liver and ventral pancreas is dependent on GATA4. *BMC developmental biology* **7**: 37.

Xia X, Ayala M, Thiede BR, Zhang SC. 2008. In vitro- and in vivo-induced transgene expression in human embryonic stem cells and derivatives. *Stem cells* **26**: 525-533.

Yamamizu K, Piao Y, Sharov AA, Zsiros V, Yu H, Nakazawa K, Schlessinger D, Ko MS. 2013. Identification of Transcription Factors for Lineage-Specific ESC Differentiation. *Stem cell reports* **1**: 545-559.

Zaret K. 1999. Developmental competence of the gut endoderm: genetic potentiation by GATA and HNF3/fork head proteins. *Developmental biology* **209**: 1-10.

Supplemental Materials

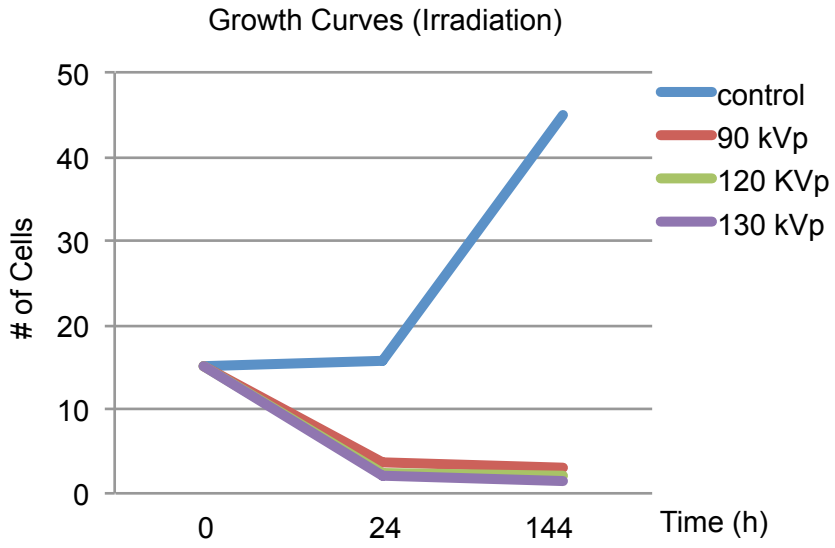
Irradiating MEFs to produce feeders for co-culture with mES cells

To determine the optimal setting on the irradiation machine, we performed growth curve assay. We tested three different kilovolt powers (90, 120, 130 kVp), each condition for 60 minutes. Compared to the controlled cells without treatment, all three treatments caused cellular senescence (Fig. 4.S1).

Immunofluorescence (IF) for HNF4α2 and HNF4α8 in mES and HCT116 inducible cell lines are similar

IF of HNF4α2 in mES cells showed some bright, large patchy staining whereas the IF of HNF4α8 looked dimmer although this was not observed before (Fig. 4.S2A and Fig. 2.6C, Chapter 2); similar staining patterns were also observed in the HCT116 inducible cells, however these staining were done only once (Fig. 4.S2B).

Fig. 4.S1

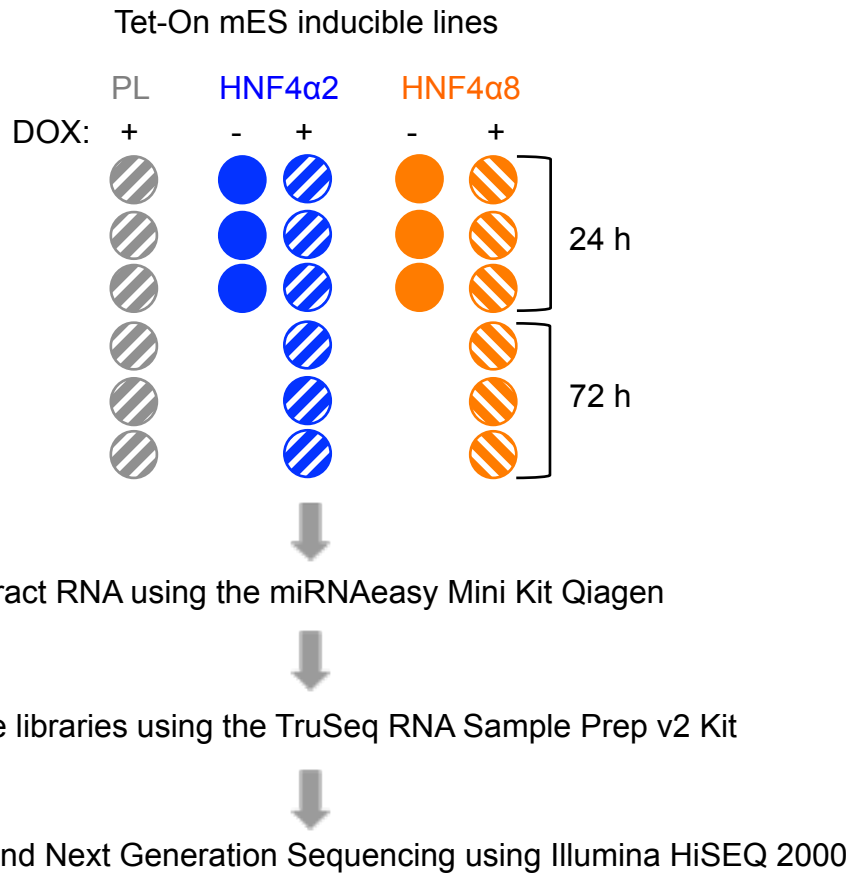


Supplemental Figure 4.1 MEF irradiation growth curves.

Cells were seeded at 1.102 x 10^6 cells per 100-mm plate and cultured for one day. At 2 x 10^6 confluency, cells were transferred to 15 mL conical tube and placed in the Faxitron machine to be irradiated at the following peak kilovolts for 60 min: 90, 102, or 130 kVp. After irradiation, cells were seeded into 1.5 x 10^5 cells per 6-well plate and cell counting using a hemocytometer was done 24 and 144 hours after plating. Control, cells were not irradiated.

Faxitron Machine information: Model: 438550; kVp: peak kilovoltage (max voltage applied on X-tray tube); Focal spot size: 1.5 mm; mA: 5; Inherent filtration: 1.6 mm Be; X-tray beam divergence: 40 deg.; Table: 15 inch.

Fig. 4.S2



Supplemental Figure 4.2 Schematic of the RNA-Seq experimental design.

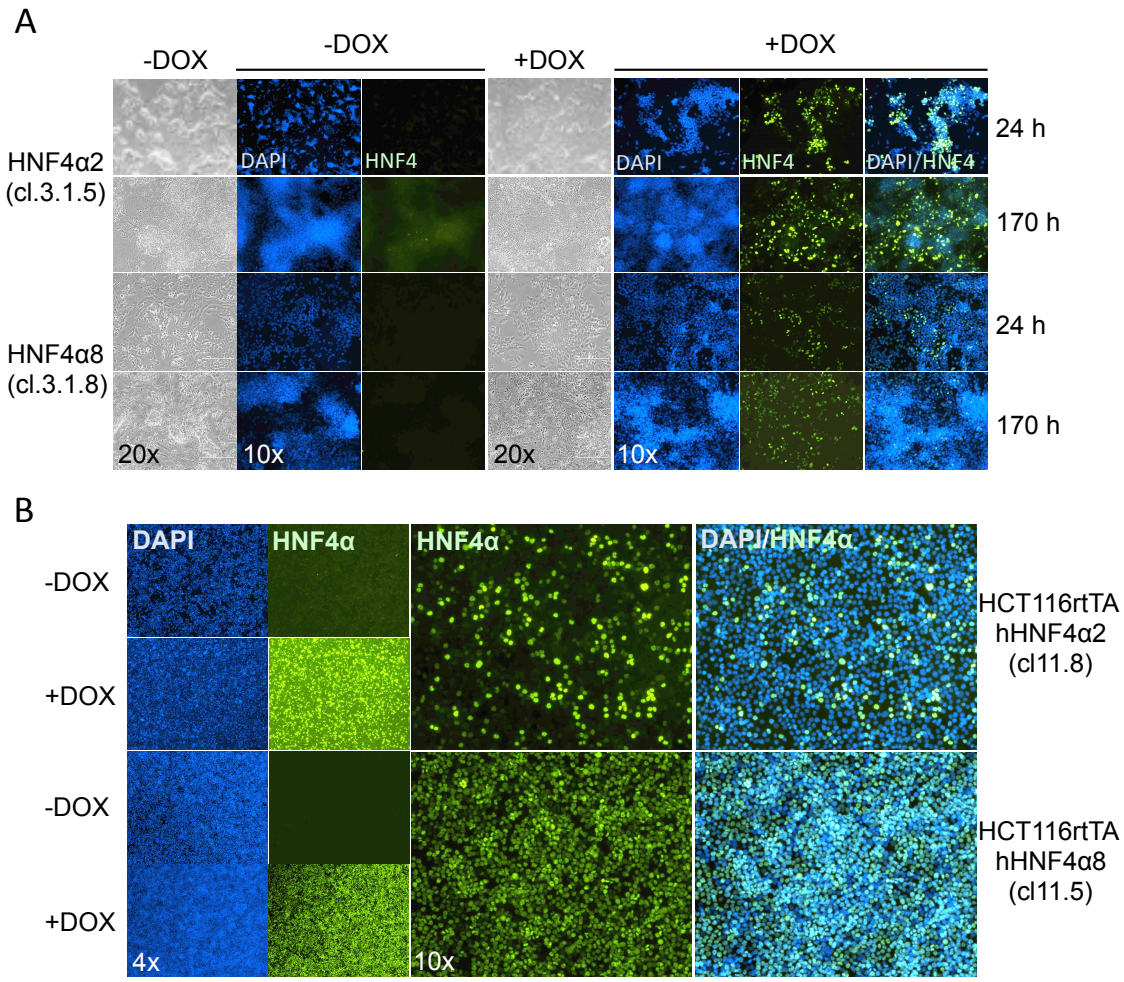
Cells seeded at 1.5 x 10^5 cells per gelatin-coated 12-well and (0.3 μg/mL) DOX treated 24 h later. RNA extracts were collected at 24 and 72 h after induction. RNA extracts were passed through an RNA column to collect total RNA. Poly A+ RNAs were selected and libraries were generated. The libraries were sent for 50 bp single end Next Generation Sequencing. Table 1 shows the number of reads from the sequencing result. See Figure 3.3B for IB analysis of HNF4α from these cells.

Table 4.S1 RNA sequencing reads (mES inducible lines)

# of Lanes	Condition	# of Replicates	Index	# of Seq reads
Lane 1	HNF4α2 (72 h, +DOX)	Rep_1	CCGTCC	9,690,763
		*Rep_2	GTCCGC	3,306,091
		Rep_3	GTGAAA	11,731,523
	Parental (24 h, +DOX)	Rep_1	CGATGT	20,455,987
		Rep_2	TGACCA	17,386,042
		Rep_3	ACAGTG	10,238,721
	HNF4α2 (24 h, -DOX)	Rep_1	GCCAAT	12,800,600
		Rep_2	CAGATC	27,657,157
		Rep_3	CTTGTA	20,738,307
	HNF4α2 (24 h, +DOX)	Rep_1	AGTCAA	15,075,179
		Rep_2	AGTTCC	19,299,341
		Rep_3	ATGTCA	19,379,625
Lane 2	HNF4α8 (72 h, +DOX)	Rep_1	CCGTCC	18,386,384
		Rep_2	GTCCGC	17,960,969
		Rep_3	GTGAAA	11,065,001
	Parental (72 h, +DOX)	Rep_1	CGATGT	11,123,888
		Rep_2	TGACCA	15,878,927
		Rep_3	ACAGTG	14,341,458
	HNF4α8 (24 h, -DOX)	Rep_1	GCCAAT	12,474,590
		Rep_2	CAGATC	11,354,915
		Rep_3	CTTGTA	23,674,515
	HNF4α8 (24 h, +DOX)	Rep_1	AGTCAA	9,321,626
		Rep_2	AGTTCC	24,595,856
		Rep_3	ATGTCA	26,675,724

*Note: the reads for one of the replicates in the α2 line that was treated with DOX for 72 hours is about 5 fold less than the reads for the other samples.

Fig. 4.S3


A
-DOX
-DOX
+DOX
+DOX
HNF4α2
(cl.3.1.5)
HNF4α8
(cl.3.1.8)
DAPI
HNF4
DAPI
HNF4
DAPI/HNF4
24 h
170 h
24 h
170 h
20x
10x
20x
10x
B
DAPI
HNF4α
HNF4α
DAPI/HNF4α
-DOX
+DOX
-DOX
+DOX
HCT116rtTA
hHNF4α2
(cl11.8)
HCT116rtTA
hHNF4α8
(cl11.5)
4x
10x

Supplemental Figure 4.3 Fluorescent staining of ectopic HNF4α isoforms in mES and HCT116 inducible cell lines.

(A) mES inducible cell lines were seeded at 1.5 x 10^5 cells per well of 0.1% gelatin-coated 12-well and (0.3 µg/mL) DOX treated 24 h later. Cells were fixed at 24 and 170 hours after induction with 3.7% formaldehyde (FA) in 1 x PBS for 15 min, followed by 1.25 M glycine incubation for 10 min, at 37^{o}C incubator. Two to three quick washes with PBS before applying blocking and blotting solutions. (B) HCT116 inducible cell lines were seeded at 3 x 10^5 cells per well of 12-well plate and (0.3 µg/mL) DOX treated 6-7 hours later. Cells were fixed 24 h after treatment with 3.7% FA in PBS for 15 min at RT and washed with PBS several times. Both lines were blocked (10% donkey serum, 10% goat serum, 0.2% Triton-X100, in PBS) for 30 min at RT and blot (1% each donkey and goat serums) in primary antibody (anti-P1/P2 1:1200) overnight at 4^{o}C and secondary antibody (Alexa-488 1:500; DAPI 1:1000) for ~1-2 h at RT.

Appendix to Chapter 4

This extended appendix includes a cell count assay on the mouse embryonic stem cells (mESCs) inducible lines and two embryoid body (EB) studies using mESCs. In the former, we further characterize the Tet-On inducible lines to determine the inhibitory activity of the HNF4α isoforms on proliferation in mESCs. We hypothesize that both HNF4α2 and HNF4α8 decrease cell number but that HNF4α2 slows cell growth more effectively than HNF4α8, as seen in the HCT116 cells. In the latter, we explore the development of the EB formation. We use EBs to recapitulate early embryogenesis and to identify the spatiotemporal expression of the promoter-driven HNF4α isoforms in early mouse development.

Materials and Methods

Plasmid Constructs

pcDNA4/TO.Flag.TCF3E was a gift from Dr. Marian Waterman at the University of California, Irvine. See Chapter 2 Materials and Methods (M&M) for details on expression vectors and establishment of inducible cell lines.

Cell culture

Media that was supplemented with 1% nonessential amino acids (NEAA), and as indicated with extra 1% sodium pyruvate and glutmax (GIBCO). Cells were maintained as described in the M&M of Chapter 4.

Cell count assay

Cells were seeded at 1.5 x 10^5 cells per well in a 12-well plate pre-coated with 0.1% gelatin and treated with 0.3 µg/mL of DOX every 24 h interval. A Coulter counter (Beckman) was used to count cells every 24 h after trypsinization.

Immunoblot (IB) assay

Protein extracts were separated as previously described (Chapter 4; M&M). Primary antibodies were mouse monoclonal anti-HNF4α P1/P2 (R&D, cat#PP-H1415-00) (1:60,000) that recognizes the C-terminus of HNF4α; monoclonal anti-HNF4α P1 and P2 (R&D, cat#PP-K9218-00 & cat#PP-H6939-00, respectively) that recognize the different N-terminal domains of the HNF4α isoforms (1:5000); anti-TCF7L2 (Millipore, cat#6H5-3) (1:1000); anti-TCF7 (Cell Signaling, C46C7) (1:1000); anti-TCF7L1 (Santa Cruz, sc166411) (1:1000); β-catenin (Abcam, ab32572) (1:5000); OCT4 (Santa Cruz, sc9081) (1:5000). Secondary antibodies were horseradish peroxidase (HRP)-conjugated goat anti-rabbit (GαR-HRP) or goat anti-mouse (GαM-HRP) from Jackson ImmunoResearch Laboratories (1:5000 to 1:20,000).

Immunofluorescence (IF) assay

mES and EB cells were fixed in 3.7-4% formaldehyde for 10 to 15 min at RT or in a 37^{o}C incubator, blocked with 10% goat and donkey serums (GS, DS) and 0.2% Triton X-100 in PBS at RT, and blotted with 1% GS and DS in PBS at 4^{o}C overnight. Primary antibodies include HNF4α P1/P2 (R&D) at 1:1200; P1-HNF4α and P2-HNF4α (R&D) at 1:500; OCT4 (Santa Cruz) at 1:200; β-catenin (Abcam) at 1:200; E-cadherin (Abcam,

ab53033) at 1:200. The secondary Alexa Fluor antibodies were used at 1:500 (Invitrogen); 4'6-diamidino-2-phenylindole (DAPI) was used at 1:1000 to visualize nuclei.

Fluorescent gel shift assay

Gel shift was performed as described in Chapter 3 M&M; the following proteins, in addition to Flag.dnTCF1 and human HNF4α2, were used in the gel shift: Flag.TCF3, Flag.TCF4, and human HNF4α8.

Embryoid body (EB) formation and IF sectioning

Formation of EBs was carried out using the hanging drop method (Höpfl et al. 2004; Kurosawa et al. 2007). Approximately 1000 cells per 27-30 µL drop were placed on the lid of a 100-mm plate. Lids with droplets were inverted on top of the bottom dishes, which were filled with 10 mL PBS. The plates were incubated at 37^{o}C for three days. On the third day, EBs were transferred to regulator bacterial, low adhesive plates in mES media without LIF for three to five days. EBs were collected and transferred to adhesive plates coated with 0.1% gelatin and incubated for another 24 h before fixing with 3.7% formaldehyde, sectioning, and staining for P1- and P2-HNF4α, β-catenin, OCT4, and E-cadherin as described above.

Results and Discussion

Ectopic expression of HNF4α2 (P1) and HNF4α8 (P2) isoforms decrease cell number in mES cells

HNF4α is known to inhibit cell proliferation (Lucas et al. 2005; Hwang-Verslues and Sladek 2008; Erdmann et al. 2007) in HEK293, HCT116, and INS-1 lines. Here we

show that ectopic expression of HNF4α not only drives differentiation but also reduces cell number in mESCs. We found that both HNF4α2 and HNF4α8 variants were able to decrease cell count but that HNF4α2 inhibited growth more efficiently and for a longer period (Fig. 4.1.1A,D; include shorter period cell count assays, Fig. 4.1.2A,B, right). Interestingly, the HNF4α protein expression also decreases with time (Fig.s 4.1.1B,E, 4.1.2A,B, left). This is seen in both stable and transient transfections of HNF4α in other cell lines as discussed in Chapter 2 and the Appendix to Chapter 3. Also shown in Chapter 3, HNF4α2 is a better tumor suppressor than HNF4α8 in HCT116 cells (Fig. 3.3C,E, Chapter 3), which is consistent with the longer sustained decrease by HNF4α2 but not HNF4α8 in mES cells. However, the precise mechanism that enables HNF4α2 to be an efficient growth inhibitor is not clear in the mESCs or in HCT116 cells.

We also tracked OCT4 expression and observed OCT4 expression decreased in the presence of both isoforms of HNF4α (Fig.s 4.1.1C,E and 4.1.2A,B, left). OCT4 is a member of the POU family of transcription factors that contain a POU DNA-binding domain, which recognizes the octamer motif ATGCAAAT (Herr and Cleary 1995; Sturm and Herr 1988; Wirth et al. 1987; Brehm et al. 1997; Parslow et al. 1984). OCT4 expression is important in maintaining pluripotency in mES cells; although, overexpression of the gene can lead to mesodermal differentiation (Niwa et al. 2000). In contrast, downregulation of the *Oct4* gene produces an epithelial-like morphology and induces expression of genes associated with endoderm differentiation such as *Gata6*, α-fetoprotein (*Afp*), and *Gtar,* which are found in fetal liver and yolk sac. (Hay et al. 2004). These results indicate that the forced expression of HNF4α in mESCs downregulates *Oct4* expression thus producing an irreversible epithelial morphology. Whether HNF4α directly regulates *Oct4* expression is not known, although other nuclear receptors such

as COUP transcription factor I (COUP-TFI), COUP-TFII (ARP-1), V-erbA-related protein 2 (EAR-2), and liver receptor homolog-1 (LRH-1) have been shown to negatively regulate *Oct4* activity (Slyvester and Schöler 1994; Schoolerlemmer et al. 1994; Gu et al. 2005). Sylvester and Schöler found several putative Sp1 binding elements overlapping with three AGGTCA direct repeats that have one or no spacing nucleotide (DR1 or DR0) in the OCT4 proximal promoter (Slyvester and Schöler 1994). A DR1 is the consensus motif for HNF4α and HNF4α has been shown to bind to Sp1-bound promoters of the cyclin-dependent kinase inhibitor 1 (CKI) $p21^{CIP1/WAF1}$ (Hwang-Verslues and Sladek 2008; Bolotin et al. 2010), suggesting that HNF4α could potentially bind, and repress, the OCT4 promoter.

mESCs express TCF3 (TCF7L1) and TCF1 (TCF1) but not TCF4 (TCF7L2)

Initially, we wanted to investigate the interplay between HNF4α and LEF/TCFs in mES cells. Since TCF3 (TCF7L1) is expressed in ES cells (Fig. 4.1.3) and plays an integral role in balancing the core regulatory circuitry of ES cells, we focused our attention on TCF3 (Pereira et al. 2006; Lluis et al. 2011; Cole et al. 2008). TCF3 has been shown to repress transcription of Nanog gene to negatively affect the rapid self-renewal of ES cells (Yi et al. 2008; Pereira et al. 2006; Cole et al. 2008; Wray et al. 2011). A region between -3 to -5 kb of the Nanog promoter contains five TCF binding motifs (A/T)(A/T)CAAAG and exhibits TCF3 repressive activity, independent of β-catenin (Yi et al. 2008). In Chapter 3, we used PBMs to identify several overlapping binding sequences that are shared by HNF4α and LEF/TCFs. We confirmed the binding of HNF4α and TCFs to one of the overlapping sequence elements using a gel shift assay (Fig. 3.6A,B, Chapter 3). We showed that three different TCFs and two HNF4α isoforms

bind the HNF4α strong and TCF strong (HstrgTstrg) element in an independent fashion. All LEF/TCFs recognize the same binding motif (van de Wetering et al. 1991, 1993; van Beest et al. 2000) (Figure 4.1.4). We speculate that in mES cells HNF4α and TCF3 work in concert to control self-renewal and drive the cells toward a differentiation state.

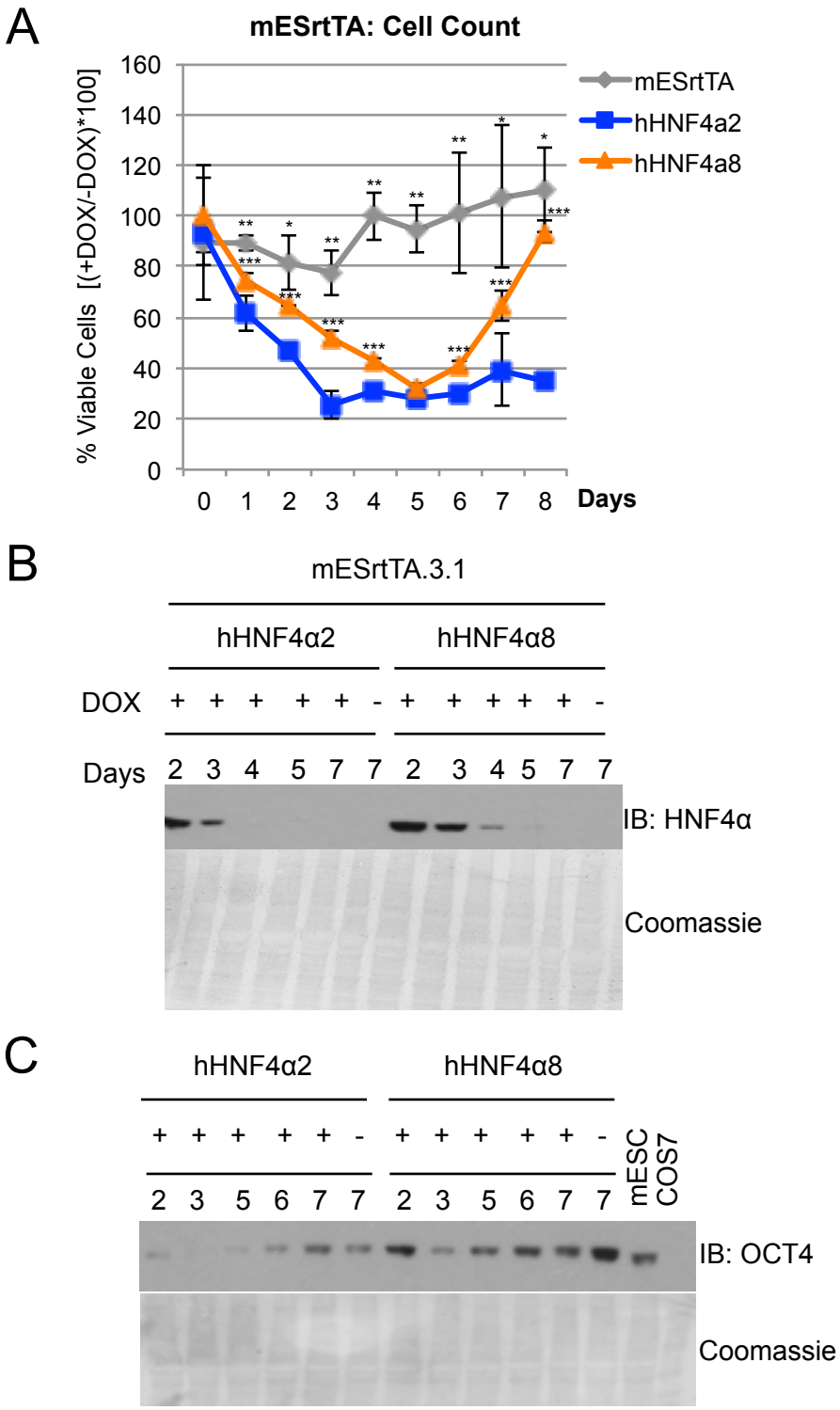
The spatiotemporal expression of the promoter-driven HNF4α isoforms in EB development

Mouse ES cells are derived from the inner cell mass (ICM) of blastocysts at embryonic day E4 of development (Evans and Kaufman 1981; Martin 1981) (Fig. 4.1.5A, left). In vitro, ES cells can be co-cultured with mouse embryonic fibroblasts (MEFs) or on 0.1% gelatin-coated plates in the presence of leukemia inhibitory factor (LIF) to maintain pluripotency (Williams et al. 1988) (Fig. 4.1.5A, center and right). Both culturing systems are sufficient to maintain the expression of the pluripotency marker, OCT4 (Schöler et al. 1989) (Fig. 4.1.5A, below). ES cells remain undifferentiated in the presence of LIF and spontaneously differentiate into three germ layers in the absence of LIF (Thomson et al. 2011; Smith et al. 1988; Mummery et al. 1990;). Spontaneous differentiation can also be achieved by growing the cells in suspension in bacterial dishes or as hanging drops without LIF (Desbaillets et al. 2000; Doetschman et al. 1985). In both cases, ES cells aggregate into spherical shapes of various sizes called embryoid bodies (EBs) that resemble mouse embryonic development (Doetschman et al. 1985, 1988) (Fig. 4.1.5B). After several days in suspension or in hanging drops, EBs can be grown on adhesive plates where they continue to expand and spread out. Cells on the periphery of the EB spread away from the densely centered mass; these cells were found to express nuclear HNF4α and membrane-bound E-cadherin (Fig. 4.1.5C).

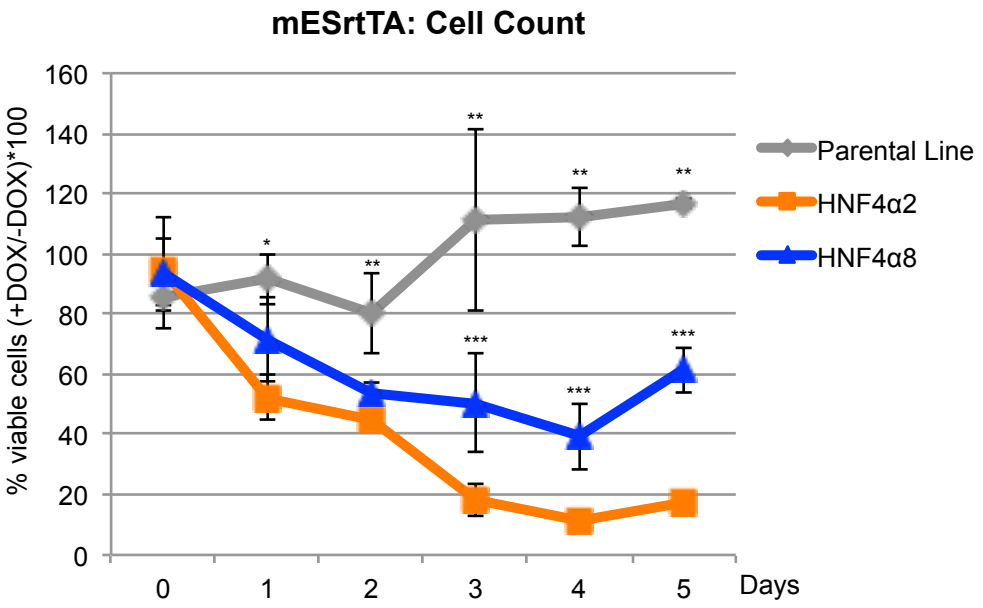
HNF4α expression is detected as early as embryonic day E4.5 of mouse development and restricted to the visceral endoderm cells of the yolk sac from E5.5 to E8.5 (Duncan et al. 1994). At E8.5, HNF4α is expressed in the embryonic tissues of the liver diverticulum and hindgut (Duncan et al. 1994). However, little is known about the expression of the promoter-driven HNF4α isoforms in development. We know, thus far, that the P2-HNF4α is present at the later stages of liver development, starting E12, and is replaced with the P1-HNF4α isoforms in the adult liver (Torres-Padilla et al. 2001; Briancon et al. 2004). Since the EBs resemble early embryogenesis, we tracked the expression of the P1- and P2-HNF4α up to day 11 using different time courses: with varying number of days in suspension or attachment (Fig.s 4.1.6A, 4.1.7A). Regardless of which method was used, we found that HNF4α was detected at day 3 after the hanging drop and continued to be expressed up to day 11 (Fig.s 4.1.6B, 4.1.7B,D). The P1-HNF4α expression seemed to come up first at around day 5, while the P2-HNF4α expression seemed to come up much later (Fig.s 4.1.6C, 4.1.7C,D) but the antibody used to detect just the P2-HNF4α consistently gives a low signal. Although we were able to detect HNF4α expression in cells on the outer perimeter of the EB, we are not the first to do so (Sajini et al. 2012; Hamazaki et al. 2004; Morrisey et al. 1998; Grover et al. 1983). In addition, we also monitored TCF3 expression and found its expression was decreased by Day 6 and gone by Day 7 just as HNF4α expression was increasing, while β-catenin expression remained unchanged (Fig. 4.1.7D). No OCT4 staining was found in cells surrounding the inner cell mass at 3 days after the hanging drop (Fig. 4.1.6B, right). Since TCF3 and HNF4α proteins were not present concurrently to any appreciable degree, deciphering the role of the two factors in EB development would not be feasible (Fig. 4.1.7D).

Embryoid bodies are very heterogeneous. It is difficult to discern any organization inside the EB. Although one group, Berge et al showed that local Wnt signaling establishes polarization and self-organization (ten Berge et al. 2008). Nevertheless, EBs may not be the perfect model to trace the expression of the HNF4α isoforms. Similar to a mouse blastocyst, EB has two layers; the cells in the outer layer which become the primitive endoderm and further differentiate into the visceral endoderm, and the cells in the inner mass which differentiate into three germ lines (Doetschman et al. 1985). However, as shown in figures 4.1.6 and 4.1.7, the cells in the inner mass are quite dense, making it impossible to distinguish individual cells and hence to discern expression of specific HNF4α isoforms. One ideal method to trace the isoforms is to recapitulate Duncan 1994 experiment, which uses in situ hybridization to track HNF4α expression in different tissues during development; we could hybridize with primers that recognize the different P1 and P2 isoforms. Another method is to stain blastocysts with the P1 antibody (P2 do not work well).

Fig. 4.1.1


A
mESrtTA: Cell Count
% Viable Cells [(+DOX/-DOX)*100]
160
140
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8
Days
mESrtTA
hHNF4a2
hHNF4a8
B
mESrtTA.3.1
hHNF4α2
hHNF4α8
DOX + + + + + - + + + + + -
Days 2 3 4 5 7 7 2 3 4 5 7 7
IB: HNF4α
Coomassie
C
hHNF4α2
hHNF4α8
+ + + + + - + + + + + -
2 3 5 6 7 7 2 3 5 6 7 7
mESC
COS7
IB: OCT4
Coomassie

D


mESrtTA: Cell Count
% viable cells (+DOX/-DOX)*100
160
140
120
100
80
60
40
20
0
0
1
2
3
4
5
Days
Parental Line
HNF4α2
HNF4α8

E

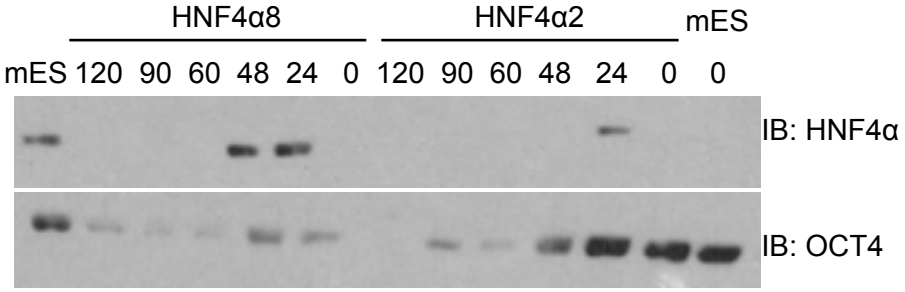

HNF4α8
HNF4α2
mES
mES 120 90 60 48 24 0 120 90 60 48 24 0 0
IB: HNF4α
IB: OCT4

Figure 4.1.1 Ectopic expression of HNF4α isoforms decrease cell count in mESCs. Cells were seeded at 1.5 x 10^5 cells per well of a 12-well plate coated with 0.1% gelatin and, 24 h later, treated with 0.3 µg/mL DOX in condition media (CM) without LIF and media were changed every 24 h interval with the same DOX concentration. Each day, cells were trypinized and counted for up to 5 or 8 days using a Coulter counter (Beckman). (A and D) Shown are graphs representing the mean±SD of percentage of viable cells in triplicates, each graph from one independent experiment. *P-value* ≤ 0.05 from T-test comparing * PL vs. α2, ** PL vs α2 & α8, *** α2 vs α8. (B, C, E) IB analysis of WCE using the P1/P2 antibody that recognizes the C-terminal end of HNF4α and OCT4 antibody of cells from one experiment in (A and D). Coomassie staining (B and C) shows equal loading of 20 µg total protein across samples.

Fig. 4.1.2

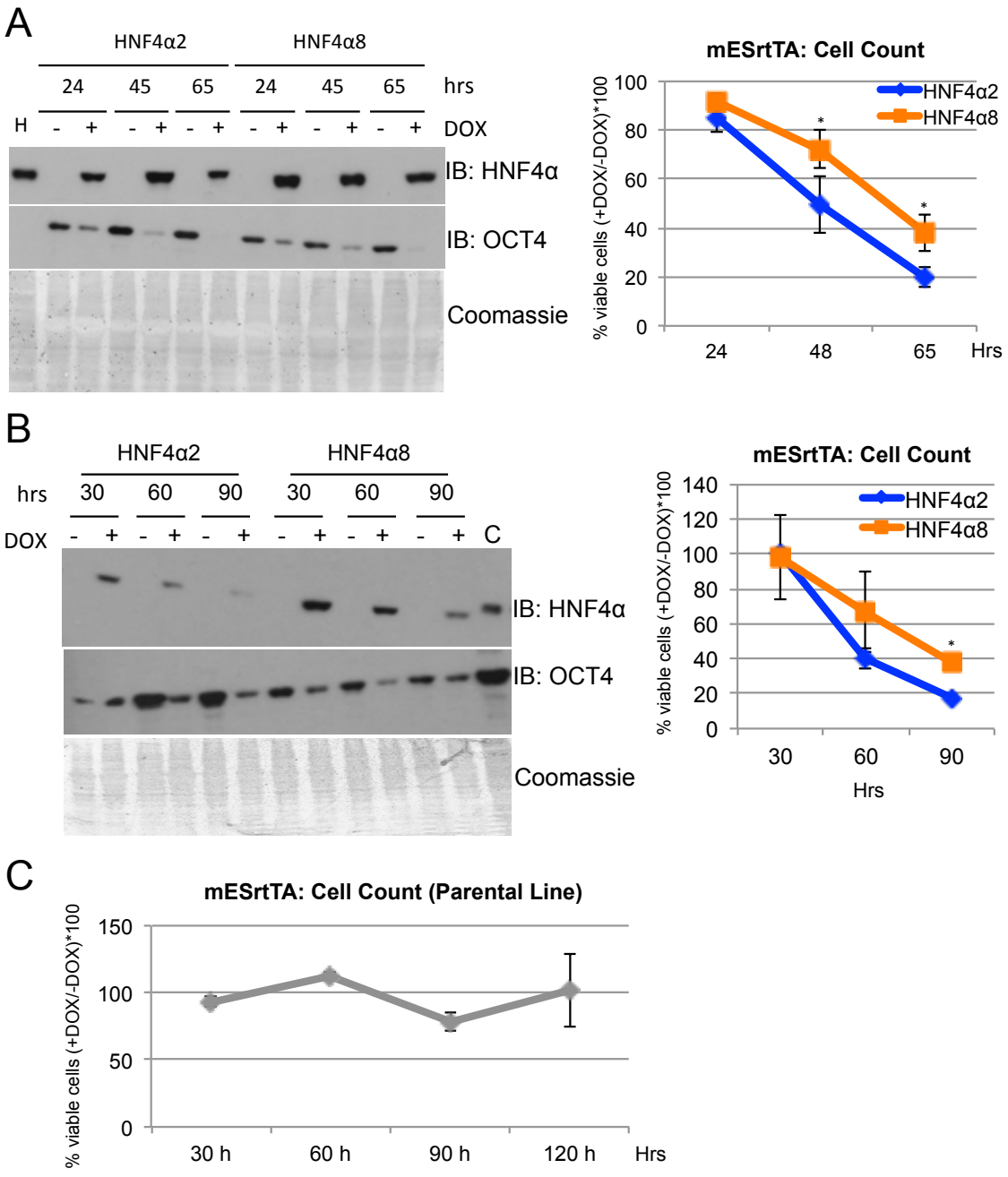

A
HNF4α2
HNF4α8
24 45 65 24 45 65 hrs
H - + - + - + - + - + - + DOX
IB: HNF4α
IB: OCT4
Coomassie
mESrtTA: Cell Count
% viable cells (+DOX/-DOX)*100
HNF4α2
HNF4α8
24 48 65 Hrs
B
HNF4α2
HNF4α8
hrs 30 60 90 30 60 90
DOX - + - + - + - + - + - + C
IB: HNF4α
IB: OCT4
Coomassie
mESrtTA: Cell Count
% viable cells (+DOX/-DOX)*100
HNF4α2
HNF4α8
30 60 90 Hrs
C
mESrtTA: Cell Count (Parental Line)
% viable cells (+DOX/-DOX)*100
30 h 60 h 90 h 120 h Hrs

Figure 4.1.2 OCT4 expression decreases over time in the presence of HNF4α.

(A and B) IB for HNF4α and OCT4 expression in WCE from the indicated inducible mES lines. Equal loading of 20 μg of total protein was monitored using Coomassie staining (parallel probes, A and B). (A and B left, and C) Include cell count assays that the extracts came from. Cell count assay on Parental line (PL) was done separately. Shown are graphs representing mean±SD percentage of viable cells in triplicates; *P-value* ≤ 0.05 from T-Test comparing * α2 vs α8. Controls: C, NE of mouse liver or mESCs; H, HepG2.

Fig. 4.1.3

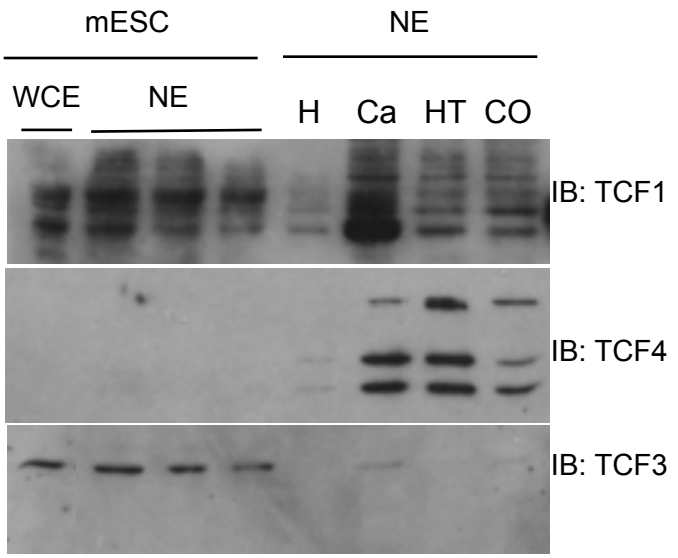


Figure 4.1.3 Embryonic stem cells express TCF3.

IB of WCE and NE (20 µg total protein per lane) of mES cells for TCF3, TCF4, and TCF1. H, HepG2; Ca, CaCo2; HT, HCT116; and CO, COS-7 cells.

Fig. 4.1.4

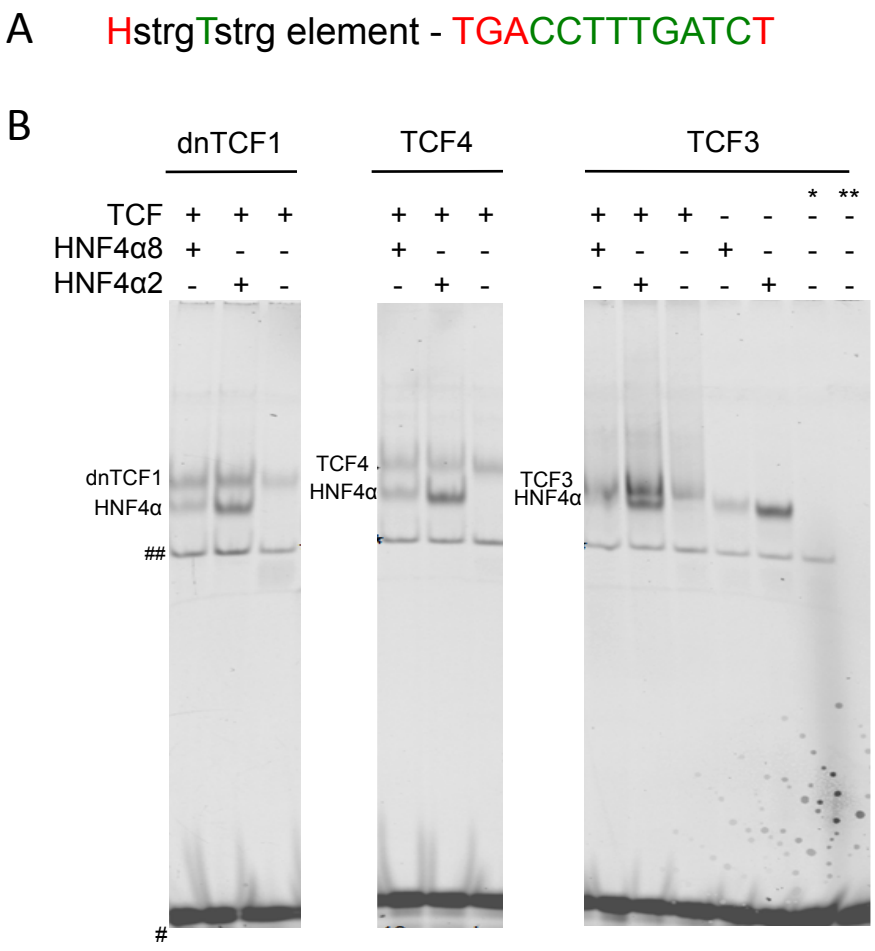
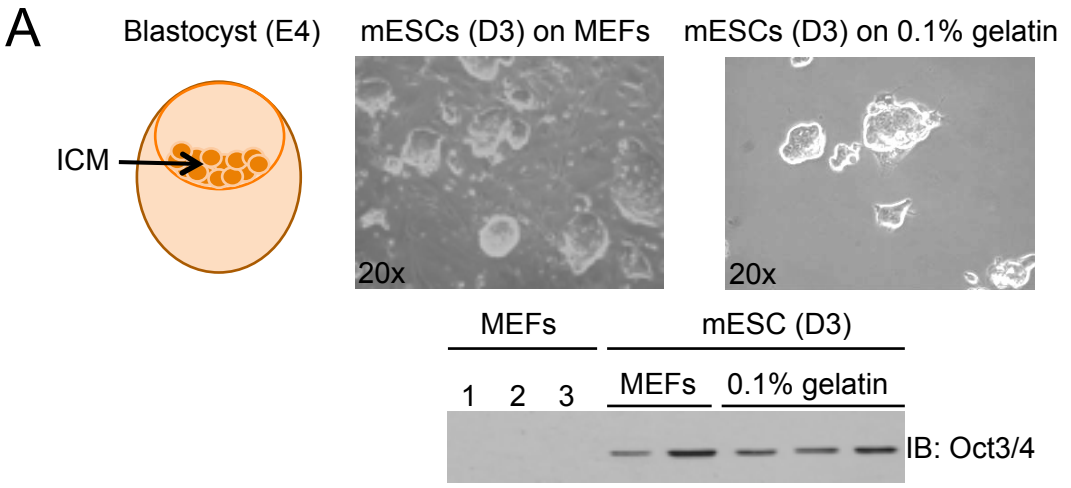
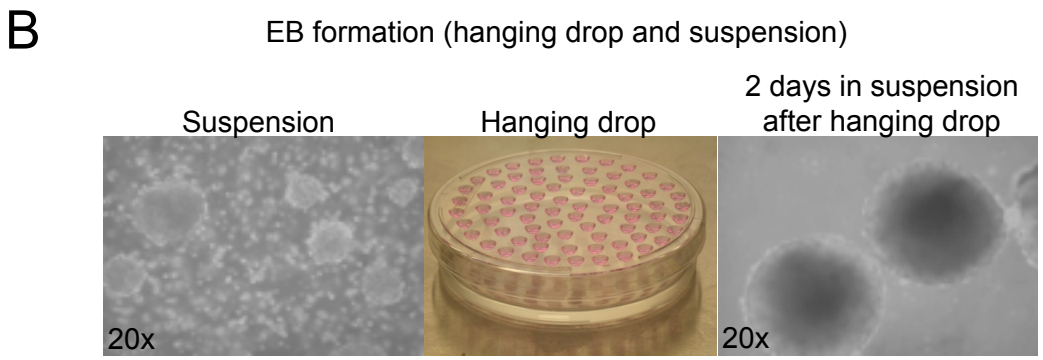

A
HstrgTstrg element - TGACCTTTGATCT
B
dnTCF1
TCF4
TCF3
TCF + + + + + + + + + - - * **
HNF4α8 + - - + - - + - - + - - -
HNF4α2 - + - - + - - + - - + - -
dnTCF1
HNF4α
##
#
TCF4
HNF4α
TCF3
HNF4α

Figure 4.1.4 HNF4α2, HNF4α8, dnTCF1, TCF3, and TCF4 bind the HNF4α strong and TCF strong (HstrgTstrg) element.

(A) Sequence of HNF4α strong and TCF strong (HstrgTstrg) element identified in PBMs described in Chapter 3 (Fig. 3.6A, left). (B) Gel shift array showing binding of HNF4α (P1-HNF4α2 and P2-HNF4α8) isoforms and TCFs (TCF3, TCF4, dnTCF1) to the HstrgTstrg element. Reaction mix includes 6 µL of protein mix (3 µL of NE and 3 µL of BSA 1mg/mL), 10 ng of probe, 1.5 µg of sonicated sperm salmon DNA in a shift reaction buffer (see Chapter 3 M&M). Reactions were not normalized to total protein. * NE of mock transfection (pcDNA3.1) plus BSA; ** Probe (plus salmon sperm DNA) only; # free probe; ## non-specific band.

Fig. 4.1.5


A
Blastocyst (E4)
mESCs (D3) on MEFs
mESCs (D3) on 0.1% gelatin
ICM
20x
20x
MEFs
mESC (D3)
1 2 3
MEFs
0.1% gelatin
IB: Oct3/4


B
EB formation (hanging drop and suspension)
Suspension
Hanging drop
2 days in suspension after hanging drop
20x
20x

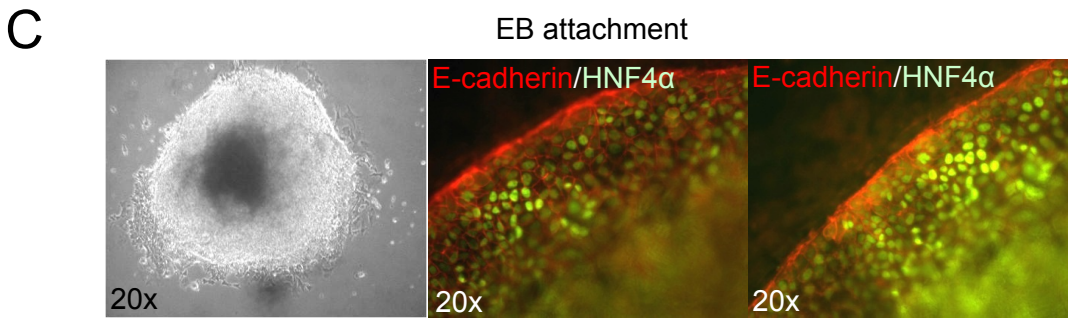

C
EB attachment
20x
E-cadherin/HNF4α
20x
E-cadherin/HNF4α
20x

Figure 4.1.5 Pluripotent mESCs can form EBs and differentiate to express nuclear HNF4α and membrane-bound E-cadherin.

(A) *Left.* Schematic of a blastocyst at E4. Cells from the inner cell mass (ICM) can be extracted and maintained on tissue culture plates; these are considered embryonic stem cells (mESCs). *Center & right*. mES-D3 cells were a gift from Dr. Prudence Talbot at the University of California, Riverside (UCR). These cells were cultured on MEFs or 0.1% gelatin-coated 6-well plates in the presence of LIF as indicated. Cells, kept under these conditions, maintained their pluripotency phenotype by expressing the pluripotency marker, OCT4 (below, IB of nuclear OCT4; 50 μg of total protein was loaded per lane). (B) Cells were subjected to a suspension culture for several days (left) or a hanging drop method for 3 days (middle) and transferred to a low adhesive plate for 2 days (right). (C) Using the latter method (B, middle), and after 2 days in suspension, EBs form spherical shapes that can be further differentiated on plates that had been coated with 0.1% gelatin. After 7 days of attachment, cells around the center mass expand outward and express nuclear HNF4α and membrane-bound E-cadherin as shown in the IF.

Fig. 4.1.6

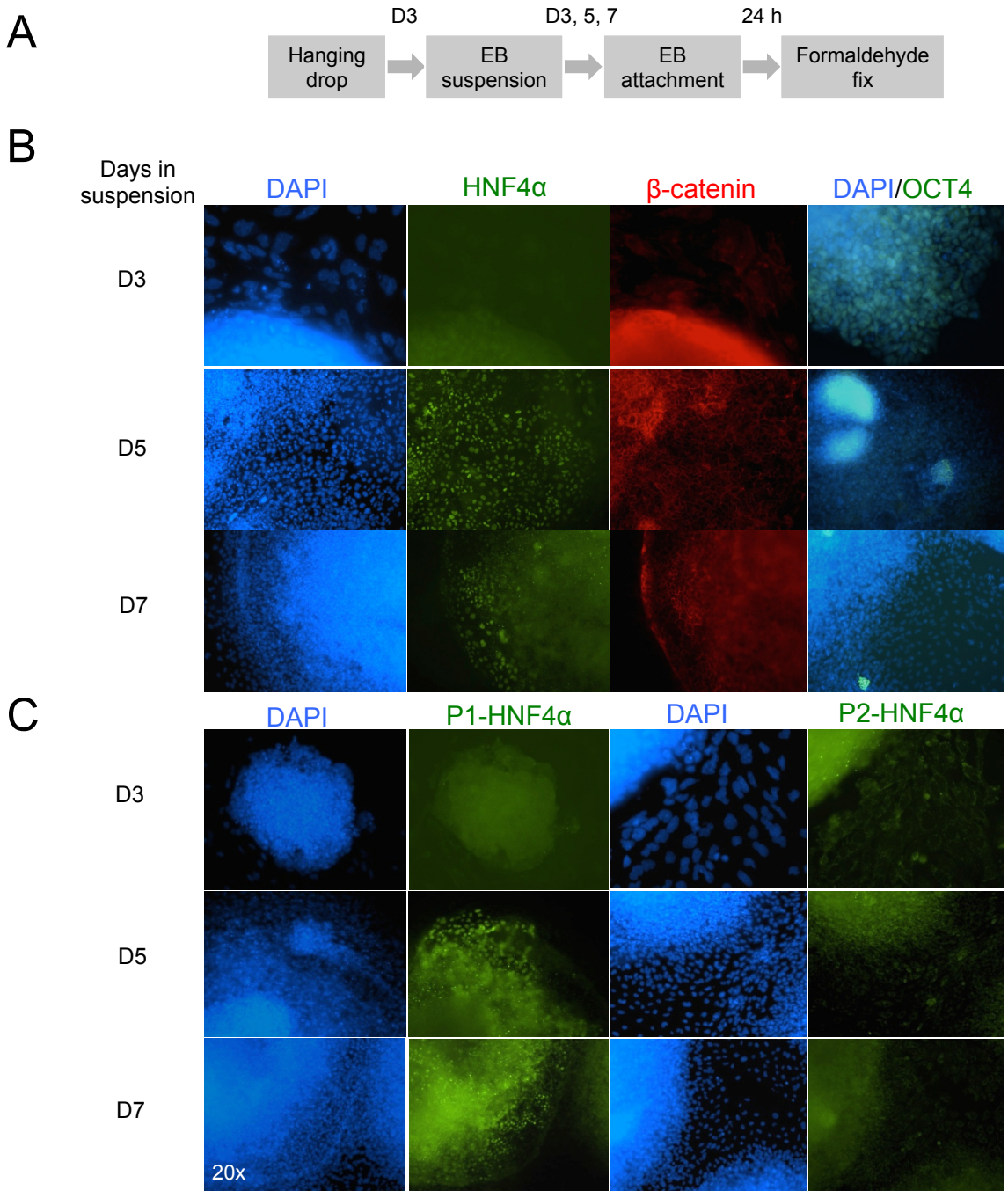
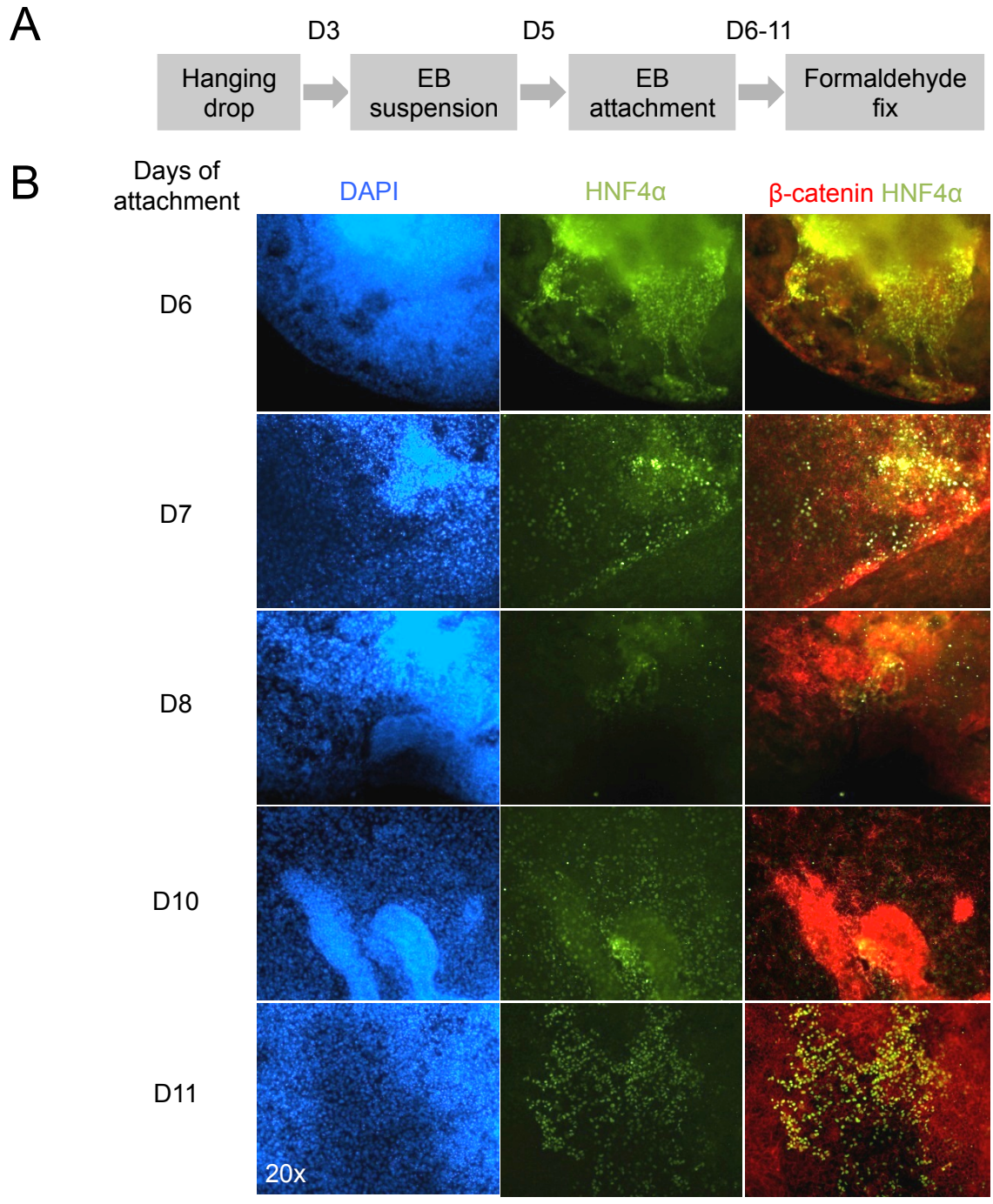
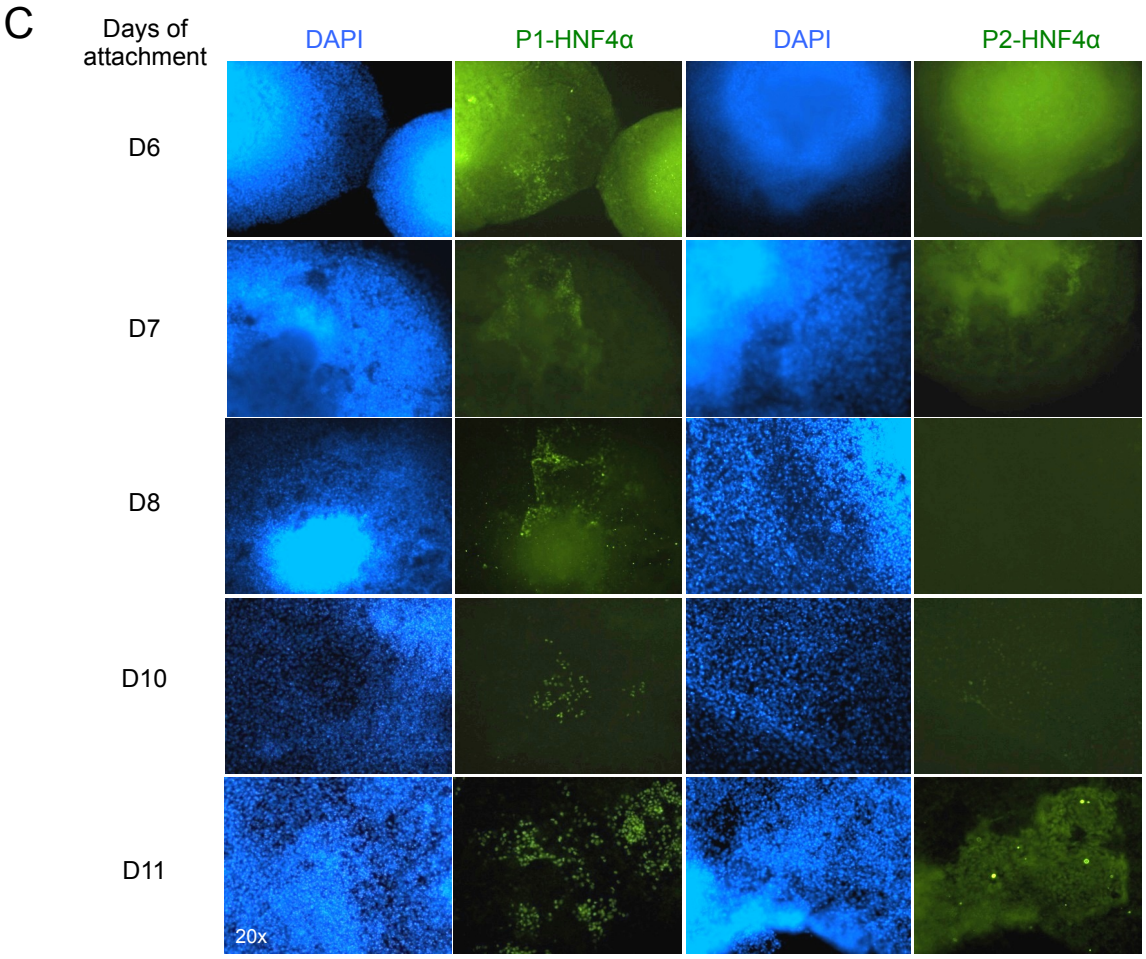

A
D3
D3, 5, 7
24 h
Hanging drop
EB suspension
EB attachment
Formaldehyde fix
B
Days in suspension
DAPI
HNF4α
β-catenin
DAPI/OCT4
D3
D5
D7
C
DAPI
P1-HNF4α
DAPI
P2-HNF4α
D3
D5
D7
20x

Figure 4.1.6 Presence of HNF4α and β-catenin and loss of OCT4 expression along the periphery of the EB.

(A) Method one of the EB formation. After cells aggregated as hanging drops, EBs were kept in suspension in media without LIF and transferred at days 3, 5, or 7 to high adhesive plates for 24 h of attachment prior to formaldehyde treatment. (B and C) IF of HNF4α isoforms, OCT4, and β-catenin expressions using the anti-P1/P2, anti-P1, anti-P2, anti-OCT4, and anti-β-catenin antibodies. OCT4 expression was lost starting at day 5 (D5) in suspension while membrane-bound β-catenin expression was detected along the periphery of the attached EB throughout the differentiation process. The anti-P1/P2 antibody gave a good signal for HNF4α starting at D5, as did the P1-specific antibody indicating expression of that isoform. The staining with the P2 antibody was less evident and included some cytoplasmic staining although additional experiments are required to verify expression and location of P2-HNF4α.

Fig. 4.1.7


A
D3
D5
D6-11
Hanging drop
EB suspension
EB attachment
Formaldehyde fix
B
Days of attachment
DAPI
HNF4α
β-catenin HNF4α
D6
D7
D8
D10
D11
20x


C
Days of attachment
DAPI
P1-HNF4α
DAPI
P2-HNF4α
D6
D7
D8
D10
D11
20x

D

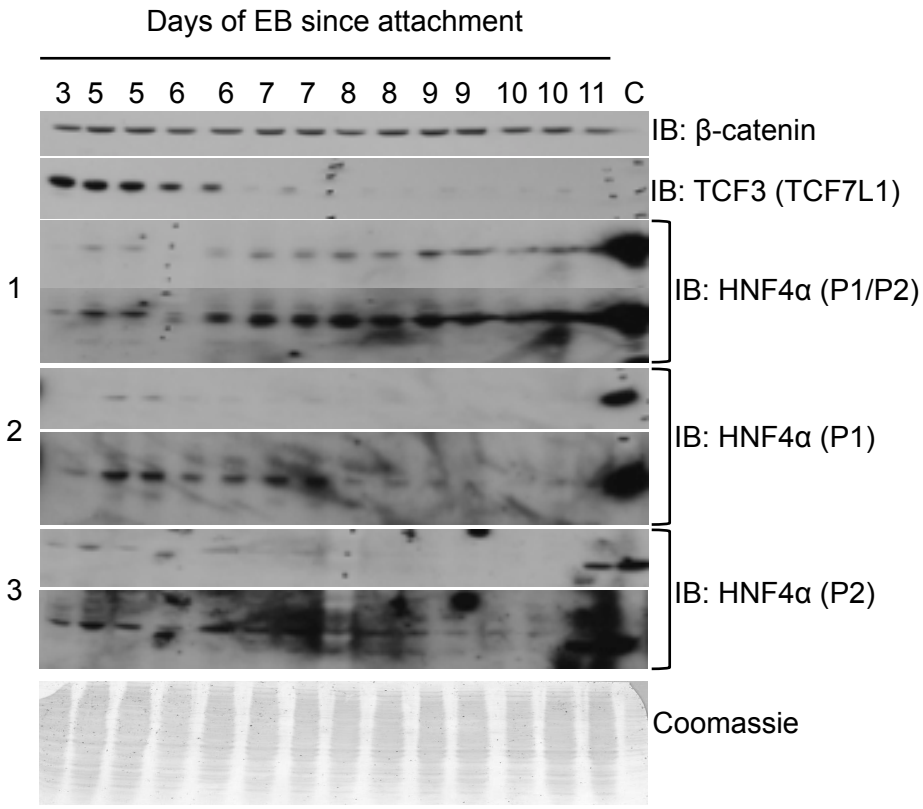

Days of EB since attachment
3 5 5 6 6 7 7 8 8 9 9 10 10 11 C
IB: β-catenin
IB: TCF3 (TCF7L1)
1
IB: HNF4α (P1/P2)
2
IB: HNF4α (P1)
3
IB: HNF4α (P2)
Coomassie

Figure 4.1.7 Tracking of the P1- and P2-HNF4α isoforms in EB development.

(A) Method two of the EB formation three days in hanging drop and another two days in suspension, EBs were transferred to high adhesive plate and allowed to differentiate for various lengths of time. Differentiated EBs were fixed at different time points (Day 6, 7, 8, 9, 10, 11). (B,C) IF showing HNF4α (P1/P2, P1, and P2) and β-catenin expression. (D) IB analysis of TCF3, HNF4α isoforms, and β-catenin from WCE of EBs at different time points of attachment. Twenty micrograms of total extract was used in the gel. These are parallel probes. 1, 2, and 3 contain short and long exposed blots. Control C, ectopic expression of HNF4α2 or HNF4α8 in COS-7. Blot 3, the control is a mix of HNF4α2 and HNF4α8 extracts from COS-7.

References

Bolotin E, Liao H, Ta TC, Yang C, Hwang-Verslues W, Evans JR, Jiang T, Sladek FM. 2010. Integrated approach for the identification of human hepatocyte nuclear factor 4alpha target genes using protein binding microarrays. *Hepatology* **51**: 642-653.

Brehm A, Ohbo K, Scholer H. 1997. The carboxy-terminal transactivation domain of Oct-4 acquires cell specificity through the POU domain. *Molecular and cellular biology* **17**: 154-162.

Briancon N, Bailly A, Clotman F, Jacquemin P, Lemaigre FP, Weiss MC. 2004. Expression of the alpha7 isoform of hepatocyte nuclear factor (HNF) 4 is activated by HNF6/OC-2 and HNF1 and repressed by HNF4alpha1 in the liver. *The Journal of biological chemistry* **279**: 33398-33408.

Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA. 2008. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes & development* **22**: 746-755.

Desbaillets I, Ziegler U, Groscurth P, Gassmann M. 2000. Embryoid bodies: an in vitro model of mouse embryogenesis. *Experimental physiology* **85**: 645-651.

Doetschman T, Williams P, Maeda N. 1988. Establishment of hamster blastocyst-derived embryonic stem (ES) cells. *Developmental biology* **127**: 224-227.

Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R. 1985. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of embryology and experimental morphology* **87**: 27-45.

Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, Darnell JE, Jr. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 7598-7602.

Erdmann S, Senkel S, Arndt T, Lucas B, Lausen J, Klein-Hitpass L, Ryffel GU, Thomas H. 2007. Tissue-specific transcription factor HNF4alpha inhibits cell proliferation and induces apoptosis in the pancreatic INS-1 beta-cell line. *Biological chemistry* **388**: 91-106.

Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**: 154-156.

Grover A, Oshima RG, Adamson ED. 1983. Epithelial layer formation in differentiating aggregates of F9 embryonal carcinoma cells. *The Journal of cell biology* **96**: 1690-1696.

Gu P, LeMenuet D, Chung AC, Mancini M, Wheeler DA, Cooney AJ. 2005. Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell differentiation. *Molecular and cellular biology* **25**: 8507-8519.

Hamazaki T, Oka M, Yamanaka S, Terada N. 2004. Aggregation of embryonic stem cells induces Nanog repression and primitive endoderm differentiation. *Journal of cell science* **117**: 5681-5686.

Hay DC, Sutherland L, Clark J, Burdon T. 2004. Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem cells* **22**: 225-235.

Herr W, Cleary MA. 1995. The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes & development* **9**: 1679-1693.

Hopfl G, Gassmann M, Desbaillets I. 2004. Differentiating embryonic stem cells into embryoid bodies. *Methods in molecular biology* **254**: 79-98.

Hwang-Verslues WW, Sladek FM. 2008. Nuclear receptor hepatocyte nuclear factor 4alpha1 competes with oncoprotein c-Myc for control of the p21/WAF1 promoter. *Molecular endocrinology* **22**: 78-90.

Kurosawa H. 2007. Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *Journal of bioscience and bioengineering* **103**: 389-398.

Lluis F, Ombrato L, Pedone E, Pepe S, Merrill BJ, Cosma MP. 2011. T-cell factor 3 (Tcf3) deletion increases somatic cell reprogramming by inducing epigenome modifications. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 11912-11917.

Lucas B, Grigo K, Erdmann S, Lausen J, Klein-Hitpass L, Ryffel GU. 2005. HNF4alpha reduces proliferation of kidney cells and affects genes deregulated in renal cell carcinoma. *Oncogene* **24**: 6418-6431.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 7634-7638.

Morrisey EE, Tang Z, Sigrist K, Lu MM, Jiang F, Ip HS, Parmacek MS. 1998. GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes & development* **12**: 3579-3590.

Mummery CL, Feyen A, Freund E, Shen S. 1990. Characteristics of embryonic stem cell differentiation: a comparison with two embryonal carcinoma cell lines. *Cell differentiation and development : the official journal of the International Society of Developmental Biologists* **30**: 195-206.

Niwa H, Miyazaki J, Smith AG. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics* **24**: 372-376.

Parslow TG, Blair DL, Murphy WJ, Granner DK. 1984. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proceedings of the National Academy of Sciences of the United States of America* **81**: 2650-2654.

Pereira L, Yi F, Merrill BJ. 2006. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Molecular and cellular biology* **26**: 7479-7491.

Sajini AA, Greder LV, Dutton JR, Slack JM. 2012. Loss of Oct4 expression during the development of murine embryoid bodies. *Developmental biology* **371**: 170-179.

Scholer HR, Balling R, Hatzopoulos AK, Suzuki N, Gruss P. 1989a. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *The EMBO journal* **8**: 2551-2557.

Scholer HR, Hatzopoulos AK, Balling R, Suzuki N, Gruss P. 1989b. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. *The EMBO journal* **8**: 2543-2550.

Schoorlemmer J, van Puijenbroek A, van Den Eijnden M, Jonk L, Pals C, Kruijer W. 1994. Characterization of a negative retinoic acid response element in the murine Oct4 promoter. *Molecular and cellular biology* **14**: 1122-1136.

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**: 688-690.

Sturm RA, Herr W. 1988. The POU domain is a bipartite DNA-binding structure. *Nature* **336**: 601-604.

Sylvester I, Scholer HR. 1994. Regulation of the Oct-4 gene by nuclear receptors. *Nucleic acids research* **22**: 901-911.

ten Berge D, Koole W, Fuerer C, Fish M, Eroglu E, Nusse R. 2008. Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell stem cell* **3**: 508-518.

Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S. 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**: 875-889.

Torres-Padilla ME, Fougere-Deschatrette C, Weiss MC. 2001. Expression of HNF4alpha isoforms in mouse liver development is regulated by sequential promoter usage and constitutive 3' end splicing. *Mechanisms of development* **109**: 183-193.

van Beest M, Dooijes D, van De Wetering M, Kjaerulff S, Bonvin A, Nielsen O, Clevers H. 2000. Sequence-specific high mobility group box factors recognize 10-12-base pair minor groove motifs. *The Journal of biological chemistry* **275**: 27266-27273.

van de Wetering M, Oosterwegel M, Dooijes D, Clevers H. 1991. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *The EMBO journal* **10**: 123-132.

van de Wetering M, Oosterwegel M, van Norren K, Clevers H. 1993. Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. *The EMBO journal* **12**: 3847-3854.

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**: 684-687.

Wirth T, Staudt L, Baltimore D. 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature* **329**: 174-178.

Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, Smith A. 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nature cell biology* **13**: 838-845.

Yi F, Pereira L, Merrill BJ. 2008. Tcf3 functions as a steady-state limiter of transcriptional programs of mouse embryonic stem cell self-renewal. *Stem cells* **26**: 1951-1960.

Chapter 5

Conclusion

The bulk of this dissertation focuses on stem cell differentiation and cancer development. We used mouse embryonic stem cells and human colorectal cancer cells as model systems to elucidate the role of the P1 and P2 promoter-driven HNF4α isoforms in proliferation and differentiation and identify factors that are involved in the mechanism that keep normal cells under control. In animal development, a single unspecialized cell, after fertilization, undergoes multiple divisions; and provided the right signaling, these cells differentiate into various cell types that make up the body. Once specialized, these differentiated cells typically do not undergo proliferation but occasionally undergo apoptosis to eliminate damaged cells. However, if one or more regulatory pathways are disrupted through accumulated mutations, the cells undergo abnormal proliferation, lose their differentiated phenotype, and fail to apoptosis, all of which are characteristics of cancer cells (Andreeff et al. 2000) (Fig. 5.1). Some consider cancer cells to be stem cell-like; both cancer and stem cells have the potential to self-renew indefinitely (reviewed in Reya et al. 2001). Understanding how these functions are controlled in normal cells and unexpectedly lost in cancer cells, in addition to understanding stem cell differentiation, are curial to cure and prevent cancer cells from rising in the first place. In the last Chapters 3 and 4, we looked closely at one of several mechanisms that control the dichotomy between proliferation and differentiation. In this Chapter, I will summarize the two projects discussed in Chapters 3 and 4, discuss whether we have addressed our hypothesis and goals, talk about the future directions for the projects, and end with a discussion on the evolution of HNF4 and TCF/LEF.

HNF4α and Colorectal Cancer (CRC)

Colorectal cancer (CRC) is the third most common cancer worldwide, affecting both sexes equally (Ferlay et al. 2004; Haggar et al. 2009). People in developed countries that adopt the high fat/low fiber Western diet are most at risk: a diet high in animal fat and meat are factors for CRC development (Boyle 2000; Larsson et al. 2006). With the increasing number of people diagnosed with CRC annually, understanding the disease will be pertinent to better prevent and/or ameliorate its impact. There is an emerging role for HNF4α in the gut although it is not yet clear exactly what that role is in colon cancer (Duncan et al. 1994; Garrison et al. 2006; Chellappa et al. 2012).

There are two promoter-driven (P1 and P2) HNF4α isoforms that differ by 38 amino acids (aa) in the N-terminal region and are expressed in different tissues. The P1-HNF4α isoform has a full length N-terminus domain and is expressed in the adult liver, whereas the P2-HNF4α isoform has a truncated N-terminus domain and is expressed in the adult pancreas (Fig. 1.2, Chapter 1). Both P1- and P2-HNF4α isoforms are found in the colon but are located in different cellular compartments of the crypt (Fig. 5.2). In the last decade or so, others have established an anti-proliferative or tumor suppressive role for HNF4α in hepatocytes and hepatocellular carcinoma cells (Späth and Weiss 1997 & 1998; Lazarevich et al. 2004; Walesky et al. 2013). More recently, The Cancer Genome Atlas (TCGA) found *HNF4A to be* amplified in CRC on the genomic level while Zhang et al found a corresponding increase in protein abundance (Cancer Genome Atlas (TCGA) 2012; Zhang et al. 2014). Zhang et al further analyzed the shRNA knockdown of both P1 and P2 isoforms or the P1 isoform alone, performed by the Achilles project, and found that knockdown of both forms negatively impacted cellular proliferation while the knockdown of P1 alone was ambiguous (Zhang et al. 2014; Cheung et al. 2011). The

findings from both groups suggested that HNF4α might function as an oncogene in colon cancer. However, neither group was able to indicate which isoform might be amplified. Since P1-HNF4α has been shown to be decreased in CRC (Chellappa et al. 2012; Tanaka et al. 2006), we hypothesized that perhaps these groups were seeing an amplification of P2-HNF4α in CRC.

In Chapter 3, we examined directly the role of P1 and P2-HNF4α (HNF4α2 and HNF4α8, respectively) in the human colon cancer. We established Tet-On inducible lines in HCT116 that expressed either the P1 or P2 isoform in a controlled manner and performed RNA-Seq and ChIP-Seq on these lines (parental/PL, HNF4α2, and HNF4α8). In the RNA-Seq, there were more genes upregulated than downregulated for both HNF4α isoforms and P2-driven HNF4α8 had slightly more dysregulated genes than HNF4α2. Comparing the two lines, HNF4α2 and HNF4α8 regulate many genes in common especially those involved in metabolism processes, drug detoxification, and oxidative stress. There were also genes that were uniquely regulated by a single isoform; HNF4α2 regulated genes involved in cell death and growth inhibition while HNF4α8 regulated genes involved in growth promotion and anti-apoptosis. In the ChIP-Seq assay for HNF4α, we found that many of the HNF4α peaks that were within 50 kb of genes, some of which were dysregulated in the RNA-Seq, have TCF4 (*TCF7L2*) motif enriched in these peaks. Closer analysis of these overlapping peaks identified at least 770 peaks where HNF4α recruits, co-localizes, or competes with TCF4. In the recruited and co-localized peaks both factors co-occupy the CTTTG core motif while the competed peaks, as well as the co-localized peaks, co-occupy the AP-1 bound TGAxTCA motif. Interestingly, only HNF4α2 (P1 isoform) can compete with TCF4, we did not find any instances of HNF4α8 (P2 isoform) competing with TCF4, although that

could be due to fewer HNF4α8 peaks than HNF4α2; these were also fewer: TCF4 (+/-DOX) peaks in the α8 line compared to the α2 line. We are not certain why this is so, but this could be due to differences in sequence depth or possible technical issues during the ChIP-Seq experiment.

In addition to examining overlapping sites between HNF4α and TCF4 in vivo ChIP-Seq assay, we also used a modified version of the high-throughput protein binding microarray (PBM) to look for more overlapping binding specificity between HNF4α and TCF and found 90 unique sequences that were bound by both HNF4α2 and TCF1, and 741 sequences that were bound by HNF4α (both HNF4α2 and HNF4α8 isoforms) and TCF4. We tested for competition between the two factors in vitro using a luciferase reporter assay and saw a competition on the element that has a higher affinity for HNF4α and TCF but not on the other elements that have a weaker affinity for one of the factors. All sequences identified by the PBM contain the CTTTG motif with variations in the flanking regions; this is consistent with the in vivo ChIP-Seq assay. Imbedded in the 741 sequences are single nucleotide polymorphisms (SNPs); we found that there are 107 instances out of the 741 in which a SNP altered the binding of one (TCF4) or both (HNF4α and TCF4) factors. Lastly, we performed functional analyses to determine the effect of HNF4α isoforms on tumor growth. We found that the α8 line does not affect the tumor size even in the presence of DOX, while the α2 line, in the presence of DOX, significantly decreased tumor size compared to the parental line (+DOX). We also found that HNF4α8 increases while HNF4α2 decreases invasion and migration.

All told, our findings provide evidence that the P1- and P2-HNF4α isoforms exhibit subtle yet significant functional differences in colon cancer cells. HNF4α8, the P2 isoform, seems to be more permissive of proliferation while HNF4α2, the P1 isoform, is

clearly tumor suppressive in colon cancer. Importantly, in contrast to the studies suggesting that HNF4α might be oncogenic due to an amplification in CRC (Cancer Genome Atlas (TCGA) 2012; Zhang et al. 2014), our results do not indicate that either isoform is oncogenic. Indeed, a cell proliferation assay (both cell counting and MTT assays) showed that HNF4α8 also decreases proliferation, as does HNF4α2; it just seemed to have less effect in vivo. We also noticed that the parental line in the presence of DOX had decreased cell numbers and viability suggesting that the DOX was probably toxic to the cells at the concentration we used (0.5 μg/mL) for the proliferation assay.

HNF4α and the Wnt/β-catenin/TCF pathway

In the last decade, there have been numerous reports on the interaction between nuclear receptors (NRs) and the canonical Wnt/β-catenin/TCF pathway. Both NRs and Wnt/β-catenin/TCF pathway are important in cellular development, differentiation, and regulation of gene expression. Disruption or dysregulation of these factors can lead to altered gene expression and pathogenesis (reviewed in Mulholland et al. 2005). Also the pro-oncogene c-Jun/AP-1 complex has been shown to interact with NRs as well (reviewed in Pfahl 1993; Schmidt et al. 1993; Biddie et al. 2011). While others have examined the interplay between HNF4α and TCF4 (Frietze et al. 2012; Gougelet et al. 2014; Yang et al. 2013), they have not looked specifically at the different isoforms and their interaction with TCF4 or AP-1 on the chromatin in colon cancer.

Molecular interplay between HNF4α, TCF4, and AP-1 on chromatin in colon cancer

In our ChIP-Seq analysis of the HCT116 inducible cell lines, we found both HNF4α2 and HNF4α8 recruit and co-localize with TCF4 on the CTTTGA consensus

sequence in colon cancer; however, it is not clear whether both factors are bound at the same time. Our gel shifts certainly did not show cooperative binding; instead, we observed separate shift bands for HNF4α and TCF1 on the element that has a strong affinity for both HNF4α and TCF (TGACCTTTGATCT). It is possible that the trimeric complex (HNF4α-TCF-DNA) is unstable in gel shift and can only be observed in vivo. Motif analysis of the ChIP-Seq peaks also showed that HNF4α co-localizes with TCF4 on the AP-1 binding site.

We are not the first to find enrichment of TCF and AP-1 consensus sequences in HNF4α peaks; another group earlier also found enrichment of AP-1 and TCF/LEF sites in HNF4α ChIP-Seq peaks in CaCo2 cells that express both P1- and P2-HNF4α isoforms (Weltmeier and Borlak 2011). Interestingly, we found that only HNF4α2 competes with TCF4 on the AP-1-bound site but not on the consensus site; we could not find any HNF4α8 and TCF4 competing peaks. Since HNF4α2 has a longer A/B domain, we predict that this domain could be responsible for the competition with TCF4 (Fig. 1.2, Chapter 1). While we have not investigated a potential direct interaction between HNF4α and AP-1 on the TGAxTCA site, one group has shown a competition between HNF4α and AP-1 on the AP-1 site juxtapose to the HNF4α site (Yang et al. 2009). Therefore, we do not know if HNF4α binds the AP-1 complex or displaces AP-1. There is a report suggesting that the AF-1, AF-2, and the LBD domains of estrogen receptor (ER) associate with AP-1 complex proteins (Webb et al. 1999). Also, there are reports showing interaction between β-catenin/TCF4 and AP-1 involved in the regulation of Wnt/β-catenin/TCF target genes (Nateri et al. 2005; Toualbi et al. 2007). However, in our ChIP-Seq, we could not find any HNF4α and TCF4 co-localized or competed peaks that overlapped with β-catenin ChIP-Seq peaks in the HCT116 (Bottomly et al. 2010); but

then again, our ChIP-Seq conditions and cell lines are different from theirs and may not have covered all regions of the chromatins. All told, our ChIP-Seq data suggest a molecular interplay between HNF4α, β-catenin/TCF4, and AP-1 but additional biochemical assays are needed to tease out the specific mechanism.

HNF4α and TCF4 DNA bending

In Chapter 3, we propose schematic models for the recruitment and co-localization of TCF4 by HNF4α (Fig. 3.11, Chapter 3). We predict that HNF4α contacts the major groove of the DNA helix and bends it towards the protein, allowing TCF4 to be recruited to the site. We hypothesize that HNF4α co-localizes to sites that are configured to a favorable conformation made by TCF4. There are studies on DNA binding for the TCF/LEF and NRs but no study has examined the two transcription factors together. The conserved high mobility group (HMG) domain shared by all TCF/LEF family members has been shown to bind to the minor groove of the double stranded (ds) DNA helix and bend it at 130^{o} towards the major groove (Love et al. 1995; Giese et al. 1992). HNF4α and all other NRs contain a highly conserved DNA binding domain (DBD) that is composed of two zinc (Zn) fingers that bind to the half-sites of response elements. In their circular permutation analysis, Jiang et al showed that the full-length HNF4α bends DNA by 80^{o}, while a truncated HNF4α bend slightly less, by 60^{o} (Jiang et al. 1997). There is crystal structure of HNF4α but it does not show any obvious bending (Chandra et al. 2013); thus it is not clear which direction HNF4α bends DNA. Since all nuclear receptors share amino acid sequence similarity (Sladek 2011; Renaud and Moras 2000), we examined in the published literature on the DNA binding properties of other NRs to predict how HNF4α might bend DNA. There are circular permutation analyses on DNA

binding properties for retinoic acid receptor (RAR)-retinoid x receptor (RXR) and thyroid hormone receptor-RXR heterodimers and RXR-RXR homodimer. The data showed that these NRs bend DNA by 50^{o} to 92^{o} toward the major groove (Lu et al. 1993). Other studies on glucocorticoid and estrogen receptor showed that a portion of the DBD inserts into the major groove of the DNA (Lumpkin et al. 1985; Schwabe et al. 1990). Therefore, we hypothesis that when HNF4α contacts the major groove, it too bends double stranded DNA towards the major groove; this would expose the minor groove, making it potentially easier and more energetically favorable for TCF4 to contact the minor groove and further bend it in the direction of HNF4α (Fig. 3.11, Chapter 3). To prove this, we would need to perform a circular permutation analysis of HNF4α and TCF4 binding to the TGACCTTTGACCT sequence. We already know how the P1- HNF4α bends DNA since Jiang et al already showed this (Jiang et al. 1997), we do not know, however, how the P2-HNF4α bends DNA. What is more important is how the factors interact on the same piece of DNA. This would provide some answers as to how HNF4α might recruit and co-localize with TCF4. Although our gel shift results have not shown the two factors interact on the same piece of DNA, we would need to test other gel shift conditions where a trimeric complex can exist in a gel shift.

HNF4α and co-regulators

TCF4 and AP-1 are not the only regulators that interact with HNF4α. HNF4α isoforms have been shown to interact differentially with other co-regulators through direct contacts with the two activation function modules of HNF4α, AF-1 and AF-2. Since the P1- and P2-HNF4α isoforms differ only in their N-terminal domains, only P1-HNF4α isoforms have an AF-1 function from the extended 24 aa in the A/B domain, while both

P1- and P2-HNF4α isoforms have an AF-2 function on the 12th helix of the ligand binding domain (LBD) (Fig. 1.2, Chapter 1). The P2-HNF4α isoforms, which have a truncated A/B domain consisting of only 16 aa, on the other hand, lack an AF-1. A GAL4 fusion of the truncated A/B portion of the P2-HNF4α7 showed that this portion did not result in transactivation on a reporter for GAL4 (Torres-Padilla et al. 2002).

AF-1 and AF-2 have been shown to interact with co-activators such as glugocortiocid receptor-interacting protein (GRIP-1), cAMP response element-binding protein-binding protein (CBP), nuclear receptor coactivator 1 (NCOA1), while AF-2 has been shown to interact with co-repressor as well including silencing mediator for retinoid and thyroid receptors (SMRT) (Sladek et al. 1999; Ruse et al. 2002; Torres-Padilla et al. 2002) (see http://www.cisreg.ca/tfe for additional list of co-regulators that have been shown to interact with HNF4α). Although P1- and P2-HNF4α both interact with SMRT, SMRT interaction with P1-HNF4α but not P2-HNF4α enhances HDAC recruitment. Also, the promoter context dictates which coregulators are recruited by HNF4α to the chromatin (Torres-Padilla et al. 2002; Torres-Padilla and Weiss 2003). With TCF4 and AP-1 now in the model, it will be important to determine what and how co-regulators are involved when HNF4α recruits TCF4 to the CTTTG consensus site, or when HNF4α co-localizes or competes with TCF4 on the AP-1 bound sites of target genes.

Single Nucleotide Polymorphisms (SNPs) and HNF4α/TCF4 binding sites

The findings from the second PBM experiment in Figure 3.7 in Chapter 3 suggest that a single nucleotide polymorphism (SNP) can alter the binding of TCF4 and/or HNF4α. It would be of interest to identify other naturally occurring SNPs in the overlapping TCF4/HNF4α ChIP peaks of the inducible HCT116 lines and determine how

many affect HNF4α and/or TCF4 binding in vivo as well as expression of nearby genes. Just recently, a group identified several CRC risk-associated SNPs that are linked to specific genes; some of the SNPs are in distal (>500 kb) and proximal enhancer binding sites, while others are in intronic or exonic regions (Yao et al. 2014). We cross-referenced our ChIP-Seq and RNA-Seq to their list of genes associated with SNPs in CRC and found eleven genes dysregulated in our RNA-Seq. Three genes (*E2F2, FOXM1, RAD51AP1*) that were found dysregulated by HNF4α2 have HNF4α2 peaks nearby and three genes (*TERC, BHLHE40, KCTD15*) that were found dysregulated by HNF4α8 do not have HNF4α8 peaks nearby. Since we only look for genes within 50 kb of the peak center, and since these genes could be regulated by enhancers that are several hundred kilobases away, as Yao et al points out, we would want to correlate genes to enhancer sites that are associated with SNPs and determine whether the SNPs alter the binding of HNF4α at that region.

Role of alterative promoters in cancer

HNF4α is one of many transcription factors (TFs) where the promoter-driven isoforms take on different roles in cancer and normal development. Depending on which alterative promoter is used, one transcript variant may act as a tumor suppressor and the other functions as an oncogene (Davuluri et al. 2008). Examples include *LIF1* – lymphoid enhancer factor protein 1, *MYC*, *RASSF1* – Ras effector, and the tumor protein (TP) of the p53 family genes (TP63, TP73) (Davuluri et al. 2008), where the different N-terminal regions have biologically different roles. The full length LEF1 in the TCF family is driven by a P1 promoter and interacts with β-catenin to activate the Wnt/β-catenin pathway in colon cancer tissue and cell lines, whereas the truncated LEF1, driven by a

second P2 promoter, cannot interact with β-catenin and is silenced in cancer (Arce et al. 2006). Similarly, the oncogene *MYC* is controlled by two alternative promoters, where the P2 promoter drives the expression of the transcript that is expressed predominantly in normal tissues, and the P1 promoter drives the expression of a less abundant transcript that is the most deregulated (Marcu et al. 1992). In the TP family, the shorter form of p73 is associated with cancer development while the longer form is not (Murray-Zmijewski et al. 2006). Likewise, the RASSF1A form is tumor suppressive, whereas RASSF1C is growth promoting (Amaar 2006). Another interesting example is the androgen receptor (AR) where the full length AR-B and the shorter AR-A isoform is present in healthy colonic mucosa, but in neoplastic mucosa, only the AR-A isoforms remain present (Catalano et al. 2000). Another example is the progesterone receptor (PR). There are two isoforms – PR-A and PR-B – that are transcribed by two distinct estrogen-inducible promoters. The PR-A lacks 164 amino acids (aa) at the N-terminus and acts as a repressor of gene expression, whereas PR-B contains those aa and acts as an activators of progesterone-response genes (Kastner et al. 1990); each isoform interacts with functionally different regulators (Giangrande et al. 1999 and 2000). The evidence of HNF4α promoter-driven isoforms taking on different roles is shown in liver development, knockin mouse model, and colon cancer. Torres-Padilla et al looked at a few genes that express either early or late in development and found that the P1- and P2-HNF4α appear to activate different sets of genes at different stages of liver development by interacting differentially with coregulators (Torres-Padilla et al. 2002; Torres-Padilla and Weiss 2003), with P2-HNF4α activating early genes while P1-HNF4α activating late genes in development (Torres-Padilla et al. 2001). Generation of the exon-swap mice by Briancon and Weiss showed dyslipidemia and expression of some

genes reduced in the “α7” only mice compared to the “α1” only mice (Briancon and Weiss 2006). Finally, Chellappa et al showed that active Src phosphorylates and destabilizes the P1-HNF4α expression but not the P2-HNF4α in colon cancer (Chellappa et al. 2012). These are the current findings where the promoter-driven P1 and P2 isoforms of HNF4α taking on different role during normal and cancer development; our work provides additional evidence of a duel role in cancer. Closer investigation on the different isoforms encoded by one gene through alterative promoter sites and splicing events in different cellular conditions will be important to understanding the process of both physiological and pathological development (Davuluri et al. 2008).

HNF4α in mouse embryogenesis

Since the discovery of mouse embryonic stem cells (mESCs) in 1981 (Evans and Kaufman 1981; Martin 1981) and the derivation of human embryonic stem cells (hESCs) in 1998 (Thomson et al. 1998), ES cells have been used for various studies including differentiation to specific lineages, genetic manipulations, drug testing, and therapeutic replacement of damaged tissues (Keller 2005). ESCs, derived from the inner cell mass of blastocyst-stage embryos at E4.0, survive – with appropriate culturing conditions – indefinitely and can differentiate into any tissue of the three germ layers. Taking advantage of their pluripotency. In Chapter 4, we generated Tet-On inducible lines that express either P1 or P2-driven HNF4α under DOX induction in mES cells and investigated the effect of ectopically expressed HNF4α isoforms on these cells. Surprisingly, expression of either HNF4α isoform alone was able to drive differentiation with 72 hours of induction even in MEF-conditioned media containing LIF (which normally helps to maintain the undifferentiated state) and without the use of any

differentiation inducing growth factors. We have performed RNA-Seq to identify global changes in gene expression and will be determining what types of cells these are.

Understanding the role of the P1 and P2-HNF4α isoforms in differentiation may help to elucidate the molecular transdifferentiation of pancreas to liver and vice versa and the mechanism by which one HNF4α isoform is silent while the other remains active in the adult liver and pancreas. Both the liver and pancreas arise from the same precursor cells during development where the P1 and P2-HNF4α isoforms are thought to be co-expressed. After birth, one of the isoforms is turned off: the adult liver expresses only the P1-HNF4α isoforms, whereas the adult pancreas expresses only the P2-HNF4α isoforms (Briancon et al. 2004; Torres-Padilla et al. 2001). Transdifferentiation between the pancreas and liver is possible because they come from the same precursor of the endodermal lineage (reviewed in Shen et al. 2003). Other examples of transdifferentiation include smooth to skeletal muscle cells, epithelial pigment cells to lens cells and esophagus epithelium to columnar epithelium resembling the small intestine also known as Barrett's esophagus (Araki and Okada 1977; Eguchi and Okada 1973; Patapoutian et al. 1995; Stairs et al. 2008; reviewed in Shen et al. 2003).

We also took advantage of the fact that mESCs are used as a model system for embryogenesis. Stem cells can form into embryoid bodies (EBs) that resemble the first few stages of early embryogenesis. We were able to stain for HNF4α in cells located at the periphery of the EB sphere. However, we are not completely certain which isoform it is; but based on preliminary data (Fig. 4.1.S6 and 7, Appendix to Chapter 4), we predict that it is the P1-HNF4α that comes on first and the P2-HNF4α comes on later. We found cells along the periphery of the EB that expressed HNF4α also expressed E-cadherin

and these cells resembled primary endodermal cells surrounding the inner cell mass (Doetschman et al. 1985).

With the development of inducible pluripotent stem (iPS) cells by Takahashi and Yamanaka in 2006 (Takahashi and Yamanaka 2006), scientists were able to take any somatic cells and revert the cells back to a pluripotent state, and thereby have the capacity to differentiate into any cell lineage in the body. One approach to identify the role of the different HNF4α isoforms in different tissues is to use iPS cells from the transgenic mice generated by Briancon and Weiss 2006 (Briancon and Weiss 2006). These are knockin 'HNF4α7' and 'HNF4α1' mice that express only one HNF4α variant in tissues that express HNF4α; despite the lack of one isoform, both the P1 and P2 promoters are still intact and these embryos can make healthy and fertile mice suggesting the HNF4α isoforms are redundant under conditions of normal development.

We have derived MEFs from control wide-type WTα7 and WTα1 embryos, as well as knockin homozygous HMZα7 and HMZα1 embryos and have induced these cells to become iPS cells using the previously described method (Takahashi and Yamanaka 2006). We have these cells in stock but have not done any in depth characterization on them as of yet. The plan is to make hanging drops, compare their ability to form EBs, determine the timing of the HNF4α isoform expression in the EB, and identify the type of cells these iPS cells differentiate into in the EB aggregate using lineage specific markers (GATA binding protein transcription factor (GATA) 4 and 6, forkhead box protein (Fox) A1 and 2, alpha fetal protein (AFP), and paired box gene (Pax) 4 and 6) (Keller 2005). These cells could be a good model system with which to identify the function of the promoter-driven HNF4α isoforms in different tissues.

Future direction

Ultimately, we want to move these two projects on the role of HNF4α isoforms in colon cancer and embryonic development forward. So how do we go about that and in what direction? What new discoveries will we find and how will those findings broaden our understanding of what we already know? These are the questions we will want to consider in the next phase of investigation.

For the HNF4α isoforms and CRC project, one possible direction is to use a mouse model to study colorectal cancer. Mouse models of CRC have been used since 1928 to study the pathology of CRC. Two well known models have been established and used extensively by others: the familial adenomatous polyposis (FAP) model, an inherited mutation in the *APC* allele, and the hereditary non-polyposis colorectal cancer (HNPCC) model, inherited mutations in one of several DNA mismatch repair (MMR) genes (Moser et al. 1990; Kucherlapati et al. 2010; Karim et al. 2013). In addition, there are various chemicals that induce CRC in mouse; these include dimethydrazine (DHM), azoxymethane (AOM), and dextran sulfate sodium (DSS) to name a few (De Robertis et al. 2011; Narisawa et al. 1974; Reddy et al. 1981; Andreassen et al. 2002; Yu et al. 2005; Karim et al. 2013).

One area of study would be to examine a potential three-way interplay between HNF4α, TCF4 and AP-1 in normal mouse colon and determine how that mechanism might be different or lost in mouse model of colitis-associated colon cancer (CAC). Both P1 and P2-HNF4α are known to be expressed in the colon (Garrison et al. 2006; Babeu et al. 2009), but our lab in unpublished work has found that P1- and P2-HNF4α are expressed in different compartments in the colonic epithelium: P1-HNF4α is expressed in cells located at the differentiated compartment while P2-HNF4α is expressed in cells

located in the proliferative compartment where Wnt/β-catenin/TCF4 activities are high and AP-1 activity is found (Giles et al. 2003; Sancho et al. 2009). The latter compartment is also where stem cells reside (but we did not see HNF4α expressed in staining) (Korinek et al. 1998). Sladek lab members have also found that the two promoter-specific variants regulate different genes and respond to stress such as colitis and CAC differently as well (manuscript under revision). They collected tumor samples from control and CAC colon tissues from WT, HMZα7, and HMZα1 mice. We could use WT samples from control and CAC colon tissues to understand the interplay between the HNF4α isoforms and the Wnt/β-catenin/TCF4/AP-1 pathways in normal vs tumor colon. We propose that in a normal WT colon, the cells in the differentiated compartment where P1-HNF4α and TCF4 are co-expressed, competition is occurring, while the cells in the proliferative compartment where P2-HNF4α is found, instead of competing, HNF4α is helping and enhancing β-catenin/TCF4 and AP-1 transcription activities (Fig. 5.2). In the same project the Sladek lab found that the colon tumors lost expression of the P1 isoform but continued to express the P2 isoform. We predict that with P1-HNF4α gone, the P2-HNF4α is present everywhere in the crypt and synergizes with TCF4 and AP-1 to active target genes to promote tumor progression. The P1-HNF4α has been shown to inhibit the P2-HNF4α expression (Briancon et al. 2004).

To test this proposal, we would first look for protein-protein interactions between HNF4α, TCF4 and AP-1 using a rapid immunoprecipitation mass spectrometry of endogenous proteins also known as RIME (Aebersold et al. 2003; Wells et al. 2002). The first step of this assay is basically the ChIP assay, but instead of purifying the DNA, we would pull out the protein-protein complex on chromatin and send the samples for mass spectrometry. We could do RIME and ChIP-Seq at the same time to identify

protein-protein interactions on promoters of target genes. We would immunoprecipitate HNF4α using HNF4α-specific antibodies and TCF4 using an antibody against the TCF4 protein. Since we have expression array data for colon tissues, we could cross-reference the ChIP-Seq to the mRNA levels to look for direct target genes and identify motifs for overlapping and non-overlapping binding of HNF4α, TCF4, β-catenin, and AP-1 in normal colon and tumor samples. We would find differential interaction between the HNF4α isoforms with TCF4 and AP-1 in normal and tumor colon tissues.

Finally, we could use the exon-swap HMZα7 and HMZα1 mice (Briancon and Weiss 2006) as models to better understand the function of the P1 and P2-HNF4α isoforms in liver and pancreas development. We could take out livers and pancreas at E12.5 (days post-coitum) and two days after birth to do ChIP-Seq for HNF4α follow by RNA-Seq to determine targets and functions of the HNF4α isoforms in these tissues. Both isoforms are expressed before birth and only one isoform remains expressed after birth in specific tissues. It would be of interest to determine the function of P1 or P2-HNF4α isoform in both liver and pancreas development when normally there is a spatial and temporal expression of both isoforms in the development of WT mice. Even though these transgenic mice are viable and fertile, suggesting that the isoforms are redundant, we expect to find subtle differences between the 'α7 only' and the 'α1 only' mice in liver and pancreas development since under conditions of stress the isoform-specific mice show dramatic differences, suggesting that the balance between the isoforms in normal development is rather tenuous. These findings could provide insight into the development of maturity onset diabetes of the young 1, a form of type II diabetes that is induced by mutations in the *HNF4A gene*.

Evolution of HNF4α and TCF/LEF

HNF4 is one of the most ancient nuclear receptors (NRs). The most primitive metazoans (cnidarians and sponges) expressed HNF4. Sponges expressed one or two NRs that had amino acid similarity to HNF4 (Sladek 2011). Since these are the simplest animal organisms with no tissues or organ systems, the NRs in sponges must carry out very basic functions. As organisms became more complex, the number of NRs increased (Sladek 2011). For example bilaterians and mammalians have 25 and 48 NRs, respectively (Bertrand et al. 2004; Robinson-Rechavi et al. 2001); but there are exceptions. *Caenorhabditis elegans*, which are less complex than mammalians, have about 270 predicted NRs (Sluder et al. 2001). It turns out that there are only 15 NR families in *C. elegans*; since an explosive duplication event of the NR HNF4 occurred some time ago, this brought about 255 variations of a single NR2A (HNF4) that exist in *C. elegans* (Robinson-Rechavi et al. 2005). Mammalians also have variations of a single HNF4; humans have HNF4α and HNF4γ (Sladek 2011), while Xenopus has a third HNF4 member, HNF4β (Holewa et al. 1997).

From a very simple *HNF4-like* gene that is about 2.3 kb that encodes a 636 amino acids protein in sponge (Larroux et al. 2006), the *HNF4A* gene in mammalians is about 30 kb and is transcribed by two promoters (P1 and P2) and 3' splicing events, producing several HNF4α isoforms that are expressed at different stages of development and in different tissues of the adult mammalians. Functionally, the isoforms interact with different co-regulators and other transcription factors to regulate gene expression during development and in the lifespan of the organism. The isoforms provide a more complex and plastic level of regulation.

Similar to HNF4, the Wnt/β-catenin/TCF pathway is also found in metazoans including sponges and *C. elegans* (Cadigan and Waterman 2012). The HMG box domain of TCF/LEF is highly conserved (Castrop et al. 1992) in metazoans. In Xenopus, there are two TCF factors (XTcf3 and XLef1) that are orthologous to mouse (Molenaar et al. 1998); Drosophila has one TCF that is closely related to mammalian TCF1 (Brunner et al. 1997); and the POP-1 found in *C. elegans* is homologous to TCF/LEF (Lin et al. 1995; reviewed in Roose and Clevers 1999). While almost all invertebrates and some vertebrates have at least one TCF gene, amphibians and mammalians have four TCF genes that encode for various TCF isoforms as a result of gene duplication (Cadigan and Waterman 2012). Depending on which co-regulators interact with these isoforms, they can activate or repress transcription to regulate specific target genes of the Wnt/β-catenin pathway in normal and cancer development (Cadigan 2008; Cadigan and Waterman 2012).

All told, both HNF4 and TCF/LEF are present in metazoans at the same time. More complex animals have multiple isoforms of these two factors that are driven by alterative promoters and 3' splicing events. Thus, the HNF4 and TCF/LEF isoforms play important and differential roles in metazoan development (Sladek 2011; Teo et al. 2006). Hence, it is possible that the two factors might evolve together to counteract each other's action.

Fig. 5.1

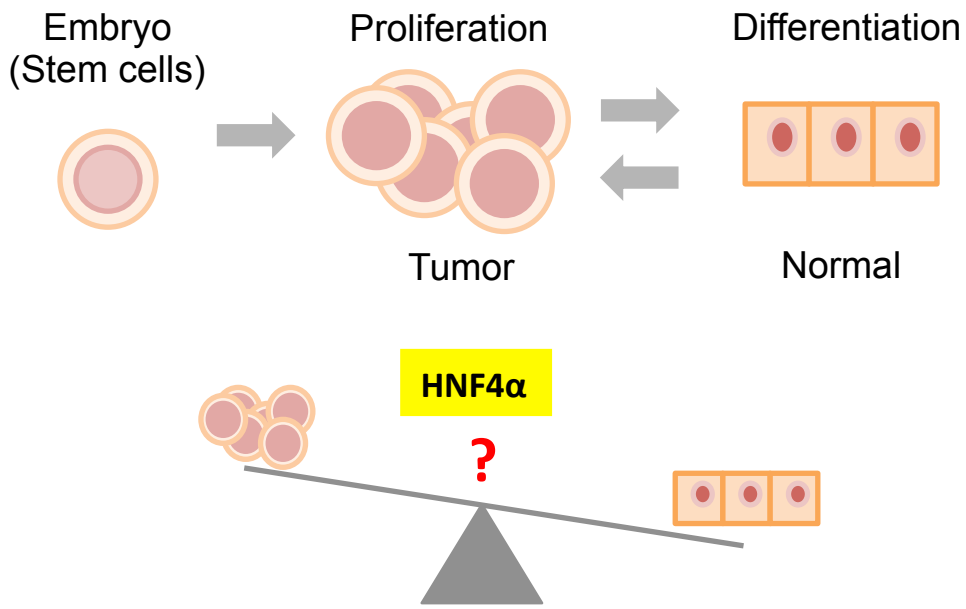


Figure 5.1 Schematic of the dichotomy between proliferation and differentiation.

An oocyte starts out as a single cell. After fertilization, the cell undergoes multiple divisions and differentiates into various cell types that make up the body. These cells remain differentiated and do not undergo proliferation; however, if one or more regulatory systems become aberrant, proliferation is uncontrollable, there is less cell-cell contact, and these abnormal cells become cancerous. Our focus is on the transcription factor HNF4α and how it balances the dichotomy between proliferation and differentiation in cells.

Fig. 5.2

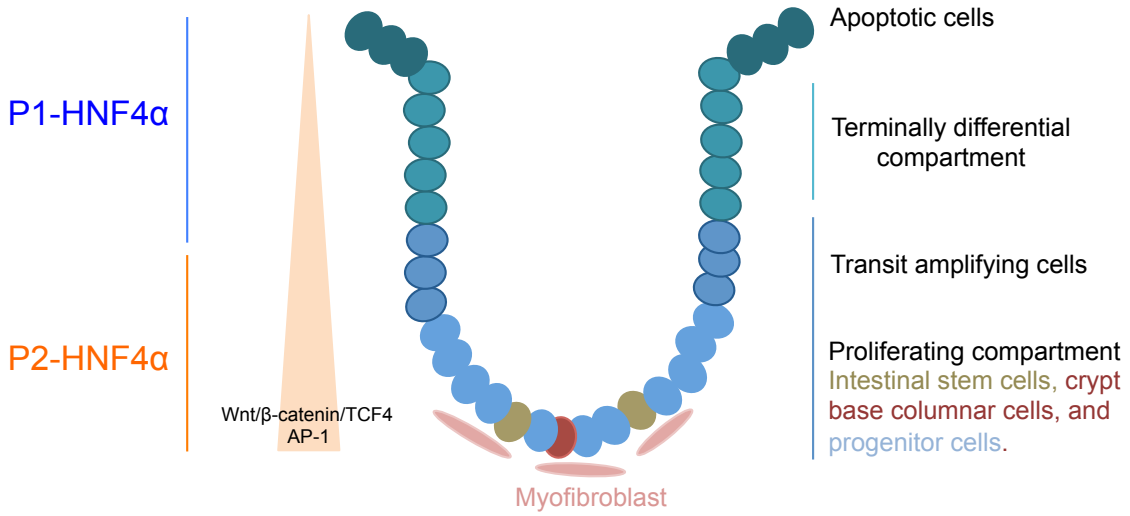


Figure 5.2 P1- and P2-HNF4α expression in the colon crypt.

Schematic design of a colonic crypt showing the Wnt/β-catenin/TCF4 gradient and the type of cells that reside in the crypt: cells at the top of the crypt are undergoing apoptosis; differentiated cells are in the terminally differential compartment; cells undergoing proliferation reside in the proliferating compartment, here includes intestinal stem cells, crypt base columnar cells, and progenitor cells; myofibroblast cells are found at the bottom of the crypt (Medema and Vermeulen 2011). The Sladek lab in unpublished work has shown that P1-HNF4α is expressed in the differentiated compartment of the colonic crypt whereas P2-HNF4α is found expressed in the proliferative compartment where there is a higher concentration of active Wnt/β-catenin/TCF4 and AP-1 activities (ven de Wetering et al. 2002; Sancho et al. 2009). Redrawn from van de Wetering et al. 2002 and Barker 2014.

References

Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. *Nature* **422**: 198-207.

Amaar YG, Minera MG, Hatran LK, Strong DD, Mohan S, Reeves ME. 2006. Ras association domain family 1C protein stimulates human lung cancer cell proliferation. *American journal of physiology Lung cellular and molecular physiology* **291**: L1185-1190.

Andreassen A, Mollersen L, Vikse R, Steffensen IL, Mikalsen A, Paulsen JE, Alexander J. 2002. One dose of 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) or 2-amino-3-methylimidazo[4,5-f]quinoline (IQ) induces tumours in Min/+ mice by truncation mutations or LOH in the Apc gene. *Mutation research* **517**: 157-166.

Andreeff M, Goodrich DW, Pardee AB. Differentiation. In: Kufe DW, Pollock RE, Weichselbaum RR, et al., editors. Holland-Frei Cancer Medicine. 6th edition. Hamilton (ON): BC Decker; 2003. Available from: http://www.ncbi.nlm.nih.gov/books/NBK13838/

Araki M, Okada TS. 1977. Differentiation of lens and pigment cells in cultures of neural retinal cells of early chick embryos. *Developmental biology* **60**: 278-286.

Arce L, Yokoyama NN, Waterman ML. 2006. Diversity of LEF/TCF action in development and disease. *Oncogene* **25**: 7492-7504.

Babeu JP, Darsigny M, Lussier CR, Boudreau F. 2009. Hepatocyte nuclear factor 4alpha contributes to an intestinal epithelial phenotype in vitro and plays a partial role in mouse intestinal epithelium differentiation. *American journal of physiology Gastrointestinal and liver physiology* **297**: G124-134.

Barker N. 2014. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature reviews Molecular cell biology* **15**: 19-33.

Bertrand S, Brunet FG, Escriva H, Parmentier G, Laudet V, Robinson-Rechavi M. 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. *Molecular biology and evolution* **21**: 1923-1937.

Biddie SC, John S, Sabo PJ, Thurman RE, Johnson TA, Schiltz RL, Miranda TB, Sung MH, Trump S, Lightman SL et al. 2011. Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. *Molecular cell* **43**: 145-155.

Bottomly D, Kyler SL, McWeeney SK, Yochum GS. 2010. Identification of {beta}-catenin binding regions in colon cancer cells using ChIP-Seq. *Nucleic acids research* **38**: 5735-5745.

Boyle P, Langman JS. 2000. ABC of colorectal cancer: Epidemiology. *Bmj* **321**: 805-808.

Briancon N, Bailly A, Clotman F, Jacquemin P, Lemaigre FP, Weiss MC. 2004. Expression of the alpha7 isoform of hepatocyte nuclear factor (HNF) 4 is activated by HNF6/OC-2 and HNF1 and repressed by HNF4alpha1 in the liver. *The Journal of biological chemistry* **279**: 33398-33408.

Briancon N, Weiss MC. 2006. In vivo role of the HNF4alpha AF-1 activation domain revealed by exon swapping. *The EMBO journal* **25**: 1253-1262.

Brunner E, Peter O, Schweizer L, Basler K. 1997. pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. *Nature* **385**: 829-833.

Cadigan KM, Waterman ML. 2012. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harbor perspectives in biology* **4**.

Cancer Genome Atlas N. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**: 330-337.

Castrop J, van Norren K, Clevers H. 1992. A gene family of HMG-box transcription factors with homology to TCF-1. *Nucleic acids research* **20**: 611.

Catalano MG, Pfeffer U, Raineri M, Ferro P, Curto A, Capuzzi P, Corno F, Berta L, Fortunati N. 2000. Altered expression of androgen-receptor isoforms in human colon-cancer tissues. *International journal of cancer Journal international du cancer* **86**: 325-330.

Chandra V, Huang P, Potluri N, Wu D, Kim Y, Rastinejad F. 2013. Multidomain integration in the structure of the HNF-4alpha nuclear receptor complex. *Nature* **495**: 394-398.

Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CL, Chan C, Dent OF, Clarke SJ, Robertson GR et al. 2012. Src tyrosine kinase phosphorylation of nuclear receptor HNF4alpha correlates with isoform-specific loss of HNF4alpha in human colon cancer. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 2302-2307.

Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC et al. 2011. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12372-12377.

Davuluri RV, Suzuki Y, Sugano S, Plass C, Huang TH. 2008. The functional consequences of alternative promoter use in mammalian genomes. *Trends in genetics : TIG* **24**: 167-177.

De Robertis M, Massi E, Poeta ML, Carotti S, Morini S, Cecchetelli L, Signori E, Fazio VM. 2011. The AOM/DSS murine model for the study of colon carcinogenesis: From pathways to diagnosis and therapy studies. *Journal of carcinogenesis* **10**: 9.

Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R. 1985. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of embryology and experimental morphology* **87**: 27-45.

Duncan SA, Manova K, Chen WS, Hoodless P, Weinstein DC, Bachvarova RF, Darnell JE, Jr. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proceedings of the National Academy of Sciences of the United States of America* **91**: 7598-7602.

Eguchi G, Okada TS. 1973. Differentiation of lens tissue from the progeny of chick retinal pigment cells cultured in vitro: a demonstration of a switch of cell types in clonal cell culture. *Proceedings of the National Academy of Sciences of the United States of America* **70**: 1495-1499.

Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**: 154-156.

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International journal of cancer Journal international du cancer* **127**: 2893-2917.

Frietze S, Wang R, Yao L, Tak YG, Ye Z, Gaddis M, Witt H, Farnham PJ, Jin VX. 2012. Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. *Genome biology* **13**: R52.

Garrison WD, Battle MA, Yang C, Kaestner KH, Sladek FM, Duncan SA. 2006. Hepatocyte nuclear factor 4alpha is essential for embryonic development of the mouse colon. *Gastroenterology* **130**: 1207-1220.

Giangrande PH, Kimbrel EA, Edwards DP, McDonnell DP. 2000. The opposing transcriptional activities of the two isoforms of the human progesterone receptor are due to differential cofactor binding. *Molecular and cellular biology* **20**: 3102-3115.

Giangrande PH, McDonnell DP. 1999. The A and B isoforms of the human progesterone receptor: two functionally different transcription factors encoded by a single gene. *Recent progress in hormone research* **54**: 291-313; discussion 313-294.

Giese K, Cox J, Grosschedl R. 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69**: 185-195.

Giles RH, van Es JH, Clevers H. 2003. Caught up in a Wnt storm: Wnt signaling in cancer. *Biochimica et biophysica acta* **1653**: 1-24.

Gougelet A, Torre C, Veber P, Sartor C, Bachelot L, Denechaud PD, Godard C, Moldes M, Burnol AF, Dubuquoy C et al. 2014. T-cell factor 4 and beta-catenin chromatin occupancies pattern zonal liver metabolism in mice. *Hepatology* **59**: 2344-2357.

Haggar FA, Boushey RP. 2009. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clinics in colon and rectal surgery* **22**: 191-197.

Holewa B, Zapp D, Drewes T, Senkel S, Ryffel GU. 1997. HNF4beta, a new gene of the HNF4 family with distinct activation and expression profiles in oogenesis and embryogenesis of Xenopus laevis. *Molecular and cellular biology* **17**: 687-694.

Jiang G, Lee U, Sladek FM. 1997. Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. *Molecular and cellular biology* **17**: 6546-6554.

Karim BO, Huso DL. 2013. Mouse models for colorectal cancer. *American journal of cancer research* **3**: 240-250.

Kastner P, Krust A, Turcotte B, Stropp U, Tora L, Gronemeyer H, Chambon P. 1990. Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. *The EMBO journal* **9**: 1603-1614.

Keller G. 2005. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes & development* **19**: 1129-1155.

Kidd AR, 3rd, Miskowski JA, Siegfried KR, Sawa H, Kimble J. 2005. A beta-catenin identified by functional rather than sequence criteria and its role in Wnt/MAPK signaling. *Cell* **121**: 761-772.

Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, Clevers H. 1998. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature genetics* **19**: 379-383.

Kucherlapati MH, Lee K, Nguyen AA, Clark AB, Hou H, Jr., Rosulek A, Li H, Yang K, Fan K, Lipkin M et al. 2010. An Msh2 conditional knockout mouse for studying intestinal cancer and testing anticancer agents. *Gastroenterology* **138**: 993-1002 e1001.

Larroux C, Fahey B, Liubicich D, Hinman VF, Gauthier M, Gongora M, Green K, Worheide G, Leys SP, Degnan BM. 2006. Developmental expression of transcription factor genes in a demosponge: insights into the origin of metazoan multicellularity. *Evolution & development* **8**: 150-173.

Larsson SC, Wolk A. 2006. Meat consumption and risk of colorectal cancer: a meta-analysis of prospective studies. *International journal of cancer Journal international du cancer* **119**: 2657-2664.

Lazarevich NL, Cheremnova OA, Varga EV, Ovchinnikov DA, Kudrjavtseva EI, Morozova OV, Fleishman DI, Engelhardt NV, Duncan SA. 2004. Progression of HCC in mice is associated with a downregulation in the expression of hepatocyte nuclear factors. *Hepatology* **39**: 1038-1047.

Lin R, Thompson S, Priess JR. 1995. pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. *Cell* **83**: 599-609.

Love JJ, Li X, Case DA, Giese K, Grosschedl R, Wright PE. 1995. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **376**: 791-795.

Lu XP, Eberhardt NL, Pfahl M. 1993. DNA bending by retinoid X receptor-containing retinoid and thyroid hormone receptor complexes. *Molecular and cellular biology* **13**: 6509-6519.

Lumpkin OJ, Dejardin P, Zimm BH. 1985. Theory of gel electrophoresis of DNA. *Biopolymers* **24**: 1573-1593.

Marcu KB, Bossone SA, Patel AJ. 1992. myc function and regulation. *Annual review of biochemistry* **61**: 809-860.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 7634-7638.

Medema JP, Vermeulen L. 2011. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* **474**: 318-326.

Molenaar M, Roose J, Peterson J, Venanzi S, Clevers H, Destree O. 1998. Differential expression of the HMG box transcription factors XTcf-3 and XLef-1 during early xenopus development. *Mechanisms of development* **75**: 151-154.

Moser AR, Pitot HC, Dove WF. 1990. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* **247**: 322-324.

Mulholland DJ, Dedhar S, Coetzee GA, Nelson CC. 2005. Interaction of nuclear receptors with the Wnt/beta-catenin/Tcf signaling axis: Wnt you like to know? *Endocrine reviews* **26**: 898-915.

Murray-Zmijewski F, Lane DP, Bourdon JC. 2006. p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell death and differentiation* **13**: 962-972.

Narisawa T, Magadia NE, Weisburger JH, Wynder EL. 1974. Promoting effect of bile acids on colon carcinogenesis after intrarectal instillation of N-methyl-N'-nitro-N-nitrosoguanidine in rats. *Journal of the National Cancer Institute* **53**: 1093-1097.

Nateri AS, Spencer-Dene B, Behrens A. 2005. Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. *Nature* **437**: 281-285.

Patapoutian A, Wold BJ, Wagner RA. 1995. Evidence for developmentally programmed transdifferentiation in mouse esophageal muscle. *Science* **270**: 1818-1821.

Pfahl M. 1993. Nuclear receptor/AP-1 interaction. *Endocrine reviews* **14**: 651-658.

Reddy BS, Ohmori T. 1981. Effect of intestinal microflora and dietary fat on 3,2'-dimethyl-4-aminobiphenyl-induced colon carcinogenesis in F344 rats. *Cancer research* **41**: 1363-1367.

Renaud JP, Moras D. 2000. Structural studies on nuclear receptors. *Cellular and molecular life sciences : CMLS* **57**: 1748-1769.

Reya T, Morrison SJ, Clarke MF, Weissman IL. 2001. Stem cells, cancer, and cancer stem cells. *Nature* **414**: 105-111.

Robinson-Rechavi M, Carpentier AS, Duffraisse M, Laudet V. 2001. How many nuclear hormone receptors are there in the human genome? *Trends in genetics : TIG* **17**: 554-556.

Robinson-Rechavi M, Maina CV, Gissendanner CR, Laudet V, Sluder A. 2005. Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. *Journal of molecular evolution* **60**: 577-586.

Roose J, Clevers H. 1999. TCF transcription factors: molecular switches in carcinogenesis. *Biochimica et biophysica acta* **1424**: M23-37.

Ruse MD, Jr., Privalsky ML, Sladek FM. 2002. Competitive cofactor recruitment by orphan receptor hepatocyte nuclear factor 4alpha1: modulation by the F domain. *Molecular and cellular biology* **22**: 1626-1638.

Sancho R, Nateri AS, de Vinuesa AG, Aguilera C, Nye E, Spencer-Dene B, Behrens A. 2009. JNK signalling modulates intestinal homeostasis and tumourigenesis in mice. *The EMBO journal* **28**: 1843-1854.

Schmidt ED, Cramer SJ, Offringa R. 1993. The thyroid hormone receptor interferes with transcriptional activation via the AP-1 complex. *Biochemical and biophysical research communications* **192**: 151-160.

Schwabe JW, Chapman L, Finch JT, Rhodes D. 1993. The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell* **75**: 567-578.

Shen CN, Horb ME, Slack JM, Tosh D. 2003. Transdifferentiation of pancreas to liver. *Mechanisms of development* **120**: 107-116.

Sladek FM. 2011. What are nuclear receptor ligands? *Molecular and cellular endocrinology* **334**: 3-13.

Sladek FM, Ruse MD, Jr., Nepomuceno L, Huang SM, Stallcup MR. 1999. Modulation of transcriptional activation and coactivator interaction by a splicing variation in the F domain of nuclear receptor hepatocyte nuclear factor 4alpha1. *Molecular and cellular biology* **19**: 6509-6522.

Sluder AE, Maina CV. 2001. Nuclear receptors in nematodes: themes and variations. *Trends in genetics : TIG* **17**: 206-213.

Spath GF, Weiss MC. 1997. Hepatocyte nuclear factor 4 expression overcomes repression of the hepatic phenotype in dedifferentiated hepatoma cells. *Molecular and cellular biology* **17**: 1913-1922.

Spath GF, Weiss MC. 1998. Hepatocyte nuclear factor 4 provokes expression of epithelial marker genes, acting as a morphogen in dedifferentiated hepatoma cells. *The Journal of cell biology* **140**: 935-946.

Stairs DB, Nakagawa H, Klein-Szanto A, Mitchell SD, Silberg DG, Tobias JW, Lynch JP, Rustgi AK. 2008. Cdx1 and c-Myc foster the initiation of transdifferentiation of the normal esophageal squamous epithelium toward Barrett's esophagus. *PloS one* **3**: e3534.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C et al. 2006. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. *The Journal of pathology* **208**: 662-672.

Teo R, Mohrlen F, Plickert G, Muller WA, Frank U. 2006. An evolutionary conserved role of Wnt signaling in stem cell fate decision. *Developmental biology* **289**: 91-99.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. 1998. Embryonic stem cell lines derived from human blastocysts. *Science* **282**: 1145-1147.

Torres-Padilla ME, Fougere-Deschatrette C, Weiss MC. 2001. Expression of HNF4alpha isoforms in mouse liver development is regulated by sequential promoter usage and constitutive 3' end splicing. *Mechanisms of development* **109**: 183-193.

Torres-Padilla ME, Sladek FM, Weiss MC. 2002. Developmentally regulated N-terminal variants of the nuclear receptor hepatocyte nuclear factor 4alpha mediate multiple interactions through coactivator and corepressor-histone deacetylase complexes. *The Journal of biological chemistry* **277**: 44677-44687.

Torres-Padilla ME, Weiss MC. 2003. Effects of interactions of hepatocyte nuclear factor 4alpha isoforms with coactivators and corepressors are promoter-specific. *FEBS letters* **539**: 19-23.

Toualbi K, Guller MC, Mauriz JL, Labalette C, Buendia MA, Mauviel A, Bernuau D. 2007. Physical and functional cooperation between AP-1 and beta-catenin for the regulation of TCF-dependent genes. *Oncogene* **26**: 3492-3502.

van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, van der Horn K, Batlle E, Coudreuse D, Haramis AP et al. 2002. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**: 241-250.

Walesky C, Edwards G, Borude P, Gunewardena S, O'Neil M, Yoo B, Apte U. 2013a. Hepatocyte nuclear factor 4 alpha deletion promotes diethylnitrosamine-induced hepatocellular carcinoma in rodents. *Hepatology* **57**: 2480-2490.

Walesky C, Gunewardena S, Terwilliger EF, Edwards G, Borude P, Apte U. 2013b. Hepatocyte-specific deletion of hepatocyte nuclear factor-4alpha in adult mice results in increased hepatocyte proliferation. *American journal of physiology Gastrointestinal and liver physiology* **304**: G26-37.

Webb P, Nguyen P, Valentine C, Lopez GN, Kwok GR, McInerney E, Katzenellenbogen BS, Enmark E, Gustafsson JA, Nilsson S et al. 1999. The estrogen receptor enhances AP-1 activity by two distinct mechanisms with different requirements for receptor transactivation functions. *Molecular endocrinology* **13**: 1672-1685.

Wells J, Farnham PJ. 2002. Characterizing transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. *Methods* **26**: 48-56.

Weltmeier F, Borlak J. 2011. A high resolution genome-wide scan of HNF4alpha recognition sites infers a regulatory gene network in colon cancer. *PloS one* **6**: e21667.

Yang J, Kong X, Martins-Santos ME, Aleman G, Chaco E, Liu GE, Wu SY, Samols D, Hakimi P, Chiang CM et al. 2009. Activation of SIRT1 by resveratrol represses transcription of the gene for the cytosolic form of phosphoenolpyruvate carboxykinase (GTP) by deacetylating hepatic nuclear factor 4alpha. *The Journal of biological chemistry* **284**: 27042-27053.

Yang M, Li SN, Anjum KM, Gui LX, Zhu SS, Liu J, Chen JK, Liu QF, Ye GD, Wang WJ et al. 2013. A double-negative feedback loop between Wnt-beta-catenin signaling and HNF4alpha regulates epithelial-mesenchymal transition in hepatocellular carcinoma. *Journal of cell science* **126**: 5692-5703.

Yao L, Tak YG, Berman BP, Farnham PJ. 2014. Functional annotation of colon cancer risk SNPs. *Nature communications* **5**: 5114.

Yu X, Kensler T. 2005. Nrf2 as a target for cancer chemoprevention. *Mutation research* **591**: 93-102.

Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S et al. 2014. Proteogenomic characterization of human colon and rectal cancer. *Nature*.