Title

Differential Effects of Hepatocyte Nuclear Factor 4α Isoforms on Tumor Growth and T-Cell Factor 4/AP-1 Interactions in Human Colorectal Cancer Cells

Journal

Molecular and Cellular Biology, 35(20)

ISSN

0270-7306

Authors

Vuong, Linh M
Chellappa, Karthikeyani
Dhahbi, Joseph M
et al.

Publication Date

2015-10-01

DOI

10.1128/mcb.00030-15

Peer reviewed

MCB
Journals.ASM.org

Differential Effects of Hepatocyte Nuclear Factor 4α Isoforms on Tumor Growth and T-Cell Factor 4/AP-1 Interactions in Human Colorectal Cancer Cells

Linh M. Vuong,[a] Karthikeyani Chellappa,[a]* Joseph M. Dhahbi,[b] Jonathan R. Deans,[a] Bin Fang,[a]* Eugene Bolotin,[a]* Nina V. Titova,[a] Nate P. Hoverter,[c]* Stephen R. Spindler,[b] Marian L. Waterman,[c] Frances M. Sladek[a]

Department of Cell Biology and Neuroscience[a] and Department of Biochemistry,[b] University of California, Riverside, Riverside, California, USA; Department of Microbiology and Molecular Genetics, University of California, Irvine, Irvine, California, USA[c]

The nuclear receptor hepatocyte nuclear factor 4α (HNF4α) is tumor suppressive in the liver but amplified in colon cancer, suggesting that it also might be oncogenic. To investigate whether this discrepancy is due to different HNF4α isoforms derived from its two promoters (P1 and P2), we generated Tet-On-inducible human colon cancer (HCT116) cell lines that express either the P1-driven (HNF4α2) or P2-driven (HNF4α8) isoform and analyzed them for tumor growth and global changes in gene expression (transcriptome sequencing [RNA-seq] and chromatin immunoprecipitation sequencing [ChIP-seq]). The results show that while HNF4α2 acts as a tumor suppressor in the HCT116 tumor xenograft model, HNF4α8 does not. Each isoform regulates the expression of distinct sets of genes and recruits, colocalizes, and competes in a distinct fashion with the Wnt/β-catenin mediator T-cell factor 4 (TCF4) at CTTTG motifs as well as at AP-1 motifs (TGAXTCA). Protein binding microarrays (PBMs) show that HNF4α and TCF4 share some but not all binding motifs and that single nucleotide polymorphisms (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 interactions with the Wnt/β-catenin/TCF4 and AP-1 pathways.

Hepatocyte nuclear factor 4α (HNF4α) (NR2A1) is a highly conserved member of the nuclear receptor superfamily found in all metazoans (1, 2) and best known as a master regulator of tissue-specific gene expression in the adult liver (3–5). HNF4α binds specific DNA sequences as a homodimer and regulates expression of genes involved in metabolism, homeostasis, differentiation, and immune response (4, 6, 7). It also plays a role in early development (8), as well as in the adult kidney, pancreas, and gut (9–15). 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 (6, 16). Recently, HNF4α was shown to be involved in colon cancer, but its precise role remains elusive (11, 12, 17, 18).

Several splice variants of HNF4α are generated via two alternative promoters (proximal promoter P1 and distal promoter P2) and two distinct 3′ splicing events (19). P1-driven HNF4α1/2, which includes the full-length N-terminal A/B domain, was cloned from adult rat liver (1), while the P2-driven HNF4α7/8 with a distinct N-terminal domain was cloned from an embryonic cell line (20) (see Fig. 1A). HNF4α2 and HNF4α8 are the predominant forms in most tissues (21). The promoter-driven HNF4α isoforms exhibit tissue-specific expression patterns: the P1-driven HNF4α1/2 is expressed in the fetal and adult liver and kidney, whereas the P2-driven HNF4α7/8 is expressed in the fetal liver and the adult stomach and pancreas; both isoforms are expressed in the large and small intestines (18, 19, 22, 23). The HNF4α gene structure, promoter sequences, and expression patterns are highly conserved between humans and mice (19), suggesting that P1- and P2-driven HNF4α play important 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 (22).

P1-HNF4α acts as a tumor suppressor in the liver (24), inhibiting hepatocyte proliferation and inflammation (25–27). Several key players in proliferation, including p53, c-Myc, T-cell factor 4 (TCF4 [*TCF7L2*]), lymphoid enhancer factor 1 (LEF1 [*LEF1*]), and cyclin D1, have all been shown to physically interact with and antagonize P1-HNF4α (12, 28–33).

The role of P2-HNF4α in cancer is less clear. Immunohistochemical staining for HNF4α in liver, colon, and stomach cancers

Received 13 January 2015 **Returned for modification** 4 February 2015
Accepted 7 July 2015

Accepted manuscript posted online 3 August 2015

Citation Vuong LM, Chellappa K, Dhahbi JM, Deans JR, Fang B, Bolotin E, Titova NV, Hoverter NP, Spindler SR, Waterman ML, Sladek FM. 2015. Differential effects of hepatocyte nuclear factor 4α isoforms on tumor growth and T-cell factor 4/AP-1 interactions in human colorectal cancer cells. Mol Cell Biol 35:3471–3490. doi:10.1128/MCB.00030-15.

Address correspondence to Frances M. Sladek, frances.sladek@ucr.edu.

* Present address: Karthikeyani Chellappa, Institute for Diabetes, Obesity, and Metabolism and Department of Physiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; Bin Fang, Smilow Center for Translational Research, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA; Eugene Bolotin, Illumina, Santa Clara, California, USA; Nate P. Hoverter, Purigen Biosystems, Pleasanton, California, USA.

L.M.V. and K.C. contributed equally to this article.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/MCB.00030-15.

doi:10.1128/MCB.00030-15

showed that there is a dysregulation of the HNF4α isoforms, with P2-HNF4α typically being expressed at higher levels than P1-HNF4α (18, 34, 35). Additionally, in a large cohort of 450 human colon cancer samples, we found a loss of nuclear P1-HNF4α, which we attributed to Src tyrosine kinase preferentially phosphorylating P1- but not P2-driven HNF4α (17). The Cancer Genome Atlas (TCGA) recently identified the region encompassing *HNF4A* (20q13.12) as being one of several amplified loci in over 255 human colon cancers (36) and found an overexpression of the HNF4α protein in a subset of those samples (37). HNF4α has also been shown to exhibit oncogenic activity in gastric cancer (38). While these findings suggest that the HNF4α gene may act as an oncogene, as well as a tumor suppressor, the relative contributions of the different HNF4α isoforms were not determined.

While HNF4α, especially the P1-HNF4α isoform, is known to drive differentiation, the Wnt/β-catenin/TCF signaling pathway is well known to promote cell proliferation. There are an increasing number of reports that indicate a potential cross talk between HNF4α and the Wnt pathway in liver zonation, hepatocellular carcinoma (HCC) development, and colorectal cancer: Physical interactions between HNF4α and TCF4 have been reported in soluble nuclear extracts (NE) as well as in chromatin-bound fractions on isolated promoters (12, 39–42). LEF1/TCF binding motifs have also been found enriched in HNF4α chromatin immunoprecipitation sequencing (ChIP-seq) peaks and vice versa (40, 42–45), suggesting a potential coregulation by these two transcription factors (TFs). The nature of that coregulation, however, is not yet clear.

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 that expresses either P1-HNF4α2 or P2-HNF4α8 under the control of doxycycline (DOX). Xenograft assays indicate that HNF4α2 is more effective at suppressing tumor growth than HNF4α8 *in vivo*. Transcriptome sequencing (RNA-seq) and ChIP-seq analyses revealed differences in gene expression and binding locations between HNF4α2 and HNF4α8, while TCF4 ChIP-seq indicated that HNF4α recruits, colocalizes, and competes with TCF4 on a substantial number of promoters. Our results also indicate that the HNF4α isoforms interact with TCF4, as well as the AP-1 complex, in a differential fashion. Finally, common and unique TCF4 and HNF4α binding motifs were identified using protein binding microarrays (PBMs), which also showed that single nucleotide polymorphisms (SNPs) in TCF4/HNF4α binding sites can affect DNA binding. Overall, our results indicate that there are specific, potentially important functional differences in the HNF4α isoforms, some of which involve distinct interactions with the Wnt/β-catenin/TCF and AP-1 pathways.

MATERIALS AND METHODS

Plasmid constructs. The full-length human HNF4α2 (NM_000457) and HNF4α8 (NM_175914) cDNAs in pcDNA3.1 were gifts from Christophe Rachez at Pasteur Institute, Paris, France (14, 46). The human ApoB.-85.-47.E4.Luc luciferase reporter construct and Flag-dnTCF1E1 (Flag.dnTCF1) and Flag-TCF4E2 (Flag.TCF4) expression vectors have been described previously (29, 47).

The doxycycline (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 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 (with restriction sites underlined): 5′-HNF4α2_Koz.EcoRI, 5′-GGAATTCCCCACCATGGATATGGCC-3′; 5′-HNF4α8_Koz.EcoRI, 5′-GGAATTCCCCACCATGGTCAGCGTG-3′; and 3′-N1C465.BamHI, 5′-GCGGGATCCCGCTAGATAACTTCCTGCTT-3′.

All oligonucleotides were synthesized by Integrated DNA Technologies (IDT).

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 were gifts from Chee-Gee Liew (48, 49).

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 medium (Iwakata and Grace modification, with L-glutamine) (Corning Cellgro catalog no. 10-050-CV) supplemented with 10% fetal bovine serum (FBS) (BenchMark catalog no. 100-106) and 100 U/ml penicillin-streptomycin (1% P/S). Cells were passaged every third or fourth day at 85 to 95% confluence. HEK 293T (ATCC, CRL-11268) and COS-7 (ATCC, CRL-1651) cells were cultured in DMEM (Dulbecco's modified Eagle's medium with 4.5 g/liter glucose, L-glutamine, and pyruvate) supplemented with 10% FBS or bovine calf serum (BCS), 100 U/ml nonessential amino acids (1% NEAA), and 1% P/S. All cell lines were maintained at 37°C and 5% CO_2.

To generate the stable lines, HCT116 cells were seeded at 3×10^6 cells per well in a 6-well plate and transfected 24 h later with 1 μg of linearized pCAG-rtTA 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. Puromycin-resistant colonies were screened for DOX inducibility 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-expressing Tet-On-inducible 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 modified McCoy's 5A medium supplemented with Tet-free 10% fetal bovine serum (FBS), 1% P/S, and 0.53 μg/ml puromycin. The HCT116rtTA HNF4α2- and HNF4α8-expressing lines were maintained in a similar fashion with the addition of 70 μg/ml G418.

Migration and invasion assay. Tet-On-inducible HCT116 (PL, HNF4α2, or HNF4α8) clones were seeded at 1.5 to 2.3×10^6 cells in a 100-mm-diameter plate and 24 h later treated with or without 0.5 μg/ml of 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), and transferred (5×10^4 cells) to the upper chamber of an invasion or migration Transwell plate (BD Biocoat). McCoy 5A medium supplemented with 20% FBS was added to the lower chamber. Forty-eight hours later, the top chambers were removed, stained with hematoxylin, and viewed under a microscope. Images were taken at ×20 magnification and printed out for visual 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. Tet-On-inducible HCT116 cells (PL, HNF4α2, or HNF4α8) were seeded in 150-mm plates. The following day, the cells were trypsinized and subcutaneously injected (3×10^6 cells) into the flank of ~8-week-old athymic nude male mice (National Cancer Institute, strain 01B74). 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 Laboratories). 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 high-concentration BD Matrigel matrix (BD Biosciences catalog no. 354248) was added to the cells at a 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).

IB analyses. For immunoblot (IB) analyses, protein extracts were separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene difluoride (PVDF) membrane (Immobilon; Millipore) as previously described (28). Signals were detected using the SuperSignal West Dura extended-duration substrate kit (Thermo Fisher Scientific). A Bradford assay (Bio-Rad) was used to measure the protein concentration: 20 to 60 μg of whole-cell lysates (WCE) (29) or 20 μg nuclear extracts (NE) (50) was loaded per lane. Coomassie staining of the blot verified equal loading of protein. The primary antibodies (Abs) were mouse monoclonal anti-HNF4α P1/P2 (R&D Systems, catalog no. PP-H1415-00) and affinity-purified anti-α445 (1), which recognize the C terminus of both the P1- and P2-HNF4α isoforms; monoclonal anti-HNF4α P1 and P2 (catalog no. PP-K9218-00 and Cat PP-H6939-00, respectively; R&D Systems), which recognize the different N termini of HNF4α; and anti-Flag (M2; Sigma-Aldrich), and anti-TCF7L2 (Millipore catalog no. 6H5-3). The secondary antibodies were horseradish peroxidase (HRP)-conjugated goat anti-rabbit (GαR-HRP) or goat anti-mouse (GαM-HRP) Abs from Jackson ImmunoResearch Laboratories.

HNF4α and TCF1/4 PBMs. Protein binding microarrays (PBMs) were carried out essentially as previously described (7, 51). In Fig. 6A, the HNF4α-centric PBM3 design described by Bolotin et al. (105) was used with NE of human HNF4α2 or dominant-negative TCF1 (dnTCF1) from transfected COS-7 cells. In Fig. 7, another custom-designed array was ordered from Agilent (SurePrint G3 Custom GE 1X1M), which contained an ~60-base oligonucleotide corresponding to sequences within 100 bp of the center of published HNF4α ChIP-seq peaks from HepG2 and CaCo2 cells (52, 53). A total of ~125,000 loci, including SNP alleles from dbSNP version 132, were spotted in quadruplicate (~125,000 loci × ~2 alleles × 4 replicates = 1 million spots of DNA) on the slide as single-stranded DNA. The DNA was made double stranded using a primer to a common linker sequence (5′-TCGACCTCTACTCTAATCTCGCTAGC-3′), deoxynucleoside triphosphates (dNTPs) (GE Healthcare), and Thermo Sequenase (Affymetrix catalog no. 78500). The binding reactions were carried out with ~6 μg of human HNF4α2 or HNF4α8 in NE from transfected COS-7 cells diluted 1:10 in desalting buffer (20 mM HEPES [pH 7.8], 60 mM KCl, 8 mM EDTA, 8 mM EGTA) and processed through a 30-kDa-cutoff column (Amicon catalog no. UFC503096) to a final concentration of 110 mM KCl and then applied to the arrays in PBM binding buffer (20 mM HEPES [pH 7.8], 110 mM KCl, 8 mM EDTA, 8 mM EGTA, 0.1% Tween 20 plus 20 μg salmon sperm DNA). WCE containing Flag.TCF4 (~600 ng) were applied directly to the array diluted in buffer (35 mM Tris-HCl [pH 7.5], 25 mM KCl, 60 mM NaCl, 0.75 mM $MgCl_2$, 2.5 mM β-mercaptoethanol, 0.25% NP-40 plus 20 μg salmon sperm DNA). Arrays were incubated for 100 min at room temperature (RT), after being washed three times for 3 min each with phosphate-buffered saline (PBS) plus 0.1% Tween 20. The bound TFs then were detected using anti-HNF4α P1/P2 or anti-Flag M2 Ab (Sigma-Aldrich catalog no. F1804) for dnTCF1 and TCF4 at a 1:100 dilution in 2% low-fat milk plus 0.1% Tween 20 in PBS overnight (ON) at RT, followed by a conjugated secondary Ab (GαM IgG [H+L] DyLight 550; Pierce catalog no. 84540) diluted 1:50 as described above for 90 min and quantification using an Agilent G2565CA microarray scanner at the UCLA DNA Microarray Core. Extraction and normalization of the data were performed as described previously (7). Position weight matrices (PWMs) were generated using SeqLogo (54) and Weblogo v2.8.2 (55).

HNF4α and TCF reporter constructs and assay. Luciferase reporter constructs containing one of three HNF4α/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 are underlined; see Fig. 6A for motif sequences.) HEK 293T cells were seeded at 1.6×10^5 to 2.0×10^5 cells per well of a 12-well plate and 24 h later transfected with 80 ng of HNF4α2 or dnTCF1 expression vector, 0.5 μg of reporter, and 0.1 μg of a β-galactosidase-expressing cytomegalovirus construct, CMV.β-gal, using Lipofectamine 2000. The following day, cells were harvested with lysis buffer (25 mM glycylglycine [pH 7.8], 15 mM $MgSO_4$, 4 mM EGTA, 1% Triton X-100), and luciferase and β-galactosidase activities were measured as previously described (28). All transfections were performed in triplicate, normalized to β-galactosidase, and performed at least four times.

HNF4α and TCF fluorescent gel shift. Double-stranded oligonucleotides (double-stranded DNA [dsDNA]) (5 μg) with 5′-adenine overhangs were fluorescently labeled using a 5-fold molar excess of Cy3-dUTP (GE Healthcare) and 5 U of Klenow fragment (New England BioLabs) in a 100-μl reaction mixture. Unincorporated label was removed using Mini Quick Spin DNA columns (Roche). The shift probes contained the motifs shown in Fig. 6A (5′-AAAACGCGC[motif sequence]GCCTA-3′). NE were prepared from COS-7 cells transfected with ~12 μg of Flag.dnTCF1 or HNF4α2 via $CaPO_4$ precipitation as previously described (50). 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) (56) and the carboxy-terminal Flag fusion protein Flag-BAP (Sigma-Aldrich).

Gel shifts were performed using a 10% nondenaturing polyacrylamide gel as previously described (50). Briefly, each shift reaction mixture contained: 1.5 μl 50 mM EDTA, 1 to 2 μl purified, labeled probe at 0.3 to 5.0 ng/μl, 4.0 μl 5× shift buffer, and 1.5 μl poly(dI-dC) or sonicated salmon sperm DNA at 1 μg/μl. NE (5 to 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 the volume up to 20 μl and total protein to 10 μg. After 20 to 30 min at RT, 6 μl of the reaction mixture was loaded per lane, and the gel was run at a constant current of 12 mA for 45 to 60 min. A Typhoon 9410 imager was used to visualize the bands on the gel.

Transcriptome sequencing (RNA-seq) analysis. Tet-On-inducible HCT116 clones (PL, HNF4α2, and HNF4α8) were seeded at ~5.5×10^5 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 minikit (Qiagen) was used to extract and purify total RNA; 4 μg of each RNA sample was used to generate a poly(A)$^+$ RNA library using the TruSeq RNA sample prep v2 kit (Illumina catalog no. RS-122-2001). Libraries were submitted for 50-bp paired-end 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 triplicate) were multiplexed and sequenced in two lanes, each of which yielded ~442 million reads (~42 million reads per sample).

Paired-end sequencing reads were aligned to the human reference genome (GRCh37/hg19) with Tophat v1.2 (57) using the default parameters, with the exception of allowing up to 10 alignments to the reference genome for a given read instead of the default value of 20. The data aligned by Tophat were processed by Cufflinks (58) to assemble transcripts and to measure their relative abundance in fragments per kilobase of exon per million fragments mapped (FPKM). 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. 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. Principle component analysis showed that all triplicates clustered together within their treatment group (data not shown). The following criteria were used to select differentially expressed genes: (i) a fold change (FC) of at least 1.5 or more, (ii) at least two replicates with an FPKM of ≥5, and (iii) the triplicates for a given condition with a coefficient of variation (CV) of ≤0.5.

ChIP-seq analysis. Tet-On-inducible HNF4α2- and HNF4α8-expressing HCT116 clones were seeded at ~8×10^6 cells per 150-mm-diameter plate and 6 h later treated with 0 or 0.3 μg/ml of DOX. Twenty-four hours after induction, cells were harvested as previously described (30, 59), with minor modifications: the cells were fixed with 1% formaldehyde in PBS for 10 min at RT, and the cross-linking was stopped with 0.125 M glycine in PBS for 10 min at RT. All subsequent steps were performed at 4°C, using ice-cold buffers. Cells were scraped in PBS plus 1 mM phenylmethylsulfonyl fluoride (PMSF) and dithiothreitol (DTT) and centrifuged at 6,000 × *g* 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 $MgCl_2$) plus 1 mM PMSF and DTT for 10 min. The nuclei were pelleted and resuspended in 0.34 ml nuclei lysis buffer (50 mM Tris-HCl [pH 8.0], 10 mM EDTA, 1% Triton X-100) plus 1 mM PMSF and DTT and 2 μg/ml leupeptin and aprotinin. The samples were sonicated using a model 500 sonic dismembrator (Fisher Scientific) to obtain DNA fragments of about 200 to 500 bp and then diluted 1:1 in immunoprecipitation (IP) dilution buffer (20 mM Tris-HCl [pH 8.0], 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton X-100) and precleared with 20 μl of packed protein G-agarose beads (Pierce) that were preblocked with 1 μg/μl BSA (fraction V; Thermo Fisher Scientific) for 30 min. The lysates (~1×10^7 to 3×10^7 cell equivalents per IP) were nutated for 2 h with one of the following Abs: 3 to 5 μg of affinity-purified anti-HNF4α (α-445) (1), anti-TCF7L2 (TCF4) (Millipore catalog no. 6H5-3), or mixed equal amounts of mouse and rabbit IgG controls (Millipore catalog no. 12-371 and Santa Cruz catalog no. sc-2027, respectively). The preblocked protein G beads (30- to 40-μl slurry in 1:1 IP dilution buffer) were added, and the samples were nutated ON at 4°C and then washed with three sequential buffers for 5 min each at RT: TSE I (20 mM Tris-HCl [pH 8.0], 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), TSE II as for TSE I but with 500 mM NaCl, and TSE III (10 mM Tris-HCl [pH 8.0], 0.25 mM LiCl, 1 mM EDTA, 1% NP-40, 1% deoxycholate). At the final wash, the IP sample was washed two times with 1× TE for 5 min at RT. The precipitated material was eluted with 150 μl IP elution buffer (0.1 M $NaHCO_3$, 1% SDS) at 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°C for 4 to 5 h to reverse the cross-links. The DNA was precipitated with 1 ml 100% ethanol ON at −20°C, washed with 70% ethanol, and resuspended in 100 μl TE. RNA and protein digestions were performed by addition of 1 μl of 10 μg/μl RNase A (Roche) and incubation at RT for 25 min followed by 11 μl of 10× proteinase K buffer (100 mM Tris-HCl [pH 8.0], 500 mM NaCl, 50 mM EDTA) and 1 μg proteinase K (IBI Scientific) for 1 h at 55°C. The GeneJET PCR purification kit (Thermo Fisher Scientific) was used to purify the DNA, and a Qubit fluorometer in the UCR Genomics Core was used to measure the DNA concentration; 5 to 20 ng of ChIP material per condition (from one 150-mm plate) was used to generate libraries using a BIOO Scientific ChIP-seq DNA library kit (NEXTflex ChIP-seq kit catalog no. 5143-02 and barcodes catalog no. 514122). Libraries were submitted for 50-bp single-end Illumina sequencing as described above. Three different IPs (TCF4 with DOX [TCF4+DOX], TCF4 without DOX [TCF4−DOX], and HNF4α with DOX [HNF4α+DOX]) for each inducible HCT116 line (HNF4α2 and HNF4α8) were performed in duplicate (12 samples total). The reads from duplicate samples were pooled and normalized to one input sample per cell line.

Sequencing reads (~30 million per condition) were preprocessed and mapped to the human reference genome (GRCh37/hg19) with Bowtie v1.8 (60) and subsequently analyzed with MACS v2 (61). The callpeak function was used to generate bedGraph files using the default minimum FDR cutoff of 0.05 (see Fig. 8A), which were further analyzed with the bdgdiff function to identify differential binding events between conditions using a $\log_{10}$ likelihood ratio cutoff of 100 instead of the default value of 1,000 (Fig. 8B). The bdgcmp function was used to deduct 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 (62), to retrieve Ensembl genes that are closest to the transcription start site (TSS [+1]).

ChIP-IB analysis. Tet-On-inducible HCT116 HNF4α2- and HNF4α8-expressing lines were maintained as described above with the addition of 1% minimal essential medium with nonessential amino acids (MEM NEAA) (Corning Cellgro catalog no. 25-025-Cl). Cells were processed as described above for ChIP. Samples were immunoprecipitated with the α445 Ab and incubated with beads ON at 4°C. The samples were washed three times with cold radioimmunoprecipitation assay (RIPA) buffer (15 mM Tris-Cl, 1% NP-40, 0.7% deoxycholate, 1 mM EDTA, 150 mM NaCl) and once with 200 μl DNase buffer (40 mM Tris-HCl, 1 mM $CaCl_2$, 20 mM NaCL, 6 mM $MgCl_2$). Samples were incubated in 100 μl of DNase buffer with or without 24 μg DNase I (Sigma-Aldrich catalog no. D5319) at 30°C for 20 min with agitation every 5 min. After DNA digestion, samples were rinsed three times with RT RIPA buffer; at the final wash (0.5 ml), three-quarters of the sample was eluted with 6 μl of SDS plus β-mercaptoethanol and 24 μl of ddH_2O, boiled for 5 min, spun at 18,000 × *g* for 5 min at RT, and analyzed by IB as described above. The following primary Abs were used: anti-TCF7L2 (TCF4), anti-P1/P2, and anti-FRA1 (FOSL1) (Santa Cruz catalog no. sc-183x). These were followed by the secondary Abs: mouse and rabbit IgG-HRP Trueblot (Rockland catalog no. 18-881731 and 18-8816-31). The other quarter of the sample was eluted with 50 μl IP elution buffer, boiled for 1 min, and nutated for 20 min at RT. NaCl (0.3 M final concentration) was added to the eluted sample, and the mixture was incubated at 65°C for ~5 h. Proteins were digested as described above. DNA was purified using a ChIP DNA Clean & Concentrator (Zymo Research catalog no. D5205), and the DNA concentration was measured on a Qubit fluorometer. The percentage of FOSL1 pulled down by HNF4α (α445) or control IgG was quantified using Image Lab software (Bio-Rad, version 5.1, build 8). The percentage of pulldown was determined using the FOSL1 signal from the appropriate input lane.

Re-ChIP-PCR. After the first ChIP with the α445 Ab, the sample was eluted with 100 μl of elution buffer, transferred to a new tube, diluted 1:3 with IP dilution buffer, and then immunoprecipitated with FRA1 or TCF7L2 Ab as described above. Primers spanning the promoter regions of the human *SIAH2* (forward, 5′-GGGAGAGTGGATAGGTCTGC-3′; reverse, 5′-AGTAGGTGGGCGAATGAGAC-3′) and *ACSL5* (forward, 5′-ACACTGCTTCTTCTTACCCCA-3′; reverse, 5′-CAGACATTGGCCAG TTGAGC-3′) genes were used in the final PCR to yield product sizes of 380 bp and 317 bp, respectively.

Bioinformatics and statistical analyses. VENNY (63) was used to compare lists of genes in the different cell lines. Gene ontology (GO) analysis was performed using DAVID Bioinformatics Resources 6.7 (64). Cisgenome (65) was used to identify HNF4α and TCF4 target genes and to 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 (66) and Integrated Genome Viewer (IGV) (67) for visualization. The R program was used to identify overlapping HNF4α2 and HNF4α8 peaks and DNase-seq peaks, defined as at least 13 nucleotides (nt), and to generate scatter plots and heat maps (68, 69). MEME analysis was used to mine motifs (70). The HNF4 Binding Site Scanner website (http://nrmotif.ucr.edu/fuzzhtmlform2.html) was used to identify HNF4 binding sites in the ChIP-seq peaks using the Support Vector Machine (SVM) option (7); any predicted sites with scores of ≥1.5 were considered to be potential binders for HNF4. The following data from the Encyclopedia of DNA Elements

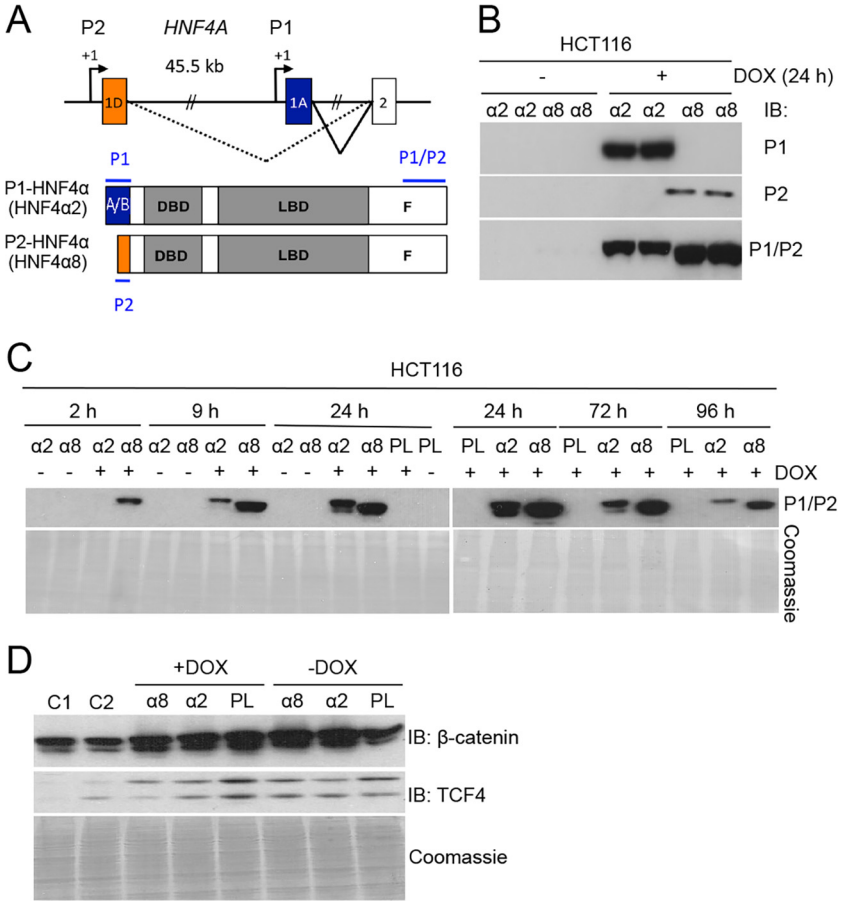


FIG 1 Establishment of stable inducible HCT116 lines expressing human HNF4α2 or HNF4α8. (A) Schematic of the human *HNF4A* gene and the isoforms generated by its two promoters (P1 and P2). Epitopes to the P1, P2, and P1/P2 antibodies (Abs) are indicated. DBD, DNA binding domain; LBD, ligand binding domain. The P1-HNF4α isoforms contain a full-length A/B domain (blue); P2 isoforms contain a truncated A/B domain (orange). (B and C) IB with Abs described in panel A of NE (B) and WCE (C) from inducible HCT116 lines expressing HNF4α2 (α2) or HNF4α8 (α8) or the parental (PL) line treated with 0.3 μg/ml DOX or not treated with DOX for the indicated times. (D) IB for β-catenin and TCF4 (two splice variants) in the indicated NEs prepared 24 h after the addition of DOX. Controls 1 and 2 (C1 and C2, respectively) are NEs from HEK293T and HepG2 cells, respectively. Coomassie staining verified equal loading.

(ENCODE) Project were used: DNase-seq (the J. A. Stamatoyannopoulos lab from the University of Washington; UCSC accession no. wgEncodeEH001162 [GSM736600 and GSM736493]) and ChIPseq TCF7L2 (TCF4) (the P. J. Farnham lab from the University of Southern California; UCSC accession no. WgEncodeEH000629 [GSM782123]), FOSL1 and JUND (R. M. Myers lab from HudsonAlpha Institute for Biotechnology, UCSC accession no. wgEncodeEH003246 [GSM1010756] and UCSC accession no. wgEncodeEH003216 [GSM1010847]); all sequencing of which was performed with the HCT116 cell line (71, 72). Line and bar graphs are plotted as means ± standard errors of the means (SEM) or standard deviations (SD) as indicated. Student's *t* test was used to calculate *P* values: $P < 0.05$ was considered significant.

Microarray data accession numbers. All RNA-seq and ChIP-seq data have been submitted to GEO under accession no. GSE62889 and GSE62890, respectively. All PBM data have been uploaded onto The Nuclear Receptor DNA Binding Project website (http://nrdbs.ucr.edu; see PBM Data and aaSNPs columns in Results tab [aaSNP data are also searchable at http://nrmotif.ucr.edu/aaSNP/]).

RESULTS

Generation and characterization of isoform-specific HNF4α-expressing inducible cell lines. To distinguish the function of the HNF4α promoter-driven isoforms (Fig. 1A) in human colon cancer, we generated Tet-On-inducible HCT116 cell lines that express either HNF4α2 or HNF4α8. The parental line, HCT116, is poorly differentiated, considered to have stem cell-like properties (73), and does not express endogenous HNF4α (Fig. 1B). It does, however, contain a mutant allele of the β-catenin gene (*CTNNB1*) resulting in a constitutively active Wnt/β-catenin pathway (74); TCF4 is the primary member of TCF family expressed in HCT116 cells (75) (Fig. 1D). The chromatin occupancies of TCF4 and β-catenin in HCT116 cells have also been characterized by the ENCODE project (44, 76). All told, HCT116 cells are a good model to examine the role of the HNF4α isoforms in colon cancer and their interaction with the Wnt pathway.

Induction and appropriate expression of the HNF4α isoforms were verified by immunoblotting (IB) with isoform-specific antibodies (Ab) (Fig. 1A and B). HNF4α8 was detected as early as 2 h after DOX induction, and the HNF4α2 protein was typically expressed at somewhat lower levels than HNF4α8 (Fig. 1C). Expression of both HNF4α2 and HNF4α8 peaked at 24 h (Fig. 1C). The expression of HNF4α did not significantly affect β-catenin or TCF4 expression (Fig. 1D).

HNF4α2 is more tumor suppressive than HNF4α8. To deter-

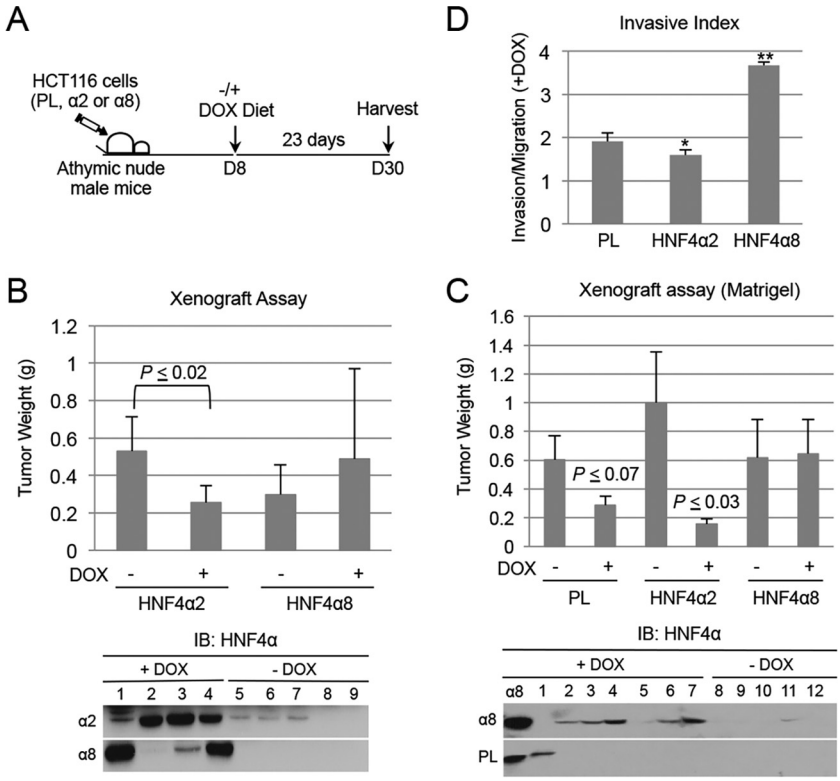


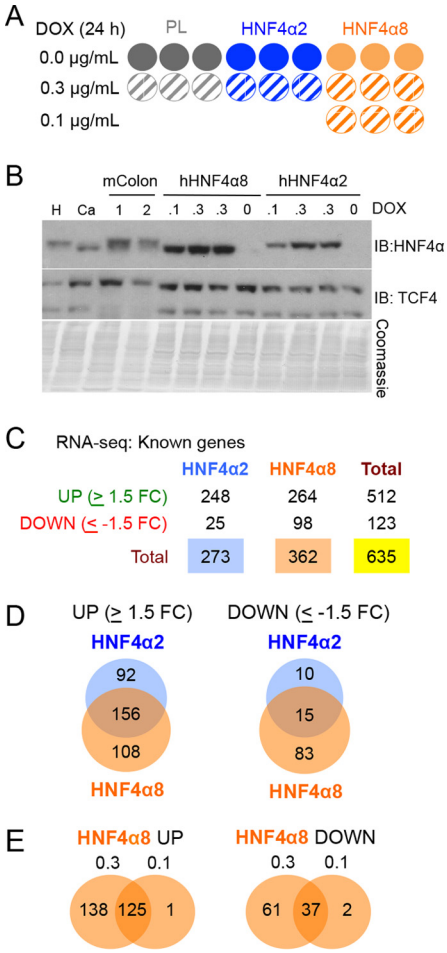
FIG 2 HNF4α2 is more tumor suppressive than HNF4α8 in colorectal cancer cells. (A) Schematic of the xenograft assay. After injection of the indicated HCT116 lines (PL, parental; α2, HNF4α2 expressing; α8, HNF4α8 expressing), half of the immunocompromised mice were switched to a DOX diet on day 8. (B and C) Tumor weight at the time of harvest from cells given an injection without (B) or with (C) Matrigel. Error bars are means ± SEM from each condition. (B) HNF4α2, $n = 5$, and HNF4α8, $n = 4$, each without (−) DOX and with (+) DOX; (C) parental (PL), $n = 8$ or 9, HNF4α8, $n = 7$ to 9, and HNF4α2, $n = 10$ for each condition. (B and C, bottom) HNF4α IB (P1/P2 Ab) of WCE of individual tumors. α8, HCT116 HNF4α8 line 24 h after DOX. Coomassie staining verified equal loading (not shown). (D) Invasive index of the parental (PL) and HNF4α2- or HNF4α8-expressing lines in the presence of DOX. Bars are means ± SD. *, $P \leq 0.04$, and **, $P \leq 0.00001$, versus PL.

mine the effect of the HNF4α isoforms on tumor growth, we subcutaneously injected the HCT116 HNF4α2- and HNF4α8-expressing cell lines into immunocompromised mice and allowed the tumors to develop for 8 days before induction of expression of HNF4α with a DOX diet (Fig. 2A). Tumor growth was monitored for another ~23 days, at which point the mice were sacrificed. The HNF4α2-expressing line in the presence of DOX resulted in a significantly decreased tumor weight, while the HNF4α8-expressing line did not (Fig. 2B); HNF4α2 and HNF4α8 were appropriately expressed in the majority of tumors (Fig. 2B, bottom).

To enhance tumor growth, we repeated the xenograft assay with Matrigel. The HNF4α2-expressing line again resulted in statistically smaller tumors ($P \leq 0.03$) in the presence of DOX (Fig. 2C, top). In contrast, there was no difference in tumor weights with or without DOX in the HNF4α8-expressing line, despite the expression of HNF4α8 in the DOX-treated tumors (Fig. 2C, bottom). The parental line lacking the HNF4α transgene (PL) also showed somewhat smaller, although not statistically significant ($P \leq 0.07$), tumors in the presence of DOX. Finally, the HNF4α2-expressing line displayed a lower invasive index and the HNF4α8-expressing line a higher one in an *in vitro* invasion/migration assay (Fig. 2D). These results indicate that while HNF4α2 is clearly tumor suppressive in human colon cancer, HNF4α8 is not.

HNF4α2 and HNF4α8 regulate unique sets of genes relevant to tumor growth. To determine the basis for the difference in tumor growth between the HNF4α2- and HNF4α8-expressing lines, we performed RNA-seq on cells treated with DOX for 24 h. We used two different concentrations of DOX (0.1 and 0.3 μg/ml) for the HNF4α8-expressing line to more closely match the induced expression of the two isoforms (Fig. 3A and B). IB analysis verified similar levels of expression of HNF4α2 and HNF4α8, which were comparable to those expressed in the normal mouse colon (Fig. 3B). There were many more genes upregulated than downregulated in both the HNF4α2- and HNF4α8-expressing lines after DOX induction ($P < 0.01$) (Fig. 3C), consistent with HNF4α acting as a positive regulator of gene expression (1, 7, 77). While roughly 40% of the upregulated genes (156 genes) were common between the isoforms, a total of 200 genes were uniquely regulated by the different isoforms (Fig. 3D). Interestingly, HNF4α8 downregulated more genes than did HNF4α2 (83 versus 10 genes, respectively). A heat map shows some of the most highly dysregulated genes with their non-log-fold change (FC) with or without DOX (Fig. 3F). The parental rtTA line had only six genes dysregulated at the 1.5-fold cutoff (not shown). Since the two concentrations of DOX used for the HNF4α8-expressing line (0.1 and 0.3 μg/ml) showed consistent results (Fig. 3E), the two data sets were combined for subsequent analysis. (For a list of all dysregulated genes, see Tables S1 to 4 in the supplemental material.).

Gene ontology (GO) analysis revealed that both HNF4α2 and



F

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

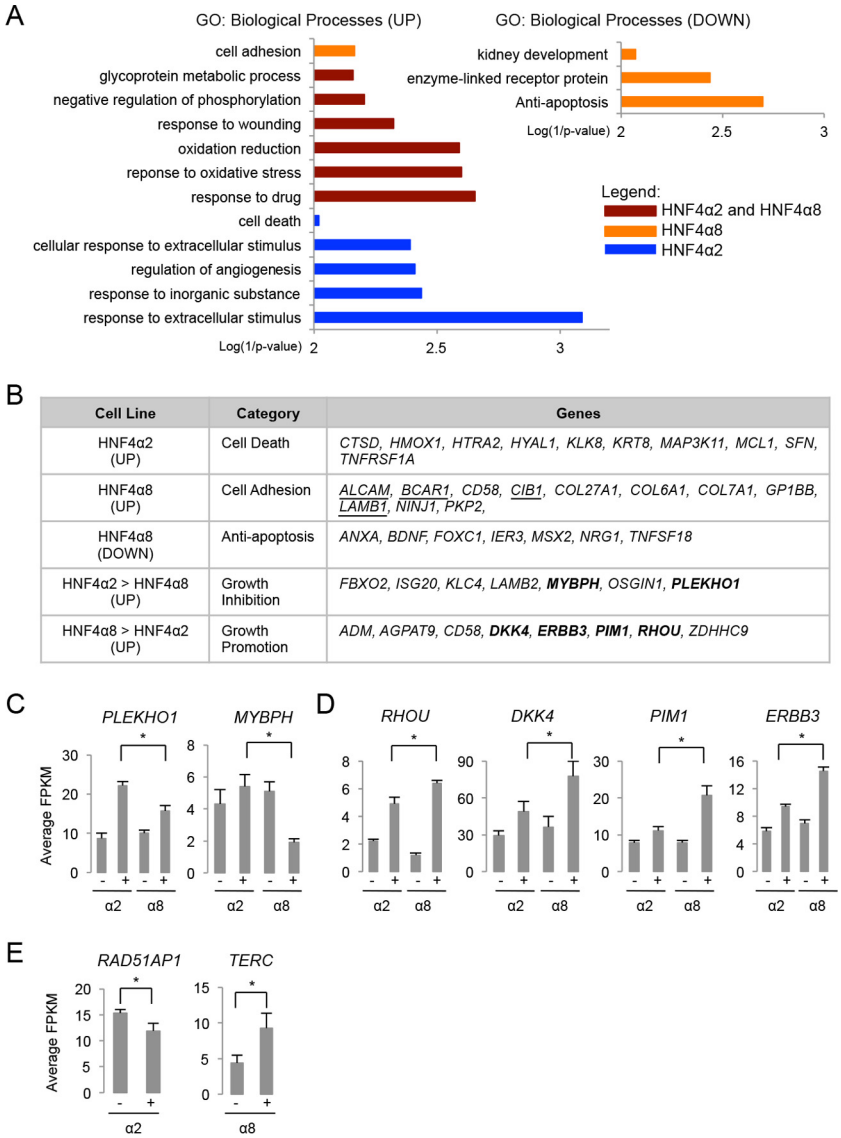
Key: 21.2 10 5 0 -5 -20 -42.9

FIG 3 Differential expression of genes in HCT116 cells by HNF4α2 and HNF4α8. (A) Schematic of the samples submitted for RNA-seq. Each circle represents 1 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. Mouse colon was taken from whole tissue and contained non-HNF4α-expressing cells as well as HNF4α-expressing cells. (C) Total number of known genes dysregulated in the HNF4α2- and HNF4α8-expressing HCT116 lines with ≥1.5-fold change (FC) upon 24 h of DOX. (D) Venn diagram of the genes in panel C. (E) As in panel D but for the two sets of HNF4α8 samples (0.1 and 0.3 μg/ml DOX). (F) Heat map showing genes with the largest FC. All values are statistically significant ($q \leq 0.05$). Zeroes are placeholders for genes without any FC.

HNF4α8 upregulated genes involved in drug metabolism, oxidative stress, negative regulation of phosphorylation, and glycoprotein metabolism (Fig. 4A). Wound healing was another common category, consistent with HNF4α playing a role in protecting the colonic epithelium from inflammatory bowel disease (10, 11, 13, 27, 78). In contrast, HNF4α2 specifically upregulated genes involved in cell death and 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 (Fig. 4B).

GO analysis did not yield a distinct category of HNF4α2-only downregulated genes, but three categories were noted for HNF4α8—kidney development (HNF4α8 is not expressed in the adult kidney), enzyme-linked receptor protein signaling, and antiapoptosis (Fig. 4A). We identified 15 genes involved in growth inhibition or promotion for which there was a significant difference between the two lines (Fig. 4B). HNF4α2 upregulated seven genes involved in growth inhibition to a greater extent than HNF4α8, while HNF4α8 more greatly upregulated eight genes involved in growth promotion. The two growth-suppressive genes preferentially upregulated by HNF4α2 were *PLEKHO1* and *MYBPH*, both of which suppress tumor progression *in vivo* (Fig. 4C) (79, 80). Two genes significantly more upregulated by HNF4α8 than HNF4α2 were *RHOU*, which regulates focal adhesion formation and cell migration (81), and *DKK4*, a known β-catenin/TCF target that enhances migration and invasion potential (82) (Fig. 4D), consistent with the higher invasive index of HNF4α8 (Fig. 2D). Also more significantly upregulated by HNF4α8 were *PIM1*, an oncogenic serine/threonine kinase gene (83), and *ERBB3*, a member of the epidermal growth factor receptor (EGFR) family of receptor tyrosine kinase genes (84) (Fig. 4D). Activation of *PIM1* promotes proliferation, inhibits apoptosis, and leads to upregulation of *ERBB3* (85).

Finally, several SNPs associated with a high probability of colorectal cancer were recently identified in the promoters of genes that are highly expressed in tumor versus normal colon tissue (86). One such gene, *RAD51AP1*, which is involved in homologous DNA repair, was downregulated only by HNF4α2, while another gene, *TERC*, which is involved in maintaining telomeres, is upregulated only by HNF4α8 (Fig. 4E).



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

FIG 4 HNF4α2 and HNF4α8 regulate different biological processes in HCT116 cells. (A) Gene ontology of genes up- or downregulated by the HNF4α isoforms (≥1.5-fold). (B) Dysregulated genes in select biological processes related to proliferation and differentiation. Greater-than signs indicate that one isoform upregulates the genes statistically more than the other. Underlining indicates cell adhesion genes that promote cell growth. Boldface indicates genes featured in panels C and D. (C and D) Average FPKM from triplicate samples (mean ± SD) of select growth-inhibiting (C) and growth-promoting (D) genes in the HNF4α2- and HNF4α8-expressing lines with or without DOX (0.3 μg/ml). (E) As in panel C but for genes containing SNPs associated with colon cancer. All FCs with or without DOX within a given line have *P* values of ≤0.05. *, $P \leq 0.05$ across cell lines.

HNF4α2 and HNF4α8 exhibit distinct chromatin occupancies *in vivo*. Considering that HNF4α2 and HNF4α8 have identical DNA binding domains (Fig. 1A), it was surprising that they regulated so many genes in a distinct fashion. To determine how many of the dysregulated genes in the RNA-seq are direct targets of the HNF4α isoforms, we performed ChIP-seq analysis 24 h after DOX induction. Although HNF4α8 continued to be expressed at a somewhat higher level than HNF4α2 (Fig. 5A), there were many more peaks in the HNF4α2-expressing line than in the HNF4α8-expressing line, as well as more genes within 50 kb of HNF4α2 peaks; there were also ~2,200 common peaks associated with ~1,600 genes (Fig. 5B). Cross-referencing the ChIP-seq with the RNA-seq data showed that among the genes with unique HNF4α2 peaks, there was considerable overlap in genes differently upregulated by HNF4α2 (101 genes), although a substantial number of the HNF4α2 unique genes had peaks for both HNF4α2 and HNF4α8 (80 genes), and many had no HNF4α peaks at all (67 genes) (Fig. 5C, top left), suggesting that they are indirect targets.

Interestingly, for the HNF4α8-upregulated genes, only one gene (*PLXNB1*) was occupied solely by HNF4α8, while many genes had peaks for both isoforms (109 genes) or only for HNF4α2 (94 genes) (Fig. 5C, top right): some of these genes were upregulated by HNF4α2 but less than the 1.5-fold cutoff used for the analysis. Examples of dysregulated genes with HNF4α ChIP peaks are shown in the lower half of Fig. 5C. One such gene (*IDH1*) is the gene coding for isocitrate dehydrogenase 1; a mutation in *IDH1/2* was recently shown to inhibit the ability of HNF4α

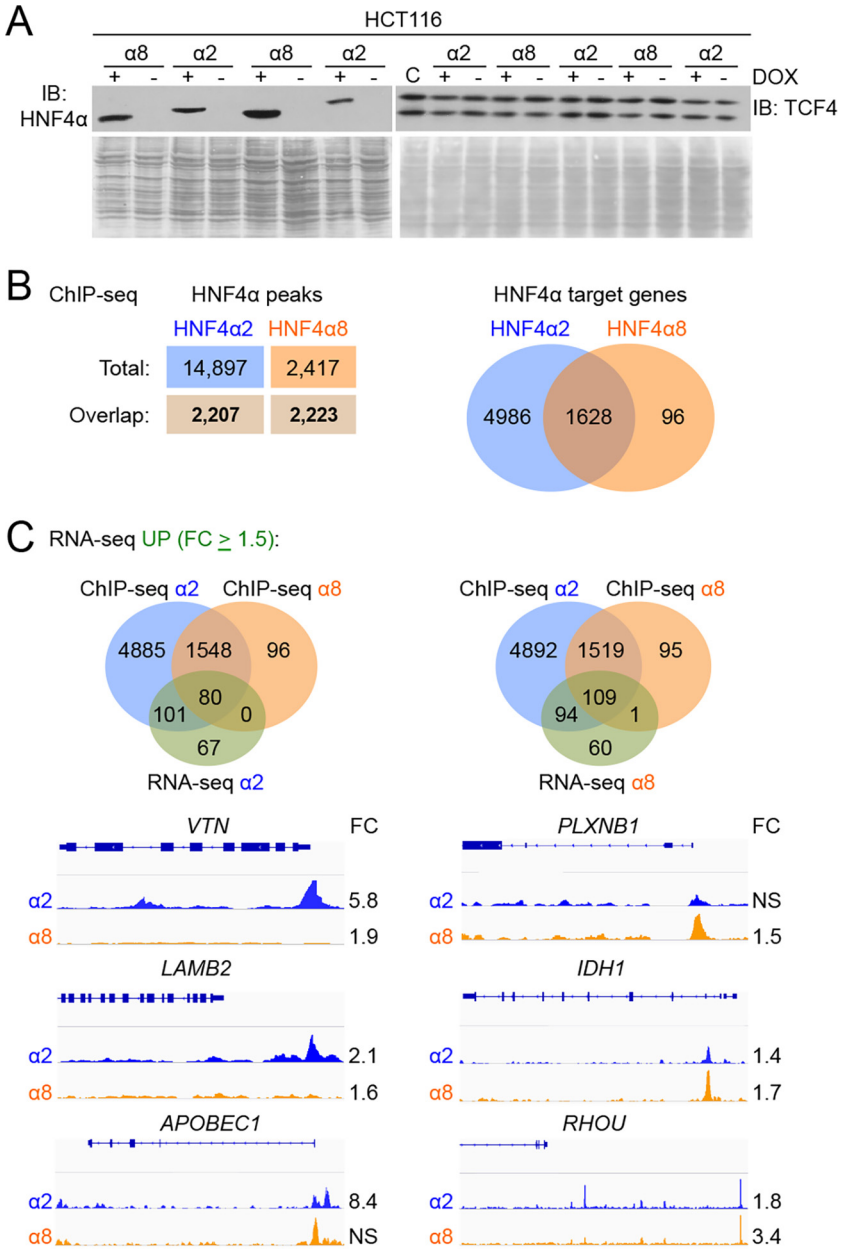


FIG 5 Integration of transcriptomic (RNA-seq) with cistromic (ChIP-seq) data in HCT116 HNF4α-inducible cell lines. (A) IB of HNF4α and TCF4 in WCE 24 h after DOX (0.3 μg/ml) induction; shown are representative samples of those used in ChIP-seq. (B, left) Total number of HNF4α ChIP-seq peaks in the +DOX samples. (B, right) Venn diagram comparing genes within 50 kb of the indicated ChIP peak. Overlapping HNF4α2 and HNF4α8 peaks were defined as those with ≥13 overlapping nucleotides. (C) Venn diagram comparing genes upregulated in RNA-seq ≥1.5-fold to genes within 50 kb of ChIP-seq peaks. Below are examples of genes with HNF4α ChIP-seq peaks and FC from RNA-seq. NS, no significant change.

to differentiate hepatocytes, thereby causing an increase in biliary cancer (87). All told, there were many more genes with HNF4α ChIP-seq peaks than were dysregulated in the RNA-seq. While this is not uncommon in genomics analysis (40), a lower cutoff (1.2-fold) and a greater distance between the peak and the putative target gene showed more overlap (not shown); a longer induction time (>24 h of DOX) also would have presumably increased the number of dysregulated genes without significantly altering the HNF4α peaks. (See Tables S5 and S6 in the supplemental material for a list of genes bound by HNF4α.).

Identification of shared and unique binding motifs for TCF/LEF and HNF4α. We next investigated whether TCF factors in Wnt/β-catenin could play a role in the differential expression by the HNF4α isoforms. Since the core of the canonical TCF binding motif (CCTTTGA) is the reverse complement of the center of the HNF4α consensus sequence (AGGTCAaAGGTCA) (43) (Fig. 6A, right), we employed a high-throughput DNA binding assay called the protein binding microarray (PBM) to determine the extent of the overlap in binding specificity between HNF4α and TCF. We identified 90 unique DNA sequences (out of ~5,000 examined)

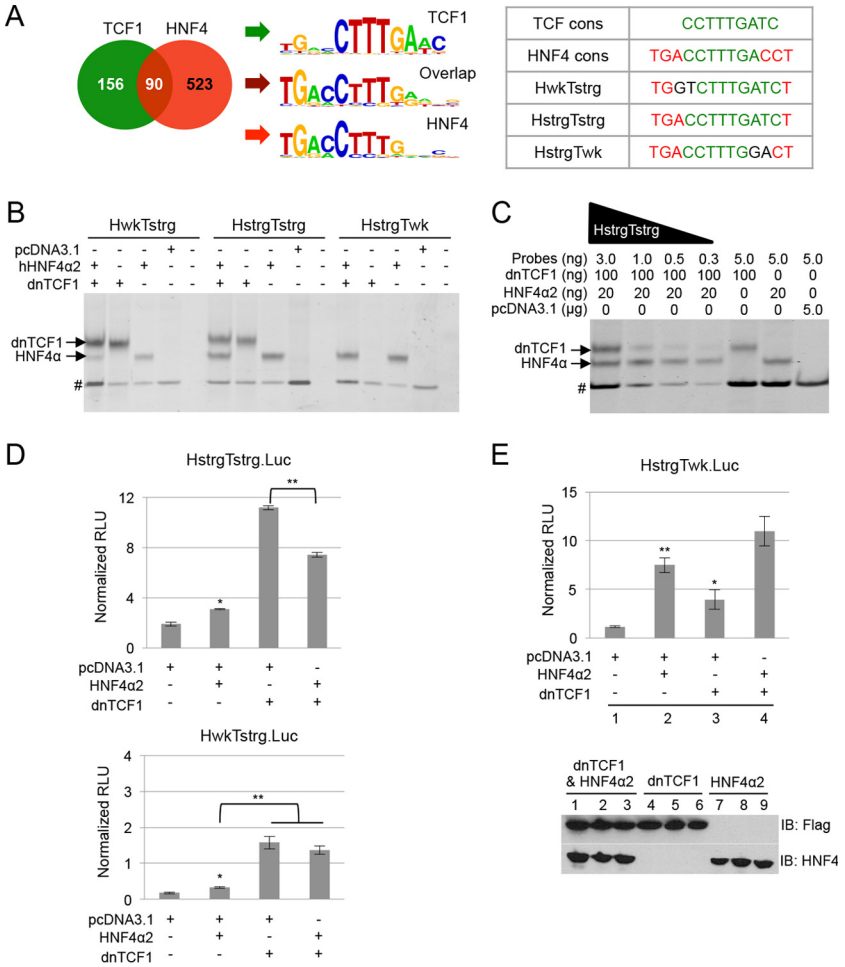


FIG 6 Overlapping DNA binding specificity of and competition between HNF4α and dnTCF1. (A, left) Venn diagram of the number of DNA sequences bound by TCF1 (dnTCF1) and HNF4α (HNF4α2) in HNF4α-centric PBM3. (A, center) Position weight matrices (PWM) showing the motif derived from the sequences bound only by TCF, only by HNF4, or by both (overlap). (A, right) Select DNA sequences with different binding strengths (strg, strong; wk, weak) for HNF4α2 (H) and dnTCF1 (T) aligned to their respective consensus sequences (cons) and color coded in green for TCF and red for HNF4α, which is typically presented in the reverse complement, AGGTCAaAGGTCA. (B) Gel shift assay with probes containing the indicated sequences defined in panel A and NE of transfected COS-7 cells expressing HNF4α2 or dnTCF1. #, nonspecific band. (C) Gel shift as in panel B but with decreasing amount of probe containing HstrgTstrg. (D and E) 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 ± SD relative light units (RLU) normalized to β-galactosidase activity (D, top, $\times 10^6$, and bottom, $\times 10^5$; E, $\times 10^4$) of triplicate samples from one of four independent experiments. (D) *, $P < 0.003$ for HNF4α2 versus pcDNA3.1 (top and bottom); **, $P \leq 0.0004$ for other comparisons as indicated. (E) *, $P \leq 0.02$ for condition 3 compared to conditions 1, 2, and 4; **, $P < 0.008$ for condition 2 compared to conditions 1 and 4. (E, lower panel) IB of WCE showing expression of dnTCF1 (Flag tagged) and HNF4α2 in the transfected cells.

that were bound by both dominant-negative TCF1 (dnTCF1) and HNF4α2 (Fig. 6A, left). Position weight matrices (PWMs) derived from sequences bound only by dnTCF1 or HNF4α2 (156 and 523, respectively), as well as those bound by both factors (90 sequences), revealed highly related binding motifs, all of which contain the CTTTG core with variations in flanking sequence (Fig. 6A, center). Note that PBM assays with 45,000 unique sequences showed that dnTCF1 (*TCF7*), TCF3 (*TCF7L1*), and TCF4 (*TCF7L2*) (all E-tail versions) have nearly identical DNA binding specificities (data are available at http://nrdbs.ucr.edu).

We selected three sequences that gave a weak (wk) or strong (strg) signal for HNF4α and TCF in the PBM (Fig. 6A, right) and designed probes for gel shift analysis. The results confirmed the relative affinity of these sequences for dnTCF1 and HNF4α2, with HwkTstrg yielding a more intense shift band for dnTCF1 than HNF4α2 and, conversely, HstrgTwk yielding an intense band for HNF4α but not dnTCF1; HstrgTstrg yielded shift bands of similar intensities for both TFs (Fig. 6B). Further characterization of the HstrgTstrg site showed that HNF4α2 bound with a higher affinity than dnTCF1 (Fig. 6C). Similar gel shift results were obtained with TCF4 (data not shown).

To determine whether dnTCF1 and HNF4α2 can compete for binding *in vivo*, we performed transient-transfection assays with luciferase constructs containing each of the three motifs used in the gel shift assay. HNF4α2 competed for transcriptional control of the constructs containing high-affinity HNF4α sites (HstrgTstrg and HstrgTwk) but not the one containing the low-affinity site (HwkTstrg) (Fig. 6D and E), suggesting that competition occurred *in vivo*. IB analysis of the transfected cells confirmed appropriate expression of the two TFs (Fig. 6E, bottom). Interestingly, dnTCF1 activated all

three of the luciferase constructs: while dnTCF1 typically acts as a repressor of transcription since it lacks the β-catenin binding domain, others have noted that it can also activate transcription (88–90). Taken together, these results indicate not only that HNF4α and TCF recognize many of the same DNA sequences 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 ~1 million spots of DNA corresponding to 250,000 sequences (in four replicates) that we mined from published HNF4α ChIP-seq data from a human liver cancer cell line, HepG2 (52), which expresses predominantly P1-HNF4α, and the colon cancer line, CaCO2 (53), which expresses predominantly P2-HNF4α (17). We found 741 DNA sequences bound by both TCF4 and HNF4α8, the majority of which also bound HNF4α2 (Fig. 7A, red spots). When we divided the sequences bound by all three TFs into three categories—TCF4 preferred (green), HNF4α preferred (red), and similar affinity (blue)—the CTTTG core was the defining feature for all categories, with some variations in the flanking nucleotides (Fig. 7A, top right). We next examined all 40,233 sequences on the PBM that contained a CTTTG motif and found ~1,100 sequences that both HNF4α2/8 and TCF4 bound well (blue spots) but nearly three times as many sequences (~3,200) that were preferred by HNF4α (red spots) and about 7,900 sequences that bound only TCF4 well (green spots). The vast majority of the 40,233 spots bound neither HNF4α nor TCF4 (gray spots), indicating that the CTTTG core is not sufficient for either TF to bind DNA in the PBM (Fig. 7B). Furthermore, a heat map of the top 1,000 TCF4 binders among the 40,233 CTTTG motifs revealed that relatively few are good HNF4α binders and vice versa (Fig. 7C). There were also relatively few qualitative differences between the HNF4α isoforms, although overall HNF4α8 tended to bind DNA with a higher affinity than HNF4α2 (Fig. 7C). (See Table S13 in the supplemental material for a list of the SNPrs numbers pertaining to the 40,233 CTTTG motifs and associated binding scores.).

The specificity and complexity of the HNF4α and TCF4 sites were further demonstrated when we examined the effect of SNPs incorporated in the PBM design. 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. 7D, black spots). Of those, 35 SNPs prevented both HNF4α2 and TCF4 from binding DNA (dual nonbinders [red spots]), while the remaining 72 interfered only with TCF4 binding (HNF4α-only binders [purple spots]). Interestingly, the majority of the dual nonbinders (22/35) had the SNP in the CTTTG core, while only 10 out of 72 of HNF4α-only binders did (Fig. 7D, lower left). In fact, the majority of the HNF4α-only binders (45/72) 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 shown in Fig. 7D include six SNPs that had notably different effects on the TCF4 and HNF4α Z-scores (bold). They also included six genes that were dysregulated by the HNF4α isoforms (*RHOU*, *SPTBN1*, *OAF*, *COBL*, *GLCE*, and *DKK1*) (Fig. 7D, right). These results suggest that TCF4 cannot tolerate any mutations in the CTTTG core, while HNF4α can, depending on the flanking sequence, and reveal subtle yet important differences in DNA binding specificity between TCF4 and HNF4α. (See Tables S11 and S12 in the supplemental material for a complete list of binding motifs altered by SNPs.).

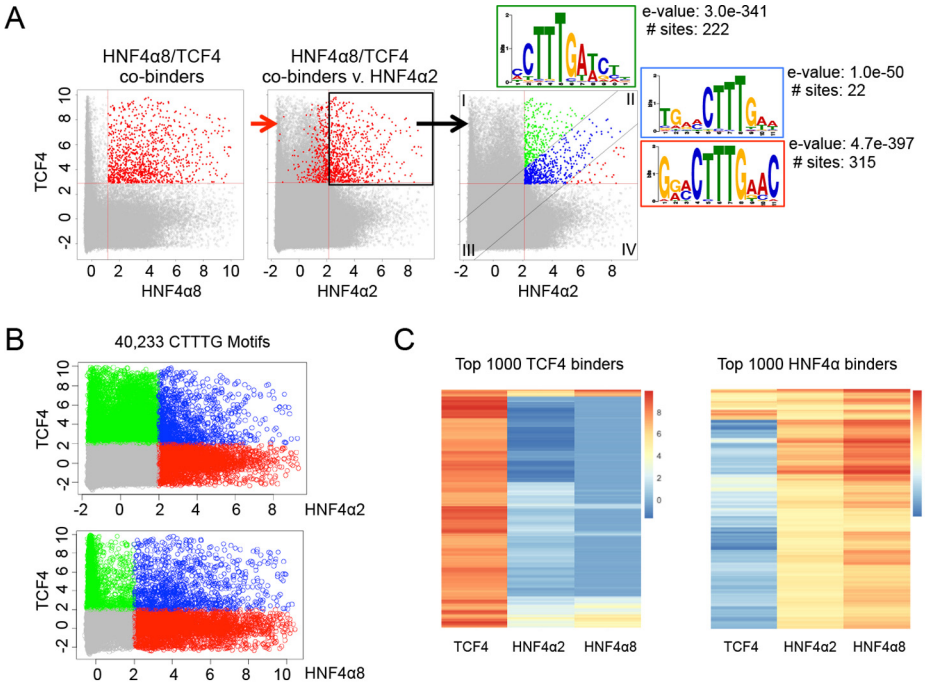
HNF4α recruits, colocalizes, and competes with TCF4 *in vivo*. To examine HNF4α and TCF4 binding *in vivo*, we determined TCF4 occupancy in the absence or presence of HNF4α (with or without DOX) in the HCT116 HNF4α2- and HNF4α8-expressing lines (Fig. 8A). We identified all differential binding peaks between TCF4−DOX and TCF4+DOX and divided them into three categories: (i) the TCF4−DOX peak is significantly smaller than the TCF4+DOX peak, (ii) the TCF4−DOX peak is roughly equal to that of TCF4+DOX peak, and (iii) the TCF4−DOX peak is larger than that of TCF4+DOX peak (Fig. 8B). When we queried how many of those peaks contained an HNF4α overlapping peak in the samples with DOX (+DOX samples), we identified three distinct binding patterns: HNF4α recruits, colocalizes, or competes with TCF4 (Fig. 8C). There were 42 TCF4 peaks that appeared only when an HNF4α2 peak was present but 78 peaks when HNF4α8 was, suggesting a preferential recruitment of TCF4 by HNF4α8. There were many more TCF4 peaks that colocalized with either of the HNF4α isoforms (485 and 126, respectively) and a difference in the function of the nearby genes: GO analysis showed that HNF4α2-unique recruiting and colocalizing peaks are associated with genes involved in metabolism and apoptotic mitochondrial changes, whereas HNF4α8-unique recruiting and colocalizing peaks regulate genes involved in cellular signaling, including the Wnt signaling pathway, underscoring a functional difference between how the HNF4α isoforms interact with TCF4 see Tables S7 and S8 in the supplemental material for a list of genes bound by TCF4 in the presence and absence of HNF4α2/HNF4α8).

An even more remarkable difference was observed between the HNF4α isoforms in the competed peaks, where the TCF4 peak was reduced in the +DOX sample at the same location that an HNF4α peak appeared. There were 60 such peaks in the HNF4α2 line, but only two were called by the MACS program in the HNF4α8 line.

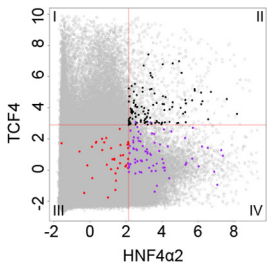
Interestingly, when we cross-referenced the HNF4α/TCF4 overlapping peaks to DNase-seq data in HCT116 from ENCODE, only about half of the recruited peaks were in regions of open chromatin, while >90% of the colocalizing and competing peaks were (except for the two HNF4α8-competed peaks) (Fig. 8C). Furthermore, the majority of the recruited peaks (32/42 for HNF4α2 and 70/78 for HNF4α8) had either more than one CTTTG site or a CTTTG site as well as a non-CTTTG HNF4α site (identified by SVM), leaving open the possibility that HNF4α binds one site and TCF4 binds another.

The vast majority of the overlapping peaks in all three categories were found close to the TSS (Fig. 8D), suggesting potential functional relevance in regulating gene expression. (See Tables S9 and S10 in the supplemental material for a list of overlapping peaks and associated genes.)

Cross-referencing the genes with TCF4/HNF4α overlapping ChIP-seq peaks with the RNA-seq data (≥1.2 FC) revealed nearly 100 dysregulated genes within 50 kb of the overlapping peak (Table 1). For the HNF4α2-competed peaks, there were six such genes, one of which was *SSH1*, a member of the slingshot homolog family of phosphatases that are important for directional cell migration (91). It was downregulated in the HNF4α2 (but not HNF4α8) RNA-seq, consistent with the HNF4α2 line having a lower invasive index (Fig. 2D). The only HNF4α2 ChIP peak in the vicinity of the *SSH1* gene is one where HNF4α2 competes with TCF4 (see Fig. 10C). In contrast, HNF4α2 appears to compete



D



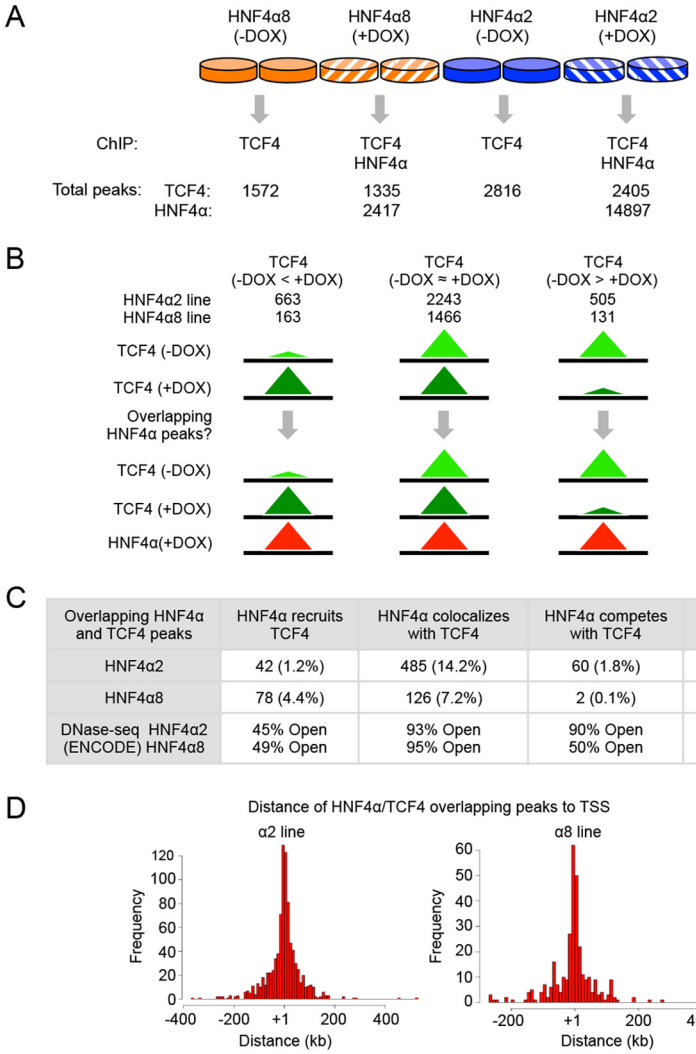
HNF4α/TCF4 cobinders to:		
Altered NT in	Dual nonbinders	HNF4α only binders
CTTTG	22	10
CTTTG flank	6	17
No CTTTG	7	45
Total:	35	72

Altered NT in:		PBM score		Z-score	
CTTTG	**sequence**	**TCF4**	**HNF4α**	**TCF4**	**HNF4α**
rs2297887	ggtaaacgtggactTtggtccggcatggc	3.1	3.7	5.8	5.8
CISD1	----------C----------	0.4	1.7		
rs77304671	accagcatgtaaccTttgtaccaattgat	3.1	4.1	3.5	6.8
MS4A4A	----------A----------	2.0	0.7		
rs1273578	ggtagagctggactTtgaacccacatctt	4.6	7.6	**6.0**	**1.3**
COQ9	----------G----------	-0.2	2.5		
rs77344780	acccccgcaatgctTtgttgtagttacac	2.9	2.2	**4.9**	**0.2**
C9orf9	----------C----------	0.1	2.1		
CTTTG flank	**sequence**	**TCF4**	**HNF4α**	**TCF4**	**HNF4α**
rs10972567	gcctgagaactttgAtccctttccattg	5.0	2.4	10.1	3.2
CREB3	----------C----------	-0.2	1.4		
rs35436776	ccagtgctctgcagTtcaaaggtccctga	6.7	3.0	4.4	6.5
RHOU	----------C----------	1.3	0.7		
rs1666561	cctttgccctgtgaTctttgttggaccct	3.7	5.3	**8.4**	**0.2**
MIR548A3	----------C----------	0.7	5.7		
rs9501868	ctgttgacattattTcaaaggtcagaatt	4.2	4.0	**7.6**	**0.1**
C6orf195	----------C----------	0.1	4.0		
No CTTTG	**sequence**	**TCF4**	**HNF4α**	**TCF4**	**HNF4α**
rs57637384	tttagtttgtctgcGgttgtttttttgtt	3.0	2.2	**3.7**	**0.2**
SPTBN1	----------A----------	1.6	2.1		
rs73007638	gggtcactggttgtTtgttcttggactgt	3.4	3.3	**4.8**	**0.1**
NCKAP5	----------C----------	1.0	3.4		

FIG 7 PBM analysis of TCF4 and HNF4α show differences in binding motifs and SNPs that alter DNA binding. (A) Scatter plots of PBM DNA binding scores for TCF4 and the HNF4α isoforms. Each of the 250,000 spots represents a unique DNA sequence and the average score of four replicates on the PBM. Red lines represent the threshold of binding, set at 2 SD above background. Red spots represent sequences bound by both HNF4α8 and TCF4. (Right) Relative binding affinity: TCF4 > HNF4α, green; TCF4 ≈ HNF4α, blue; and HNF4α > TCF4, red. Shown are the top-scoring PWM generated by MEME for each color group in quadrant II with the e-value and number of sites (sequences) used to generate the PWM. (B) As in panel A but for 40,233 sequences that contain a CTTTG motif. Blue spots represent sequences bound by both HNF4α and TCF4, red spots represent those bound preferentially by HNF4α, and green spots represent those bound preferentially by TCF4. (C) Heat maps of the top 1,000 highest binding scores among the 40,233 CTTG-containing sequences for either TCF4 (left) or HNF4α (right). (D) Scatter plot showing spots in quadrant II for which the alternate SNP allele significantly reduced DNA binding (black spots [107 total]). Red spots represent SNPs that resulted in nonbinders for both TCF4 and HNF4α2 (dual nonbinders), and purple spots represent SNPs that resulted in nonbinders for TCF4 only (HNF4α-only binders). (Bottom left) The 107 black spots are categorized based on the presence/absence of a CTTTG and whether the SNP is in the CTTTG core or flank. (Right) Examples of SNPs in each category along with the PBM score and Z-score (difference between the two alleles). rs numbers of SNPs are color coded to match the spots in the scatter plot. Boldface indicates notable differences in Z-scores.

with TCF4 to activate *FGGY*, a member of a kinase family that phosphorylates carbohydrates (92) (see Fig. 10C). There were also dysregulated genes near recruited peaks (~34% of the 86 total genes) and colocalized peaks (~15% of the 408 total genes), including several in the Wnt pathway (*WNT3*, *DKK1*, and *LRP5*) (Table 1).

Interplay between HNF4α, TCF4, and AP-1 *in vivo*. MEME analysis found the canonical CTTTG core motif as the most enriched sequence in the recruited and colocalized peaks (Fig. 9A and B), consistent with the PBM analysis (Fig. 7). Surprisingly, however, the CTTTG core was not found to be enriched in the HNF4α2-competed peaks by MEME: the only significantly en-



Overlapping HNF4α and TCF4 peaks	HNF4α recruits TCF4	HNF4α colocalizes with TCF4	HNF4α competes with TCF4	Total overlapping peaks
HNF4α2	42 (1.2%)	485 (14.2%)	60 (1.8%)	3411
HNF4α8	78 (4.4%)	126 (7.2%)	2 (0.1%)	1760
DNase-seq HNF4α2 (ENCODE) HNF4α8	45% Open 49% Open	93% Open 95% Open	90% Open 50% Open	

FIG 8 Overlapping TCF4 and HNF4α ChIP-seq peaks in HCT116 HNF4α-inducible cell lines. (A) Schematic of samples analyzed in ChIP-seq with the total number of HNF4α+DOX and TCF4+DOX and TCF4−DOX peaks in the HNF4α8 and α2 lines. (B) Three distinct categories of TCF4+DOX and TCF4−DOX are shown with the corresponding total number of peaks; each category was queried for an overlapping HNF4α peak. (C) Absolute numbers of HNF4α and TCF4 overlapping peaks in each category (recruited, colocalized, and competed) and percentage of peaks compared to total overlapping peaks. DNase-seq results in the HCT116 cells (from ENCODE) were compared to peaks in each category to determine the percentage of peaks in open or closed chromatin. (D) Histograms of all overlapping peaks for the HNF4α2 and HNF4α8 lines plotted relative to TSS (+1).

riched motif was TGAXTCA (1.3e−14) (Fig. 9C). Nonetheless, visual inspection revealed that 47 out of the 60 competing peaks do in fact contain a CTTTG motif and that 38 have one or more TGAXTCA motifs in the vicinity of a CTTTG motif. The TGA XTCA motif was also significantly enriched in the colocalized peaks for both HNF4α2 and HNF4α8 (Fig. 9B). MEME detected the AP-1 motif only in the HNF4α8 recruited peaks, but the E value was not significant ($7.1e^{+0}$).

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 to 60% of colocalized peaks and 70% of the competed peaks overlap AP-1 peaks (Fig. 9B and C, AP-1 ChIP). The HNF4α2 recruited peaks harbor CTTTG but not TGAXTCA motifs (e.g., *OAF* and *WNT3*), while the HNF4α8 peaks often contain both motifs (e.g., *SIAH2* and *ACSL5*) (Fig. 10A). Interestingly, while HNF4α2 binds the same region of *SIAH2* and *ASCL5* as HNF4α8, the ChIP-seq did not reveal any recruitment of TCF4 by HNF4α2 (Fig. 10A). This was confirmed by a re-ChIP experiment in which HNF4α8, but not HNF4α2, cooccupied the DNA with TCF4 (Fig. 9D and E, lane 12). In contrast, the re-ChIP showed that HNF4α2, but not HNF4α8, interacts with FOSL1 on *ACSL5* (Fig. 9E lane 11). A co-IP using cross-linked samples also showed an interaction between HNF4α2 and FOSL1 in both the absence (Fig. 9F) and presence (data not shown) of DNase treatment, while HNF4α8 seemed to interact less well with FOSL1. We only once observed a very weak TCF4 signal in the co-IP with HNF4α in the absence of DNase: a parallel sample treated with DNase gave no signal at all (data not shown).

On some genes, TCF4 colocalized with both HNF4α2 and HNF4α8 (e.g., *ALK* and *DKK1*), while on others, TCF4 colocalized with either HNF4α2 (e.g., *ACSL1*) or HNF4α8: CTTTG

TABLE 1 Genes with overlapping HNF4α and TCF4 peaks dysregulated in RNA-seq

Characteristic (total no. of genes)	Genes[a]	
	Upregulated	Downregulated
Competition: HNF4α2 (33)	***FGGY***, *MICAL2*	*CPOX*, *LRRTM4*, ***SSH1***, *TIAM1*
Recruitment		
HNF4α2 (33)	*ALDH3A2*, *EPDR1*, *FAM169A*, *NFE2L3*, ***OAF****, *PFKP*, *PRKAG2*, ***WNT3***	
HNF4α8 (53)	***ACSL5***, *CBLB**, *CLEC16A*, *EPDR1*, *FAM169A*, *FRMD6*, *GPD2*, *HDGF*, *IRF2BP2*, *KIAA1671*, *LRP5*, *NSMCE1*, *PPAP2B*, *PRKAG2*, ***SIAH2***, *SLC2A1*, *SPTBN1**	*C16orf45*, *CENPF*, *EFHD2*, *PHGDH*
Colocalization		
HNF4α2 (324)	***ACSL1***, *AMOTL1*, *ASS1*, *CD59*, *COBL**, *CTSB*, *EVPL*, *F2RL1*, *FAM129B*, *GLCE**, *GRK5*, *HMGCL*, *ITGB5*, *ITPKA*, *LDLR*, *NEK6*, *OSBP*, *PDE2A*, *PFKP*, *PTPRH*, *SERINC2*, *SERPINB1*, *SLC2A1*	*ATP6AP1L*, ***DKK1****, *ETV1*, *IER3*, *KIAA1430*, *LRP6*, *MAP2*, *MET*, *MID1*, *PCDH7*, *PPP3CA*, *PRKACB*, *RAP2B*, *SASS6*, *SCG2*, *STEAP1*, *SYTL3*, *TEAD1*, *TNIK*, *TSPANS*, *WDR62*
HNF4α8 (84)	*ABHD2*, *AMN1*, *AMOTL1*, *CDC14B*, *ETFB*, *GRK5*, *IRF2BP2*, *LPCAT3*, *MICAL2*, *MLEC*, *SERINC2*, *SLC35D2*	***DKK1****, *IER3*, *MOSPD1*, *MSRB3*, *TEAD1*, *TIAM1*, *ZNRF3*

[a] Boldface indicates genes shown in Fig. 10. Total genes with a TSS within 50 kb of the peak center are indicated. Asterisks indicate genes identified in Fig. 7 that contain a binding site bound by both HNF4α and TCF4.

and TGAXTCA motifs could be found in all of these peaks and coincided with the FOSL1/JUND peaks on *ALK* and *ACSL1* (Fig. 10B). In the TCF4 peaks competed by HNF4α2 there are TGAXTCA motifs, as well as CTTTG motifs (*NUDT13*, *FGGY*, and *ABHD2*) (Fig. 10C). Interestingly, on *ABHD2*, HNF4α8 colocalized with TCF4, while HNF4α2 modestly competed with TCF4. On *SSH1*, HNF4α2 but not HNF4α8 binds well and there is a single TGAXTCA motif.

DISCUSSION

In order to elucidate apparently contradictory roles of HNF4α in cancer in different tissues, we examined the effects of the two major HNF4α isoforms driven by the P1 and P2 promoters, HNF4α2 and HNF4α8. While there are a few reports on the differential activity of the HNF4α isoforms (22, 93–95), to our knowledge, this is the first in-depth functional comparison of the HNF4α isoforms in a colon cancer line. Likewise, while there are reports of interactions between HNF4α and TCF4 (12, 31, 40–45), this is the first report to examine the effect of the presence of the HNF4α isoforms on TCF4 chromatin binding, identify a potential three-way interaction between HNF4α, TCF4, and AP-1, and examine in great depth the DNA binding specificity of the HNF4α isoforms and TCF4.

Using a human colon cancer cell line (HCT116) with inducible expression of a single HNF4α isoform, we show that while P1-driven HNF4α2 clearly suppresses the growth of tumors in colon cancer cells, P2-driven HNF4α8 does not (Fig. 2). RNA-seq analysis suggests that this functional difference is due to differential expression of certain target genes, with HNF4α2 upregulating genes involved in growth suppression and cell death and HNF4α8 upregulating genes involved in cell proliferation and antiapoptosis (Fig. 3 and 4).

A high-throughput *in vitro* DNA binding assay (PBM) of 250,000 distinct human genomic sequences identified 741 unique DNA sequences that were bound by both HNF4α and TCF4 (Fig. 7A). Nearly all contained the common CTTTG core found in the TCF and HNF4α consensus sequences (Fig. 6A and 7A), as did an even larger number of sequences bound by either TCF4 or HNF4α alone (Fig. 7B). Analysis of individual sequences revealed 107 instances in which a SNP significantly altered the binding of one or both TFs (Fig. 7D), indicating that DNA binding specificity is remarkably complex and can be exquisitely sensitive to sequence alterations.

Comparison of HNF4α and TCF4 chromatin binding *in vivo* identified ~793 TCF4/HNF4α overlapping peaks that could be grouped into three categories: HNF4α (i) recruits, (ii) colocalizes, or (iii) competes with TCF4 (Fig. 8 to 10). There were notable differences in the enriched motifs in the peaks in the different categories as well as the relative distributions of overlapping peaks—HNF4α8 recruited TCF4 more frequently than HNF4α2, but HNF4α2 competed with TCF4 more frequently. The overlapping peaks are relatively close to genes (Fig. 8D), many of which were dysregulated in the RNA-seq (Table 1), suggesting that the interaction is indeed functional. Luciferase assays confirmed a competition between HNF4α and TCF4 at the level of transactivation (Fig. 6D and E).

HNF4α recruits TCF4 to the chromatin. The TCF4-recruited peaks for both HNF4α2 and HNF4α8 harbor the common CTTTG core, which raises two issues: why does TCF4 not bind these sequences on its own, and can HNF4α and TCF4 bind the same site at the same time? Since ~50% of the TCF4-recruited peaks are in regions of closed chromatin, one possibility is that HNF4α acts as a pioneering factor, binding first to the chromatin and subsequently making the site more accessible to TCF4 (Fig. 11A). In regions of open chromatin, binding by HNF4α to a high-affinity site may somehow enhance binding by TCF4 to a nearby low-affinity site (assisted loading model): the PBM analysis revealed many sites containing the CTTTG motif that have high affinity for HNF4α but low affinity for TCF4 (Fig. 7B, red spots). A third possibility (not shown) is that HNF4α and TCF4 actually do bind the same site at the same time: TCF4 is known to bind in the minor groove and nuclear receptors in the major groove, and both TCF4 and HNF4α have been shown to bend DNA (96–99). While we did not observe a trimeric complex between HNF4α, TCF4, and DNA in our gel shift experiments, we only assayed a couple of different sequences. Additional studies will be required to determine whether any of the 3,000 or more sequences that are high affinity for HNF4α and low affinity for TCF4 could possibly bind both TFs at the same time.

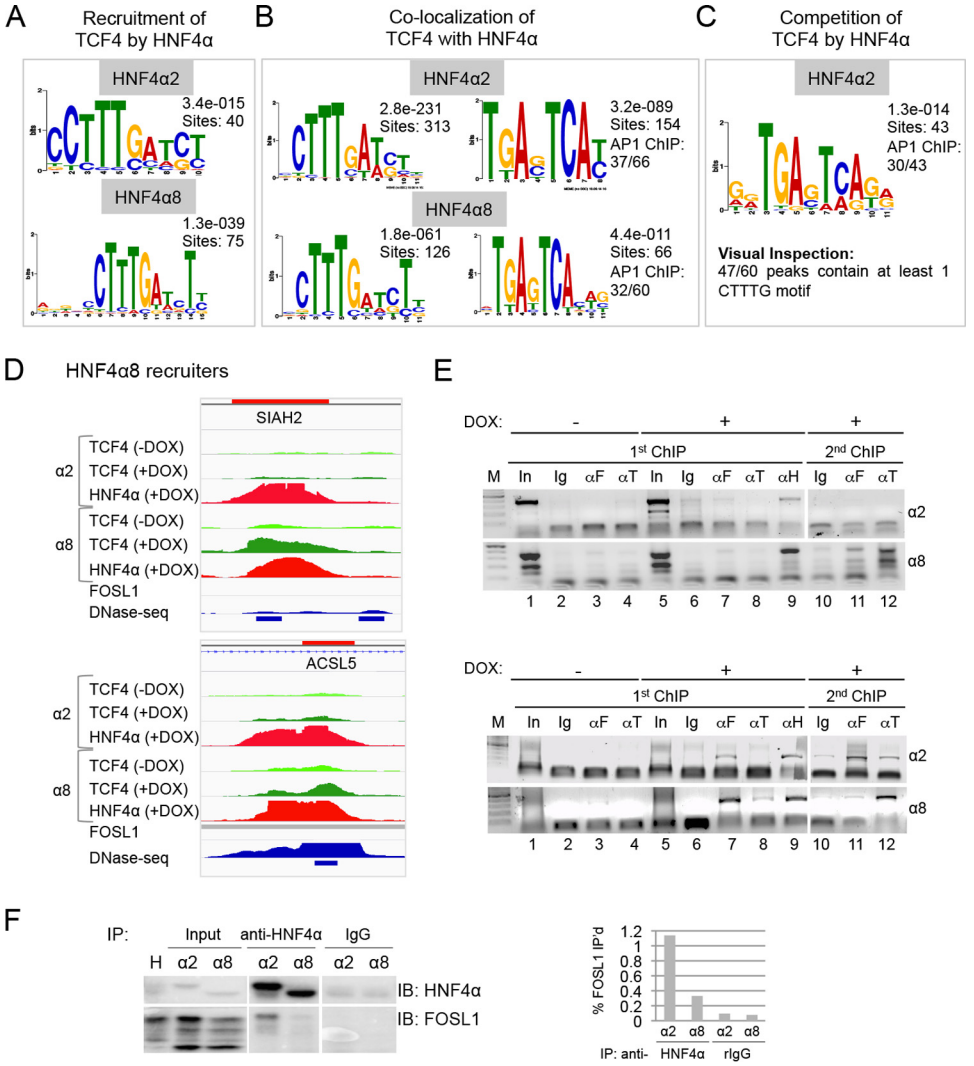


FIG 9 CTTTG and TGAXTCA motifs in TCF4/HNF4α overlapping CHIP-seq peaks and evidence of HNF4α2 and HNF4α8 differential interactions with TCF4 and AP-1. Shown are PWMs along with the E values and number of sites (sequences) used to generate the PWMs mined from the TCF4 recruited peaks (A), colocalized peaks (B), and competed peaks (C) described in the legend to Fig. 8. Peaks in each category containing a TGAXTCA motif were manually cross-referenced to FOSL1 and JUND ChIP-seq in the HCT116 ENCODE database. Given is the number of TCF4/HNF4α peaks that overlap both FOSL1 and JUND (AP-1) ChIP-seq peaks divided by the total number of peaks examined. (D and E) Snapshot from IGV of HNF4α (+DOX) and TCF4 (+DOX and −DOX) ChIP-seq peaks for *SIAH2* (D) and *ACSL5* (E). (D, left) FOSL1 ChIP-seq and DNase-seq peaks from ENCODE in HCT116 cells are shown as gray bars and blue peaks, respectively. (E) ChIP–re-ChIP followed by PCR with primer sets that amplify the region in red in the IGV. The first ChIP for IgG (Ig), FOSL1 (αF), or TCF4 (αT) (±DOX) and HNF4α (αH and α445) (+DOX) was followed by a second ChIP for IgG, FOSL1, or TCF4 (+DOX) after the HNF4α ChIP. In, input; M, molecular weight marker. Shown is one of two or three PCRs from one ChIP–re-ChIP experiment. (F, left) IB of DOX-induced HNF4α (0.3 μg/ml DOX for 24 h) followed by IB for endogenous FOSL1 in NEs from cross-linked α2 and α8 lines. The input was 1 to 2% of the total amount used in IP. Shown are the results from one of four independent experiments. (Right) Percentage of FOSL1 pulled down by the indicated Ab using the Image Lab software system (Bio-Rad).

HNF4α2 competes with TCF4 for chromatin binding: a role for AP-1? 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 contained the CTTTG core by computational analysis. Rather, the HNF4α2-competed peaks were highly enriched for AP-1 binding motifs (TGCxTCA) and were frequently bound by both FOSL1 and JUND in HCT116 cells (Fig. 9, 10, and 11B). Further visual inspection showed that there was indeed one or more CTTTG motifs in the majority of the competed peaks, underscoring a limitation of motif mining algorithms.

While we could find only one report in the literature of HNF4α interacting with AP-1 (100), other nuclear receptors, such as the glucocorticoid and estrogen receptor, have long been known to interact with AP-1 bound to DNA (101, 102). Both TCF4 and β-catenin have also been shown to interact with AP-1 at TGAXTCA motifs (103, 104). Here, we showed that HNF4α2 also interacts with FOSL1, while HNF4α8 interacts less well (Fig. 9E and F), which could explain why we did not see more HNF4α8-competed peaks (Fig. 8C). We propose that HNF4α2, but not HNF4α8, displaces TCF4 in the AP-1 complex (Fig. 11B, top). It is also possible that there is a direct competition by TCF4 and HNF4α for shared binding motifs, although those motifs would have to be specific to HNF4α2 and not bind HNF4α8 (Fig. 11B, bottom).

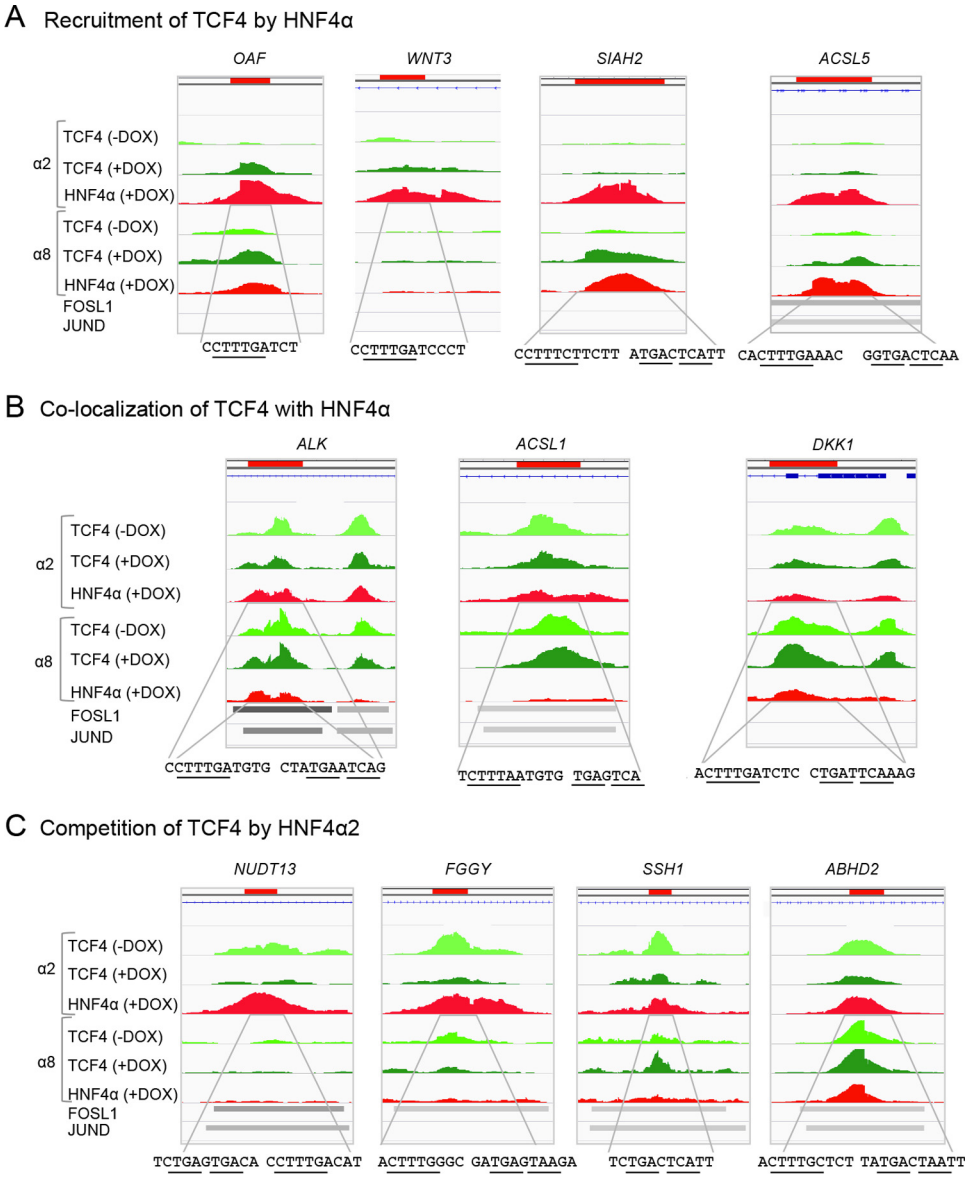


FIG 10 Examples of TCF4 peaks recruited, colocalized, and competed by HNF4α. Snapshots from IGV of HNF4α (+DOX) and TCF4 (+DOX and −DOX) ChIP-seq peaks (recruited [A], colocalized [B], and competed [C] as defined in Fig. 8) from HCT116 HNF4α2- and HNF4α8-inducible lines and FOSL1/JUND ChIP-seq peaks in HCT116 from ENCODE with the corresponding Refseq genes. The CTTTG and/or TGAXTCA motifs with surrounding nucleotide sequence are given below the peaks.

We found at least 12 examples of HNF4α2-competed peaks where TCF4 colocalizes with HNF4α8 (e.g., *ABHD2* in Fig. 10C), suggesting that the interaction between HNF4α8 and TCF4/AP-1 may be fundamentally distinct from that of HNF4α2. Finally, since the components of AP-1 (FOS and JUN) are potent proto-oncogenes, a differential interaction with the HNF4α isoforms could lead to differences in tumor growth (Fig. 2). In addition, interactions between the HNF4α isoforms and TCF4 and AP-1 on the DNA could affect interactions with coregulators and hence alter transcription, as we observed in the RNA-seq (Fig. 3 to 5). (It is not clear at this point whether the TCF4 coactivator β-catenin is present in any of these complexes nor what its effect on transactivation might be: we found examples of genes with competing peaks both up- and downregulated in the RNA-seq [Table 1].)

HNF4α colocalizes with TCF on the chromatin. The final category of overlapping HNF4α and TCF4 peaks (colocalization) 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. 9B). There are two features, however, that distinguish the colocalization and recruitment categories. The first is that there were many more genes downregulated in the colocalization category than in the recruitment category (28 versus 4, respectively) (Table 1). The second is that we observed a smaller percentage of HNF4α8 peaks than HNF4α2 peaks in the colocalized category (7.2% versus 14.2%, respectively), in contrast to the competition category, where we found nearly exclusively HNF4α2 peaks and the recruitment category that had more HNF4α8 than HNF4α2 peaks (Fig. 8C). The colocalized peaks for both

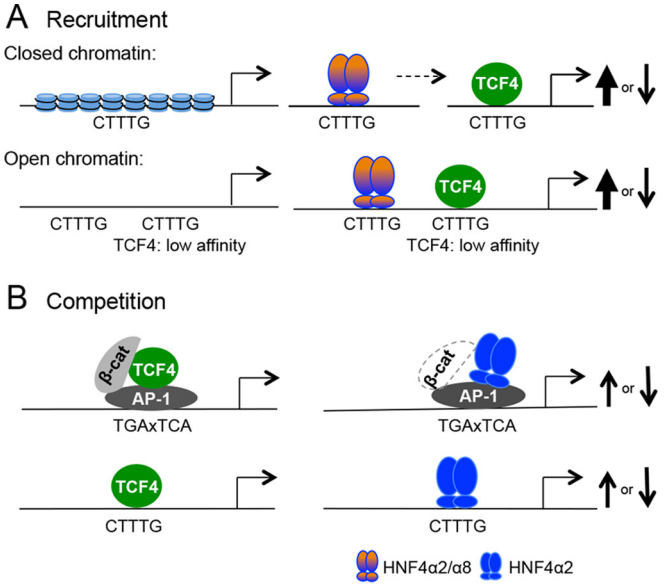


FIG 11 Model of HNF4α, TCF4, and AP-1 interplay in HCT116 human colon cancer cells. Two categories of interplay (recruitment and competition) between HNF4α and TCF4 are shown. Potential interactions with AP-1 (FOS/JUN heterodimer) are based on our ChIP-seq (Fig. 8 to 10) and ChIP-seq by ENCODE. The nucleosome array in panel A represents closed chromatin. HNF4α isoforms are color coded. Arrows indicate relative number of genes anticipated to be up- or downregulated in the recruited or competed peaks based on RNA-seq data (Fig. 3). See the text for details.

HNF4α2 and HNF4α8 frequently contained TGAXTCA motifs and overlapping AP-1 peaks, as well as CTTTG motifs (Fig. 9B), making it difficult to determine which of the motifs is relevant for the colocalization.

Another unanswered question is why, given the remarkable consensus of the common CTTTG core, are there not more HNF4α and TCF4 overlapping peaks? One potential answer is that the CTTTG core is not sufficient for DNA binding: the flanking sequence is still very important, as demonstrated by the wide range in DNA binding scores for CTTTG-containing sequences in the PBM (Fig. 7C). Furthermore, even a 1-nt change can abolish DNA binding, at least *in vitro* (Fig. 7D). This sort of fine tuning ensures that even though both HNF4α and TCF regulate many hundreds of genes, they will interact on only a subset of them and in any one of three distinct fashions, thereby allowing the cell to maintain tight control of gene expression and hence homeostasis.

In summary, this study shows that the HNF4α isoforms driven by the P1 and P2 promoters show subtle yet significant differences in chromatin binding and gene expression as well as tumor growth, suggesting that an imbalance of the isoforms may be not just a consequence of cancer, as has been observed previously (17, 18), but also a cause. Novel interactions between HNF4α and the Wnt/β-catenin/TCF pathway as well as AP-1, and a staggering complexity in DNA binding specificity, which can be affected by SNPs, were also revealed. Additional studies will be required to elucidate the molecular basis of those interactions and determine whether the results from the inducible model system employed here will translate to human colon cancer *in vivo*.

ACKNOWLEDGMENTS

We thank Z. Chen at UCLA for help with the scanner, J. Evans for injecting nude mice, D. Mane-Padros for help with TCF PBMs, and P. Deol for critical reading of the manuscript.

This study was supported by grants from the NIH: an F31 fellowship to L.M.V. (HD068611) and R01 grants DK053892 and DK094707 to F.M.S. and CA108697 and CA096878 to M.L.W. This work was supported by gifts from anonymous donors.

The funders had no role in the study design, data collection or analysis, decision to publish, or preparation of the manuscript. We have no competing financial interests to declare.

REFERENCES

1. **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 Dev **4:**2353–2365. http://dx.doi.org/10.1101/gad.4.12b.2353.
2. **Sladek FM.** 2011. What are nuclear receptor ligands? Mol Cell Endocrinol **334:**3–13. http://dx.doi.org/10.1016/j.mce.2010.06.018.
3. **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. Mol Cell Biol **21:**1393–1403. http://dx.doi.org/10.1128/MCB.21.4.1393-1403.2001.
4. **Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, Fraenkel E, Bell GI, Young RA.** 2004. Control of pancreas and liver gene expression by HNF transcription factors. Science **303:**1378–1381. http://dx.doi.org/10.1126/science.1089769.
5. **Kaestner KH.** 2010. Making the liver what it is: the many targets of the transcriptional regulator HNF4alpha. Hepatology **51:**376–377. http://dx.doi.org/10.1002/hep.23487.
6. **Sladek FM, Seidel SD.** 2001. Hepatocyte nuclear factor 4a, p 309–361. *In* Burris TP, McCabe ERB (ed), Nuclear receptors and genetic disease. Academic Press, London, United Kingdom.
7. **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. http://dx.doi.org/10.1002/hep.23357.
8. **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 Dev **8:**2466–2477. http://dx.doi.org/10.1101/gad.8.20.2466.
9. **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.
10. **Ahn SH, Shah YM, Inoue J, Morimura K, Kim I, Yim S, Lambert G, Kurotani R, Nagashima K, Gonzalez FJ, Inoue Y.** 2008. Hepatocyte nuclear factor 4alpha in the intestinal epithelial cells protects against inflammatory bowel disease. Inflamm Bowel Dis **14:**908–920. http://dx.doi.org/10.1002/ibd.20413.
11. **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. Am J Physiol Gastrointest Liver Physiol **297:**G124–G134. http://dx.doi.org/10.1152/ajpgi.90690.2008.
12. **Cattin AL, Le Beyec J, Barreau F, Saint-Just S, Houllier A, Gonzalez FJ, Robine S, Pincon-Raymond M, Cardot P, Lacasa M, Ribeiro A.** 2009. Hepatocyte nuclear factor 4alpha, a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. Mol Cell Biol **29:**6294–6308. http://dx.doi.org/10.1128/MCB.00939-09.
13. **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. http://dx.doi.org/10.1371/journal.pone.0007609.
14. **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 transcrip-

tional potentials than P2 promoter-driven isoforms. Endocrinology **144:** 1686–1694. http://dx.doi.org/10.1210/en.2002-0024.
15. **Martovetsky G, Tee JB, Nigam SK.** 2013. Hepatocyte nuclear factors 4alpha and 1alpha regulate kidney developmental expression of drug-metabolizing enzymes and drug transporters. Mol Pharmacol **84:**808–823. http://dx.doi.org/10.1124/mol.113.088229.
16. **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. http://dx.doi.org/10.1038/384458a0.
17. **Chellappa K, Jankova L, Schnabl JM, Pan S, Brelivet Y, Fung CL, Chan C, Dent OF, Clarke SJ, Robertson GR, Sladek FM.** 2012. Src tyrosine kinase phosphorylation of nuclear receptor HNF4alpha correlates with isoform-specific loss of HNF4alpha in human colon cancer. Proc Natl Acad Sci U S A **109:**2302–2307. http://dx.doi.org/10.1073/pnas.1106799109.
18. **Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C, Umezu H, Hirayama Y, Midorikawa Y, Hippo Y, Watanabe A, Uchiyama Y, Hasegawa G, Reid P, Aburatani H, Hamakubo T, Sakai J, Naito M, Kodama T.** 2006. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. J Pathol **208:**662–672. http://dx.doi.org/10.1002/path.1928.
19. **Bolotin E, Schnabl JM, Sladek FM.** 2012. HNF4α. *In* Yusuf D, Butland SL, Swanson MI, Bolotin E, Ticoll A, Cheung WA, Zhang XY, Dickman CT, Fulton DL, Lim JS, Schnabl JM, Ramos OH, Vasseur-Cognet M, de Leeuw CN, Simpson EM, Ryffel GU, Lam EW, Kist R, Wilson MS, Marco-Ferreres R, Brosens JJ, Beccari LL, Bovolenta P, Benayoun BA, Monteiro LJ, Schwenen HD, Grontved L, Wederell E, Mandrup S, Veitia RA, Chakravarthy H, Hoodless PA, Mancarelli MM, Torbett BE, Banham AH, Reddy SP, Cullum RL, Liedtke M, Tschan MP, Vaz M, Rizzino A, Zannini M, Frietze S, Farnham PJ, Eijkelenboom A, Brown, PJ, Laperriere D, Leprince D, de Cristofaro T, Prince KL, et al (ed), The transcription factor encyclopedia. Genome Biol **13:**R24. http://dx.doi.org/10.1186/gb-2012-13-3-r24.
20. **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 Res **26:** 497–504.
21. **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. http://dx.doi.org/10.2337/db07-1742.
22. **Briancon N, Weiss MC.** 2006. In vivo role of the HNF4alpha AF-1 activation domain revealed by exon swapping. EMBO J **25:**1253–1262. http://dx.doi.org/10.1038/sj.emboj.7601021.
23. **Dean S, Tang JI, Seckl JR, Nyirenda MJ.** 2010. Developmental and tissue-specific regulation of hepatocyte nuclear factor 4-alpha (HNF4-alpha) isoforms in rodents. Gene Expr **14:**337–344. http://dx.doi.org/10.3727/105221610X12717040569901.
24. **Ning BF, Ding J, Yin C, Zhong W, Wu K, Zeng X, Yang W, Chen YX, Zhang JP, Zhang X, Wang HY, Xie WF.** 2010. Hepatocyte nuclear factor 4 alpha suppresses the development of hepatocellular carcinoma. Cancer Res **70:**7640–7651. http://dx.doi.org/10.1158/0008-5472.CAN-10-0824.
25. **Bonzo JA, Ferry CH, Matsubara T, Kim JH, Gonzalez FJ.** 2012. Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4alpha in adult mice. J Biol Chem **287:**7345–7356. http://dx.doi.org/10.1074/jbc.M111.334599.
26. **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. Am J Physiol Gastrointest Liver Physiol **304:**G26–G37. http://dx.doi.org/10.1152/ajpgi.00064.2012.
27. **Hatziapostolou M, Polytarchou C, Aggelidou E, Drakaki A, Poultsides GA, Jaeger SA, Ogata H, Karin M, Struhl K, Hadzopoulou-Cladaras M, Iliopoulos D.** 2011. An HNF4alpha-miRNA inflammatory feedback circuit regulates hepatocellular oncogenesis. Cell **147:**1233–1247. http://dx.doi.org/10.1016/j.cell.2011.10.043.
28. **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. Mol Endocrinol **16:**402–410. http://dx.doi.org/10.1210/mend.16.2.0769.
29. **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. Biochem J **400:**303–313. http://dx.doi.org/10.1042/BJ20060614.
30. **Hwang-Verslues WW, Sladek FM.** 2008. Nuclear receptor hepatocyte nuclear factor 4alpha1 competes with oncoprotein c-Myc for control of the p21/WAF1 promoter. Mol Endocrinol **22:**78–90. http://dx.doi.org/10.1210/me.2007-0298.
31. **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. http://dx.doi.org/10.1053/j.gastro.2009.05.038.
32. **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. http://dx.doi.org/10.4161/cc.21019.
33. **Sladek FM.** 2012. The yin and yang of proliferation and differentiation: cyclin D1 inhibits differentiation factors ChREBP and HNF4alpha. Cell Cycle **11:**3156–3157. http://dx.doi.org/10.4161/cc.21721.
34. **Oshima T, Kawasaki T, Ohashi R, Hasegawa G, Jiang S, Umezu H, Aoyagi Y, Iwanari H, Tanaka T, Hamakubo T, Kodama T, Naito M.** 2007. Downregulated P1 promoter-driven hepatocyte nuclear factor-4alpha expression in human colorectal carcinoma is a new prognostic factor against liver metastasis. Pathol Int **57:**82–90. http://dx.doi.org/10.1111/j.1440-1827.2006.02061.x.
35. **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. Pathol Int **59:**462–470. http://dx.doi.org/10.1111/j.1440-1827.2009.02394.x.
36. **Cancer Genome Atlas Network.** 2012. Comprehensive molecular characterization of human colon and rectal cancer. Nature **487:**330–337. http://dx.doi.org/10.1038/nature11252.
37. **Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S, Davies SR, Wang S, Wang P, Kinsinger CR, Rivers RC, Rodriguez H, Townsend RR, Ellis MJ, Carr SA, Tabb DL, Coffey RJ, Slebos RJ, Liebler DC, NCI CPTAC.** 2014. Proteogenomic characterization of human colon and rectal cancer. Nature **513:**382–387. http://dx.doi.org/10.1038/nature13438.
38. **Chang HR, Nam S, Kook MC, Kim KT, Liu X, Yao H, Jung HR, Lemos R, Jr, Seo HH, Park HS, Gim Y, Hong D, Huh I, Kim YW, Tan D, Liu CG, Powis G, Park T, Liang H, Kim YH.** 19 November 2014. HNF4alpha is a therapeutic target that links AMPK to WNT signalling in early stage gastric cancer. Gut http://dx.doi.org/10.1136/gutjnl-2014-307918.
39. **Yang M, Li SN, Anjum KM, Gui LX, Zhu SS, Liu J, Chen JK, Liu QF, Ye GD, Wang WJ, Wu JF, Cai WY, Sun GB, Liu YJ, Liu RF, Zhang ZM, Li BA.** 2013. A double-negative feedback loop between Wnt-beta-catenin signaling and HNF4alpha regulates epithelial-mesenchymal transition in hepatocellular carcinoma. J Cell Sci **126:**5692–5703. http://dx.doi.org/10.1242/jcs.135053.
40. **Gougelet A, Torre C, Veber P, Sartor C, Bachelot L, Denechaud PD, Godard C, Moldes M, Burnol AF, Dubuquoy C, Terris B, Guillonneau F, Ye T, Schwarz M, Braeuning A, Perret C, Colnot S.** 2014. T-cell factor 4 and beta-catenin chromatin occupancies pattern zonal liver metabolism in mice. Hepatology **59:**2344–2357. http://dx.doi.org/10.1002/hep.26924.
41. **Neve B, Le Bacquer O, Caron S, Huyvaert M, Leloire A, Poulain-Godefroy O, Lecoeur C, Pattou F, Staels B, Froguel P.** 2014. Alternative human liver transcripts of TCF7L2 bind to the gluconeogenesis regulator HNF4alpha at the protein level. Diabetologia **57:**785–796. http://dx.doi.org/10.1007/s00125-013-3154-z.
42. **Norton L, Chen X, Fourcaudot M, Acharya NK, DeFronzo RA, Heikkinen S.** 2014. The mechanisms of genome-wide target gene regulation by TCF7L2 in liver cells. Nucleic Acids Res **42:**13646–13661. http://dx.doi.org/10.1093/nar/gku1225.
43. **Hatzis P, van der Flier LG, van Driel MA, Guryev V, Nielsen F,**

Denissov S, Nijman IJ, Koster J, Santo EE, Welboren W, Versteeg R, Cuppen E, van de Wetering M, Clevers H, Stunnenberg HG. 2008. Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. Mol Cell Biol **28:**2732–2744. http://dx.doi.org/10.1128/MCB.02175-07.

44. **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 Biol **13:**R52. http://dx.doi.org/10.1186/gb-2012-13-9-r52.
45. **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. http://dx.doi.org/10.1371/journal.pone.0021667.
46. **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. http://dx.doi.org/10.1016/0378-1119(94)90079-5.
47. **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. Mol Cell Biol **32:**3648–3662. http://dx.doi.org/10.1128/MCB.06769-11.
48. **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. http://dx.doi.org/10.1634/stemcells.2006-0634.
49. **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. http://dx.doi.org/10.1634/stemcells.2007-0710.
50. **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. Mol Cell Biol **15:**5131–5143.
51. **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 Res **40:**5343–5356. http://dx.doi.org/10.1093/nar/gks190.
52. **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 Res **37:**7498–7508. http://dx.doi.org/10.1093/nar/gkp823.
53. **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. Dev Cell **19:**713–726. http://dx.doi.org/10.1016/j.devcel.2010.10.006.
54. **Schneider TD, Stephens RM.** 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res **18:**6097–6100. http://dx.doi.org/10.1093/nar/18.20.6097.
55. **Crooks GE, Hon G, Chandonia JM, Brenner SE.** 2004. WebLogo: a sequence logo generator. Genome Res **14:**1188–1190. http://dx.doi.org/10.1101/gr.849004.
56. **Ruse MD, Jr, Privalsky ML, Sladek FM.** 2002. Competitive cofactor recruitment by orphan receptor hepatocyte nuclear factor 4alpha1: modulation by the F domain. Mol Cell Biol **22:**1626–1638. http://dx.doi.org/10.1128/MCB.22.6.1626-1638.2002.
57. **Trapnell C, Pachter L, Salzberg SL.** 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25:**1105–1111. http://dx.doi.org/10.1093/bioinformatics/btp120.
58. **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. Nat Biotechnol **28:**511–515. http://dx.doi.org/10.1038/nbt.1621.
59. **Ma H, Baumann CT, Li H, Strahl BD, Rice R, Jelinek MA, Aswad DW, Allis CD, Hager GL, Stallcup MR.** 2001. Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. Curr Biol **11:**1981–1985. http://dx.doi.org/10.1016/S0960-9822(01)00600-5.
60. **Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol **10:**R25. http://dx.doi.org/10.1186/gb-2009-10-3-r25.
61. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS.** 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol **9:**R137. http://dx.doi.org/10.1186/gb-2008-9-9-r137.
62. **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. http://dx.doi.org/10.1186/1471-2105-11-237.
63. **Oliveros JC.** 2007. VENNY. An interactive tool for comparing lists with Venn diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html.
64. **Huang DW, Sherman BT, Lempicki RA.** 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc **4:**44–57. http://dx.doi.org/10.1038/nprot.2008.1211.
65. **Ji H, Jiang H, Ma W, Wong WH.** 2011. Using CisGenome to analyze ChIP-chip and ChIP-seq data. Curr Protoc Bioinformatics Chapter 2:Unit2.13.
66. **Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D.** 2002. The human genome browser at UCSC. Genome Res **12:**996–1006. http://dx.doi.org/10.1101/gr.229102.
67. **Thorvaldsdottir H, Robinson JT, Mesirov JP.** 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform **14:**178–192. http://dx.doi.org/10.1093/bib/bbs017.
68. **R Core Team.** 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
69. **Kolde R.** 2015. pheatmap: pretty heatmaps. R package version 1.0.2. http://CRAN.R-project.org/package=pheatmap.
70. **Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS.** 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res **37:**W202–W208. http://dx.doi.org/10.1093/nar/gkp335.
71. **ENCODE Project Consortium.** 2012. An integrated encyclopedia of DNA elements in the human genome. Nature **489:**57–74. http://dx.doi.org/10.1038/nature11247.
72. **Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, Giste E, Johnson AK, Johnson EM, Kutyavin T, Lajoie B, Lee BK, Lee K, London D, Lotakis D, Neph S, Neri F, Nguyen ED, Qu H, Reynolds AP, Roach V, Safi A, Sanchez ME, Sanyal A, Shafer A, Simon JM, Song L, Vong S, Weaver M, Yan Y, Zhang Z, Zhang Z, Lenhard B, Tewari M, Dorschner MO, Hansen RS, Navas PA, Stamatoyannopoulos G, Iyer VR, Lieb JD, Sunyaev SR, Akey JM, Sabo PJ, Kaul R, Furey TS, Dekker J, Crawford GE, Stamatoyannopoulos JA.** 2012. The accessible chromatin landscape of the human genome. Nature **489:**75–82. http://dx.doi.org/10.1038/nature11232.
73. **Yeung TM, Gandhi SC, Wilding JL, Muschel R, Bodmer WF.** 2010. Cancer stem cells from colorectal cancer-derived cell lines. Proc Natl Acad Sci U S A **107:**3722–3727. http://dx.doi.org/10.1073/pnas.0915135107.
74. **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. http://dx.doi.org/10.1126/science.275.5307.1787.
75. **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. http://dx.doi.org/10.1126/science.275.5307.1784.
76. **Bottomly D, Kyler SL, McWeeney SK, Yochum GS.** 2010. Identification of β-catenin binding regions in colon cancer cells using ChIP-Seq. Nucleic Acids Res **38:**5735–5745. http://dx.doi.org/10.1093/nar/gkq363.
77. **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. http://dx.doi.org/10.1371/journal.pone.0005609.
78. **Chahar S, Gandhi V, Yu S, Desai K, Cowper-Sal Lari R, Kim Y, Perekatt AO, Kumar N, Thackray JK, Musolf A, Kumar N, Hoffman A, Londono D, Vazquez BN, Serrano L, Shin H, Lupien M, Gao N, Verzi MP.** 2014. Chromatin profiling reveals regulatory network shifts and a protective role for hepatocyte nuclear factor 4alpha during colitis. Mol Cell Biol **34:**3291–3304. http://dx.doi.org/10.1128/MCB.00349-14.
79. **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 Res **67:**9666–9676. http://dx.doi.org/10.1158/0008-5472.CAN-07-1050.

80. **Hosono Y, Yamaguchi T, Mizutani E, Yanagisawa K, Arima C, Tomida S, Shimada Y, Hiraoka M, Kato S, Yokoi K, Suzuki M, Takahashi T.** 2012. MYBPH, a transcriptional target of TTF-1, inhibits ROCK1, and reduces cell motility and metastasis. EMBO J **31:**481–493. http://dx.doi.org/10.1038/emboj.2011.416.
81. **Chuang YY, Valster A, Coniglio SJ, Backer JM, Symons M.** 2007. The atypical Rho family GTPase Wrch-1 regulates focal adhesion formation and cell migration. J Cell Sci **120:**1927–1934. http://dx.doi.org/10.1242/jcs.03456.
82. **Pendas-Franco N, Garcia JM, Pena C, Valle N, Palmer HG, Heinaniemi M, Carlberg C, Jimenez B, Bonilla F, Munoz A, Gonzalez-Sancho JM.** 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. http://dx.doi.org/10.1038/onc.2008.88.
83. **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. http://dx.doi.org/10.1593/neo.13172.
84. **Sollome JJ, Thavathiru E, Camenisch TD, Vaillancourt RR.** 2014. HER2/HER3 regulates extracellular acidification and cell migration through MTK1 (MEKK4). Cell Signal **26:**70–82. http://dx.doi.org/10.1016/j.cellsig.2013.08.043.
85. **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.
86. **Yao L, Tak YG, Berman BP, Farnham PJ.** 2014. Functional annotation of colon cancer risk SNPs. Nat Commun **5:**5114. http://dx.doi.org/10.1038/ncomms6114.
87. **Saha SK, Parachoniak CA, Ghanta KS, Fitamant J, Ross KN, Najem MS, Gurumurthy S, Akbay EA, Sia D, Cornella H, Militadous O, Walesky C, Deshpande V, Zhu AX, Hezel AF, Yen KE, Straley KS, Travins J, Popovici-Muller J, Gliser C, Ferrone CR, Apte U, Llovet JM, Wong KK, Ramaswamy S, Bardeesy N.** 2014. Mutation IDH inhibits HNF-4a to block hepatocyte differentiation and promote biliary cancer. Nature **513:**110–114. 10.1038/nature13441:1–5. http://dx.doi.org/10.1038/nature13441.
88. **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. Mol Cell Biol **16:**745–752.
89. **Grumolato L, Liu G, Haremaki T, Mungamuri SK, Mong P, Akiri G, Lopez-Bergami P, Arita A, Anouar Y, Mlodzik M, Ronai ZA, Brody J, Weinstein DC, Aaronson SA.** 2013. β-Catenin-independent activation of TCF1/LEF1 in human hematopoietic tumor cells through interaction with ATF2 transcription factors. PLoS Genet **9:**e1003603. http://dx.doi.org/10.1371/journal.pgen.1003603.
90. **Sprowl S, Waterman ML.** 2013. Past visits present: TCF/LEFs partner with ATFs for beta-catenin-independent activity. PLoS Genet **9:**e1003745. http://dx.doi.org/10.1371/journal.pgen.1003745.
91. **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. J Cell Biol **171:**349–359. http://dx.doi.org/10.1083/jcb.200504029.
92. **Zhang Y, Zagnitko O, Rodionova I, Osterman A, Godzik A.** 2011. The FGGY carbohydrate kinase family: insights into the evolution of functional specificities. PLoS Comput Biol **7:**e1002318. http://dx.doi.org/10.1371/journal.pcbi.1002318.
93. **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. J Biol Chem **277:**44677–44687. http://dx.doi.org/10.1074/jbc.M207545200.
94. **Torres-Padilla ME, Weiss MC.** 2003. Effects of interactions of hepatocyte nuclear factor 4alpha isoforms with coactivators and corepressors are promoter-specific. FEBS Lett **539:**19–23. http://dx.doi.org/10.1016/S0014-5793(03)00174-1.
95. **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. Biol Chem **388:**91–106.
96. **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. http://dx.doi.org/10.1038/376791a0.
97. **Jiang G, Lee U, Sladek FM.** 1997. Proposed mechanism for the stabilization of nuclear receptor DNA binding via protein dimerization. Mol Cell Biol **17:**6546–6554.
98. **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. http://dx.doi.org/10.1016/0092-8674(92)90129-Z.
99. **Lu XP, Eberhardt NL, Pfahl M.** 1993. DNA bending by retinoid X receptor-containing retinoid and thyroid hormone receptor complexes. Mol Cell Biol **13:**6509–6519.
100. **Yang J, Kong X, Martins-Santos ME, Aleman G, Chaco E, Liu GE, Wu SY, Samols D, Hakimi P, Chiang CM, Hanson RW.** 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. J Biol Chem **284:**27042–27053. http://dx.doi.org/10.1074/jbc.M109.047340.
101. **Pearce D, Matsui W, Miner JN, Yamamoto KR.** 1998. Glucocorticoid receptor transcriptional activity determined by spacing of receptor and nonreceptor DNA sites. J Biol Chem **273:**30081–30085. http://dx.doi.org/10.1074/jbc.273.46.30081.
102. **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. http://dx.doi.org/10.1126/science.277.5331.1508.
103. **Nateri AS, Spencer-Dene B, Behrens A.** 2005. Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. Nature **437:**281–285. http://dx.doi.org/10.1038/nature03914.
104. **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. http://dx.doi.org/10.1038/sj.onc.1210133.
105. **Bolotin E, Chellappa K, Hwang-Verslues W, Schnabl JM, Yang C, Sladek FM.** 2011. Nuclear receptor HNF4α binding sequences are widespread in Alu repeats. BMC Genomics **12:**560. http://dx.doi.org/10.1186/1471-2164-12-560.