Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

MOL #110619

Cooperative regulation of intestinal UDP-glucuronosyltransferases 1A8, -1A9 and 1A10 by CDX2 and HNF4α is mediated by a novel composite regulatory element

Nurul Mubarokah, Julie-Ann Hulin, Peter I Mackenzie, Ross A McKinnon, Alex Z Haines, Dong Gui Hu, and Robyn Meech

Discipline of Clinical Pharmacology, College of Medicine and Public Health, Flinders University, Flinders Medical Centre, Bedford Park, SA 5042, Australia. (NM, JAH, PIM RAM, AZH, DGH, RM)

Flinders Centre for Innovation in Cancer, College of Medicine and Public Health, Flinders University, Bedford Park, SA 5042, Australia. (PIM, RM, RAM, DGH)

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Running Title: Regulation of UGTs by CDX2 and HNF4α

Address correspondence to:

Robyn Meech, Department of Clinical Pharmacology, Flinders Medical Centre, Bedford Park, SA 5042, Australia.

Telephone: +61-8-82045394.

Fax: +61-8-82044795.

Email: robyn.meech@flinders.edu.au

Number of text pages: 14

Number of tables: 0

Number of figures: 9

Number of references: 31

Number of words in *Abstract*; 206

Number of words in *Introduction*: 776

Number of words in *Discussion*: 1205

Abbreviations used: bp, base pair; Cdx, caudal-related homeodomain protein; ChIP, chromatin immunoprecipitation; EMSA, electrophoretic mobility shift assay; gastrointestinal tract, GIT; HNF1, hepatocyte nuclear factor 1; HNF4, hepatocyte nuclear factor 4; PCR, polymerase chain reaction; UGT, UDP-glucuronosyltransferase

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

The gastrointestinal tract (GIT) expresses several UDP-glucuronosyltransferases (UGTs) that act as a first line of defence against dietary toxins, and contribute to the metabolism of orally administered drugs. The expression of *UGT1A8*, *UGT1A9*, and *UGT1A10* in GI tissues is known to be at least partly directed by the caudal homeodomain transcription factor, CDX2. We sought to further define the factors involved in regulation of the *UGT1A8-1A10* genes and identified a novel composite element located within the proximal promoters of these three genes that binds to both CDX2 and the hepatocyte nuclear factor HNF4α, and mediates synergistic activation by these factors. We also show that HNF4α and CDX2 are required for the expression of these UGT genes in colon cancer cell lines, and show robust correlation of UGT expression with CDX2 and HNF4α levels in normal human colon. Finally we show that these factors are involved in the differential expression pattern of UGT1A8 and UGT1A10, which are intestinal-specific, and that of UGT1A9, which is expressed in both intestine and liver. These studies lead to a model for the developmental patterning of *UGT1A8*, *UGT1A9*, and *UGT1A10* in hepatic and/or extrahepatic tissues involving discrete regulatory modules that may function (independently and cooperatively) in a context-dependent manner.

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

INTRODUCTION

UDP-glucuronosyltransferases (UGTs) render lipophilic small molecules more hydrophilic by conjugation with sugars, and are hence important for the inactivation and elimination of a wide variety of exogenous and endogenous chemicals. The human UGT superfamily comprises four families, each encoded at a separate genomic locus. The UGT1 locus has an unusual shared exon structure, containing 13 individual exons 1 located upstream of a set of shared exons 2–5 (Gong et al., 2001). A promoter located 5′ to each unique exon 1 drives independent transcription of separate nascent RNA transcripts. Subsequent cis-splicing of each exon 1 to the shared exons creates mRNAs with unique 5′ regions but identical 3′ ends (Ritter et al., 1992). The *UGT1A* genes can be grouped into clusters based on sequence identity; for example the adjacent *UGT1A7, UGT1A8, UGT1A9* and *UGT1A10* genes are >70% similar in their first exon sequences, whereas they are <60% similar to the other *UGT1A* genes (Gong et al., 2001).

UGTs resident in the gastrointestinal tract (GIT) play significant roles in metabolism of dietary chemicals and orally-delivered drugs. UGT1A7, UGT1A8 and UGT1A10 are considered extrahepatic and are mainly expressed in the GIT. UGT1A7 is restricted to the upper GIT (oesophagus and stomach), whilst UGT1A8 and UGT1A10 are detected at low to high levels in jejunum and ileum, and at moderate to high levels in colon (reviewed in (Ritter, 2007), with considerable inter-individual variation. UGT1A9 is expressed in the GIT as well as in liver and kidney; GIT expression appears to be mainly in the small intestine (duodenum, jejunum and ileum) with minimal levels in colon (Ritter, 2007). Collectively the enzymes encoded by *UGT1A8-1A10* are involved in significant intestinal metabolism of numerous drugs including morphine, naloxone, propranolol, acetaminiphen, ketoprofen, mycophenolic acid, raloxifen, resveratrol, and quertcetin (Ritter, 2007).

The intestine is sustained by a stem cell population located in the crypts that give rise to transit-amplifying cells that differentiate into absorptive cells (enterocytes) and various secretory cell types as they migrate from the crypt to the villus. Genes involved in xenobiotic

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

and drug metabolism are up-regulated during differentiation (Mariadason et al., 2002) and UGT protein is observed predominantly in villus enterocytes (Strassburg et al., 2000). Caudal related homeobox 2 (CDX2) is a transcription factor expressed in small intestine and colon epithelium in both proliferative crypt cells and differentiated villus cells (Suh and Traber, 1996). It activates intestine-restricted genes and is often termed a master regulator of intestinal identity (Fujiwara et al., 2009; Silberg et al., 2002). Conditional deletion of *Cdx2* in adult mice prevents expression of genes critical to intestinal cell differentiation leading to loss of essential absorptive functions (Hryniuk et al., 2012; Verzi et al., 2010). CDX2 has a number of transcriptional partners including HNF1 and GATA factors (Boudreau et al., 2002; San Roman et al., 2015; Ting et al., 2010). Recent work has revealed a critical role for Hnf4α as a partner for Cdx2 in intestinal specific gene expression (San Roman et al., 2015; Verzi et al., 2013). Genome wide ChIP-seq in mouse intestine identified widespread co-recruitment of Cdx2 and Hnf4α to adjacent sites in chromatin (Verzi et al., 2013). Simultaneous deletion of both *Hnf4α* and *Cdx2* led to fatal malnutrition due to greatly impaired survival and maturation of villus enterocytes, and revealed a role for these two factors in control of brush border formation, and absorption (San Roman et al., 2015). Moreover, CDX2 binds to the HNF4α promoter and regulates gene expression (Boyd et al., 2010; Verzi et al., 2013), reinforcing the cooperativity of these factors.

In addition to the high degree of conservation in the protein coding regions of *UGT1A8-1A10* (> 80%), their promoter regions are also closely conserved, particularly within the proximal region ~500 bp upstream of the transcription start site (TSS) (Cheng et al., 1998; Mojarrabi and Mackenzie, 1998; Strassburg et al., 1998). The *UGT1A8*, *-1A9*, and *-1A10* promoters were previously interrogated in Caco-2 colon cancer cells identifying Hepatocyte nuclear factor 1 (HNF1α) and CDX2 as regulators (Gregory et al., 2004). Although CDX2 recognition motifs were identified in the *UGT1A8, -1A9* and *-1A10* promoters, binding of CDX2 to these motifs could be demonstrated only for *UGT1A8* and *-1A10*; sequence differences in the presumptive 'CDX2 motif' in the *UGT1A9* promoter appeared to prevent CDX2 binding (Gregory et al., 2004), leaving the mechanism of

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

UGT1A9 regulation by CDX2 unresolved.

The current study shows that *UGT1A8, -1A9* and *-1A10* expression is programmed by the CDX2/HNF4α regulatory axis, and identifies a novel composite promotor element that mediates synergistic activation by these factors. Further we propose a model for regulation of intestinal/hepatic *UGT1A9* by both CDX2 and HNF4α that differs mechanistically from that of the intestine-specific *UGT1A8* and *-1A10* genes.

MATERIALS AND METHODS

UGT1A8, *-1A9*, and *-1A10* Promoter-Luciferase Constructs and mutagenesis.

The *UGT1A8, -1A9*, and *-1A10* promoter constructs in pGL3basic vector were described previously (Gregory et al., 2003) including variants containing mutations of the CDX2 binding site. Additional mutations including those in the novel HNF4/CDX2 element were generated using the QuikChange site-directed mutagenesis protocol (Stratagene, La Jolla, CA) with the primers shown in Supplemental Table 1.

Cell Culture and Transfection.

Caco2 cells obtained from the American Type Culture Collection (Manassas, VA) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 1 mM sodium pyruvate, 0.1 mM mixture of nonessential amino acids (Invitrogen, Carlsbad, CA) at 37°C in 5% CO_2. Cells were plated into 48-well plates at a density of 4 × 10^4 cells/well and transfected the following day with 0.2 μg of each pGL3basic promoter-reporter construct and 0.02 μg of the *Renilla reniformis* vector pRL-null (Promega, Madison, WI) using 2 μl/well Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen). For cotransfections, 0.2 μg of HNF4α, CDX2, or both HNF4α and CDX2 expression vectors (effectors) were added to the above reaction mix and normalized to a total of 0.4 μg DNA with empty expression vector pCMV5, before incubation with 1.2 μl/well Lipofectamine 2000. After 48h, the cells were harvested in 50 μl of 1× passive lysis buffer and 20 μl assayed for

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

firefly and *Renilla* luciferase activities using the Dual-Luciferase Reporter Assay System (Promega). Luminescence was measured using a Packard TopCount luminescence and scintillation counter (Mt. Waverly, Victoria, Australia). Firefly luciferase readings were normalized to the *Renilla* luciferase readings; the activities of each promoter construct transfected with each effector were normalized to the activities with pCMV5 cotransfection. Data is shown as mean and standard deviation (SD) from three replicates unless otherwise stated in the legend. Significance was assessed using ANOVA and post hoc Tukey's test.

The HNF4α plasmid was generated in house in the pCMX vector. The Cdx2 expression plasmid was kind gift from Dr. Cathy Mitchelmore (University of Copenhagen, Copenhagen, Denmark).

For analyses of endogenous UGT mRNA levels in response to expression of CDX2 and HNF4α cDNAs, we transfected cells with the various expression plasmids either using Lipofectamine LTX according to the manufacturer's recommendations, or by electroporation. Transfection of siRNAs targeting these transcription factors used Lipofectamine 2000 according to the manufacturer's protocol; a scrambled siRNA sequence was used as a negative control in all siRNA experiments.

To assess the reduction in CDX2 and HNF4α protein levels after siRNA transfection, cell lysates were subjected to immunoblotting analysis using anti-CDX2, anti-HNF4α and β-actin antibodies as reported elsewhere (Hu et al., 2014b). Immunoblot band densitometry was carried out using Multi Gauge Ver3.0 software (FUJIFILM, Tokyo, Japan). Immunoblot data shown is from a representative experiment.

RNA preparation and RT-PCR analysis

RNA was prepared from cells using TRIzol (Life Technologies, Carlsbad, California, www.lifetechnologies.com); after DNase treatment, cDNA was synthesized using NxGen M-MuLV reverse transcriptase (Lucigen, Wisconsin, www.lucigen.com) and random primers (New England Biolabs, Ipswich, Massachusetts, www.neb.com). Quantitative RT-PCR was

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

performed using a Corbett Rotorgene (Qiagen, Venlo, Limburg, Netherlands, www.qiagen.com) and GoTaq SYBR green (Promega). Data were normalized to the mRNA abundance of the housekeeping glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Data is shown as mean and SD from three replicates. Significance was assessed using ANOVA and post hoc Tukey's test.

Chromatin Immunoprecipitation (ChIP)

ChIP-qPCR was carried out essentially as described previously (Hu and Mackenzie, 2009). In brief, Caco2 cells were transfected with the HNF4α expression plasmid or empty pCMX plasmid using Lipofectamine LTX; 48 hours later, media was removed and cells were treated with 1% formaldehyde for 10 min to crosslink DNA and proteins, followed by quenching with glycine at a final concentration of 125 mM. Cells were harvested, sonicated, and isolated chromatin subjected to immunoprecipitation with 10µg of antibody. Rabbit antibodies against HNF4α (sc-6556) and the rabbit pre-immune IgG control (sc-2027) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA) Rabbit antibodies against CDX2 were from Biogenex (Biogenex, San Ramon, CA). The resultant immune-precipitates were captured by Protein A Sepharose CL-4B beads (GE Healthcare), washed and eluted as reported (Hu and Mackenzie, 2009). Eluates were incubated at 65°C overnight to disassociate the DNA/protein complexes and then digested with proteinase K to remove protein, followed by phenol-chloroform extraction and ethanol precipitation to purify the DNA. The DNA pellets were dissolved in 100 µl of Tris-EDTA buffer and 2 µl used as template for qPCR to detect the relevant promoter loci or the control locus using primers shown in Supplemental Table 1. Data is shown as mean and SD from three replicates unless otherwise stated in the legend. Significance was assessed using ANOVA and post hoc Tukey's test.

Electrophoretic Mobility Shift Assays (EMSAs)

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Caco2 cells were transfected with the HNF4α expression plasmid or empty pCMX plasmid using Lipofectamine LTX. Nuclear extracts were prepared as reported previously (Meech and Mackenzie, 2010). Oligonucleotide probe sequences are shown in Supplemental Table 1. The labelled probes were generated using the non-radioactive LUEGO protocol (Jullien and Herman, 2011) that combines two complementary target-specific oligonucleotides with a cy5-labelled universal oligonucleotide (Integrated DNA Technologies). EMSAs were performed as reported previously (Makarenkova et al., 2009) and analysed using the Typhoon Imaging System (GE Life Sciences). For supershift analysis we used rabbit antibodies to HNF4α (sc-6556; Santa Cruz Biotechnology) and CDX2 (Biogenex, San Ramon, CA) at 1 µg per reaction.

Analyses of Colon Adenocarcinoma (COAD) Transcriptomic Data.

The Colon Adenocarcinoma (COAD) transcriptome profiling (RNAseq) dataset generated by the Cancer Genome Atlas (TCGA) Research Network (http://cancergenome.nih.gov/) was downloaded from TCGA data portal (https://gdc-portal.nci.nih.gov/). The COAD RNAseq expression data from 41 normal colon samples and 480 colon adenocarcinoma samples were represented in the form of high-throughput sequencing counts. Genes (protein coding and noncoding) with a mean of less than 10 counts were discarded; the counts of the remaining genes were normalized using the upper quantile normalization method. Spearman's correlation analyses between the expression levels of two UGT genes (e.g. UGT1A8, -1A10) and two transcription factors (CDX2, HNF4α) in a cohort of either 41 normal tissues or 480 cancerous tissues were conducted and graphed using GraphPad Prism 7.03 software (GraphPad Inc., La Jolla, CA). A p value of 0.05 was considered statistically significant.

RESULTS

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Synergistic regulation of the* UGT1A8 *promoter by CDX2 and HNF4α

In previous work we showed that HNF1α and CDX2 cooperatively regulated the *UGT1A8, -1A9* and *-1A10* genes, and identified a functional CDX2 binding site in the *UGT1A8* and *-1A10* proximal promoters (Gregory et al., 2004). Recently HNF4α has been shown to cooperate with CDX2 in the regulation of many intestinal genes (San Roman et al., 2015; Verzi et al., 2013); our bioinformatic analysis together with previous functional analyses (Gardner-Stephen and Mackenzie, 2007), predicted potential HNF4α recognition motifs in the proximal promoters of *UGT1A8, -1A9* and *-1A10* suggesting that this paradigm may also be applicable to intestinal-expressed UGTs (see Supplemental Figure 1 for sequence alignments and motifs). To test this idea, we began by examining the roles of CDX2 and HNF4α in regulation of the prototypical intestinal-specific UGT, *UGT1A8*. The *UGT1A8* 1kb promoter contains one previously functionally defined CDX2 binding site (CDX2RE at -70bp) (Gregory et al., 2004). There are three motifs upstream in the *UGT1A8* promoter that are partially conserved with the HNF4α binding sites previously defined in *UGT1A9* (Gardner-Stephen and Mackenzie, 2007), (at -798, -360, and -290bp in *UGT1A8*). These motifs were shown to be non-functional in *UGT1A8* in the liver cell line HepG2; however they have not been functionally tested in an intestinal cell context (Figure 1A). We also predicted a new HNF4α binding motif in the proximal region of *UGT1A8* (at -44bp). To test whether HNF4α may be involved in regulation of the *UGT1A8* promoter in intestinal cells, and whether this may involve CDX2, we co-transfected Caco2 cells with promoter-luciferase reporters containing three different lengths of the *UGT1A8* promoter, with CDX2, HNF4α, or the combination of CDX2 and HNF4α. As shown in Figure 1B, the promoters were not transactivated by HNF4α alone, however they were each transactivated by CDX2. Moreover the combination of HNF4α and CDX2 synergistically activated all three promoter constructs (Figure 1B).

The ability of HNF4α and CDX2 to synergize on all three *UGT1A8* promoter constructs suggested that the new predicted HNF4αRE at -44bp and the CDX2RE contained

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

within the proximal region (-190bp from the TSS) are primarily involved in synergy. Consistent with this idea, mutation of the proximal (-44bp) HNF4αRE within the 1kb promoter construct ablated the synergistic induction by CDX2 and HNF4α (Figure 1C); ablation of distal (-811bp) HNF4αRE had no effect (not shown).

The proximity of the -44bp HNF4αRE to the previously identified CDX2 binding site (at -70bp) (Gregory et al., 2004), suggested that this CDX2 site mediates the synergy with HNF4α. To test this idea, we mutated the -70bp CDX2 site within the -190bp *UGT1A8* promoter construct, and tested for induction by CDX2, HNF4α, or the combination of CDX2 and HNF4α. Unexpectedly, whilst this mutation prevented induction by CDX2 alone, there was still synergistic activation by CDX2 and HNF4α (Figure 1D). Finally, we tested the ability of a *UGT1A8* promoter variant with a mutation in the initiator-like element (Sp1/Inr) to be activated by these transcription factors. Again, this mutation prevented induction by CDX2 alone, but there was still synergistic activation by CDX2 and HNF4α (Figure 1E). These data indicate that both the -70bp CDX2RE and the Sp1/Inr element are redundant for HNF4α/CDX2synergy.

Identification of a novel composite element that binds both CDX2 and HNF4α

It was previously reported that HNF4α interacts with CDX2 (Verzi et al., 2010), thus we considered the possibility that the *UGT1A8* -70bp CDX2 element is redundant for HNF4α/CDX2 synergy (Figure 1D) because CDX2 might be recruited directly to the *UGT1A8* -44bp HNF4αRE via interaction with HNF4α. To examine this possibility, we performed EMSA with a probe corresponding to the -44bp HNF4αRE. Nuclear extracts from cells transfected with HNF4α alone, or the combination of HNF4α and CDX2, were tested for binding to the probe; antibody blockade/supershift and/or mutation of the probe were used to interrogate the complexes formed. A consensus HNF4αRE probe was also used as a positive control.

As shown in Figure 2A, HNF4α formed a strong complex on the consensus

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

HNF4αRE that was supershifted by HNF4α antibody (lanes 1, 2). The HNF4α extract formed a comparatively weaker complex on the -44bp HNF4αRE probe (lane 3) but mutation of the HNF4α core recognition motif prevented this complex from forming (lane 4) indicating specificity; blockade of this complex with HNF4α antibody is also shown in Supplemental Figure 2. Of note, previous studies showed that binding of HNF4α to the functional upstream HNF4αREs in *UGT1A9* was also much weaker than to a consensus HNF4α probe (Gardner-Stephen and Mackenzie, 2007). Extracts containing both HNF4α and CDX2 formed an additional faster migrating complex on the -44bp HNF4αRE probe (lanes 7, 8) that they did not form on the consensus HNF4αRE probe (lanes 5, 6). This complex was not ablated by mutation of the core HNF4α recognition motif (lane 8).

We speculated that this faster migrating complex contained CDX2; hence we next tested whether extracts containing CDX2 alone could bind to the -44bp HNF4αRE using EMSA/supershift analysis (Figure 2B). The -70bp CDX2RE probe was used as a positive (consensus) control for CDX2 binding. CDX2 formed a robust complex with the -70bp CDX2RE probe that could be shifted by CDX2 antibody (lanes 3, 4). The CDX2 extract formed a comparatively weaker complex on the -44bp HNF4αRE probe that was also shifted by CDX2 antibody (lanes 1, 2) (Figure 2B). These data, together with that shown in Figure 2A, suggest that CDX2 might bind to the -44bp HNF4α RE probe independently of HNF4α. We also examined whether HNF4α might bind to the -70bp CDX2RE (Figure 2C). CDX2 formed a robust complex with this probe that was shifted by CDX2 antibody (lanes 1, 2); however there were no additional complexes formed by extracts that contain both CDX2 and HNF4α (lanes 3, 4). This result indicates that while CDX2 binds to the new element that we have designated the -44bp HNF4αRE, HNF4α does not bind to the previously defined -70bp CDX2RE; this finding is consistent with the redundancy of the -70bp CDX2RE for CDX2-HNF4α synergy (Fig 1D).

Overall, these data suggest that the -44bp HNF4αRE, which we have identified as

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

mediating a novel synergistic response to HNF4α and CDX2, binds to both HNF4α and CDX2. Further analysis of the sequence of this element showed that it contains a cryptic CDX2-like binding motif with the sequence TATT (Figure 3A). To test whether this motif might mediate binding to CDX2, we mutated the motif in the -44bp HNF4αRE probe and performed EMSA with extracts containing both HNF4α and CDX2. As shown in Figure 3B, mutation of the HNF4α motif blocked formation of the HNF4α complex but not the CDX2 complex (lanes 2, 3) whereas mutating the CDX2 motif (two different mutations) completely blocked formation of the CDX2 complex (lanes 4, 5). We further used unlabelled oligonucleotide competition to confirm the role of these two motifs in binding to CDX2 and HNF4α respectively (Figure 3C). The -44bp HNF4α probe formed both the CDX2 and HNF4 complexes (lane 1); a consensus HNF4αRE competitor blocked formation of the HNF4α complex but had only a modest effect on the CDX2 complex (lane 2). The consensus CDX2RE competitor blocked formation of the CDX2 complex but not the HNF4α complex (lane 3), whereas the -44bp HNF4αRE (self) competitor blocked both complexes (lane 4). A -44bp HNF4αRE competitor with a mutated HNF4α motif did not block the HNF4α complex but reduced the CDX2 complex (lane 5), in contrast a -44bp HNF4αRE competitor with a mutated CDX2 motif had little effect on the CDX2 complex but blocked the HNF4α complex (lanes 6, 7). These data further confirm that the -44bp HNF4αRE is a composite of two motifs that likely mediate adjacent binding of HNF4α and CDX2.

To assess the functional significance of the cryptic CDX2 motif in the -44bp HNF4αRE, we mutated this motif in the context of the -190bp *UGT1A8* promoter construct (Figure 4A) and assessed activation by CDX2 and HNF4α. Mutation of the cryptic CDX2 motif inhibited the synergistic activation of the promoter by CDX2 and HNF4α as effectively as mutating the HNF4α motif, showing that both motifs are required for synergy (Figure 4B).

A conserved CDX2 /HNF4α composite binding element regulates UGT1A8, -1A9 and -

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

1A10 promoters in Caco2 cells.

The sequence of the -44bp HNF4αRE in *UGT1A8* is fully conserved in the *UGT1A9* and *-1A10* proximal promoters (Figure 5A). Consistent with this conservation, chromatin immunoprecipitation (ChIP) assays using HNF4α antibody indicates that this region of all three promoters recruits HNF4α in Caco2 cells (Figure 5B). *UGT1A8* and *UGT1A10* also bear the canonical -70bp CDX2 RE; however the equivalent -70bp CDX2RE in *UGT1A9* was reported to be unable to bind CDX2 *in vitro* due to sequence divergences (mutations) (Gregory et al., 2004). *UGT1A9* also contains several HNF4α motifs distal to this proximal promoter segment (but within the 1kb promoter region) that were previously shown to be involved in regulation in liver cells (Barbier et al., 2005; Gardner-Stephen and Mackenzie, 2007). Hence we predicted that *UGT1A8* or *-1A10* would show mechanistically similar regulation by CDX2 and HNF4α; whereas *UGT1A9* may be regulated differently.

To confirm that *UGT1A10* is regulated in an equivalent manner to *UGT1A8*, we mutated the equivalent HNF4α/CDX2 composite element in *UGT1A10* (-47bp HNF4αRE in *UGT1A10*) in both the -190bp and -1kb *UGT1A10* promoters, and tested their induction in Caco2 cells (Figure 5C). Both the 190bp and 1kb *UGT1A10* promoters showed greater (synergistic) activation by CDX2 and HNF4*α* than by either factor alone, and the synergy was ablated by mutation of the -47bp HNF4αRE (Figure 5C).

We next mutated the equivalent HNF4α/CDX2 composite element in *UGT1A9* (-57bp HNF4αRE in *UGT1A9*) in both the -190bp and -1kb *UGT1A9* promoters, and tested their induction in Caco2 cells (Figure 5D). The wild-type -190bp proximal promoter construct showed no induction by CDX2 alone, but modest synergistic induction by HNF4α and CDX2. The lack of induction by CDX2 alone is in contrast to *UGT1A8* and *UGT1A10,* and is consistent with the reported non-functional/mutated CDX2 motif at approximately -70bp (Gregory et al., 2004). Importantly, synergistic activation by CDX2 and HNF4α was lost when the -57bp HNF4αRE was mutated, indicating that the HNF4α/CDX2 composite element in the proximal *UGT1A9* promoter can function similarly to that in *UGT1A8* and

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

UGT1A10. The longer -1kb *UGT1A9* promoter was transactivated by HNF4α alone (unlike the -1kb *UGT1A8* promoter), presumably due to the previously described functional upstream HNF4α sites (Barbier et al., 2005; Gardner-Stephen and Mackenzie, 2007). The -1kb *UGT1A9* promoter did not show induction by CDX2 alone, and interestingly, co-expression of CDX2 and HNF4α reduced activation of the -1kb promoter relative to HNF4α alone. This latter result may indicate competition between binding of HNF4α to the upstream HNF4αREs and the proximal -57bp HNF4αRE, as discussed later.

Overall, the data presented here indicate that the new HNF4α/CDX2 composite element can mediate HNF4α/CDX2 synergy on the *UGT1A8, -1A9* and *-1A10* proximal promoters. The discovery of this new composite element suggests a mechanism by which CDX2 might influence *UGT1A9* promoter activity in intestinal cells, given that the previously identified 'canonical' -70bp CDX2 RE was found to be non-functional (Gregory et al., 2004).

CDX2 and HNF4α regulate endogenous UGT1A8, -1A9 and -1A10 in Caco2 cells

Given the clear role for CDX2 and HNF4α in regulating the *UGT1A8, -1A9* and *-1A10* promoters, it was important to define their role in regulating the endogenous *UGT* genes. We determined that Caco2 cells express moderate-high levels of both HNF4α and CDX2, hence we elected to use siRNA-mediated knockdown of these factors to assess their roles in regulation of these *UGT* genes. The efficacy of the HNF4α and CDX2 siRNAs in reducing their target mRNA and protein levels is shown in Supplemental Figure 3. As shown in Figure 6A, HNF4α siRNA produced a 20-30% decrease of all three *UGT* genes, whilst CDX2 siRNA produced a 50-70% decrease of all three genes. We also tested the ability of these siRNAs to alter UGT expression in HT29 cells which have higher levels of both HNF4*α* and CDX2 than Caco2 cells. Both HNF4α and CDX2 siRNA produced a 40-50% decrease of all three genes. Treatment of cells with the HNF4*α* inhibitor BI6015 (Kiselyuk et al., 2012) also reduced expression of all three *UGT* genes in Caco2 cells, although the effect was only significant for *UGT1A8* and *-1A9* (Figure 6C). Overall these data indicate that both CDX2

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

and HNF4α are needed to maintain the expression level of endogenous *UGT1A8, -1A9* and *-1A10* in intestinal derived Caco2 and HT29 cells.

Previous work in mice showed that loss of CDX2 impaired HNF4α binding at co-occupied loci in intestinal cells (but not vice versa). To examine the dependence of these factors in regulation of *UGT1A8*, we used ChIP to test whether binding of exogenously-expressed HNF4α to the *UGT1A8* promoter would be affected by knockdown of endogenous CDX2. We transfected the HNF4α expression plasmid with either CDX2-siRNA or scrambled control-siRNA, and then performed ChIP using antibodies to CDX2 and HNF4α. Binding of exogenous HNF4α to the *UGT1A8* proximal promoter was inhibited after knockdown of endogenous CDX2. As expected, binding of endogenous CDX2 was also prevented by CDX2 knockdown (Figure 6D).

HNF4α is reported to be regulated by CDX2 (Verzi et al., 2013); consistent with this report, we found that CDX2 siRNA reduced not only CDX2 mRNA levels but also HNF4α mRNA levels (Supplemental Figure 3). Interestingly however, HNF4α siRNA reduced not only HNF4α levels but also CDX2 levels (Supplemental Figure 3 and 4), and a similar result was seen after treatment of cells with the HNF4α inhibitor BI6015 (Supplemental Figure 3). The regulation of CDX2 expression by HNF4α has not been previously reported. However, CDX2 was shown to bind to its own gene promoter in Caco2 cells (Boyd et al., 2010), and CDX2 and HNF4α interact, thus it is plausible that knocking down HNF4α affects CDX2 autoregulation.

Regulation of UGT1A9 by CDX2 and HNF4α is mechanistically different than regulation of UGT1A8 and -1A10.

Our data using different length promoter constructs suggests that the regulation of *UGT1A9* by HNF4α and CDX2 has two distinct components. The composite HNF4α/CDX2 element shared between *UGT1A8, -1A9* and *-1A10* appears to mediate mechanistically similar synergistic regulation of all three proximal promoters. However, the HNF4α sites

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

located further upstream in the *UGT1A9* promoter (that are not conserved in *UGT1A8* and -*1A10*) appear to mediate independent regulation of this gene by HNF4α. Differential use of these regulatory modules may play a key role in the different expression pattern of *UGT1A9* (which is both intestinal and hepatic) relative to intestinal-specific *UGT1A8* and -*1A10*. To further explore this idea we first asked whether overexpression of CDX2 and HNF4α had a different effect on endogenous UGT1A8 and -1A9 mRNA levels in intestinal (Caco2) and liver (HepG2) cell lines. In Caco2 cells, UGT1A8 mRNA was induced by transfection of a CDX2 expression plasmid alone, and synergistically by HNF4α and CDX2 together, consistent with our luciferase promoter assays. In HepG2 cells, CDX2 alone could not increase UGT1A8 mRNA; however there was slight induction by CDX2 and HNF4α together (Figure 7A). CDX2 could not induce UGT1A9 mRNA in either Caco2 or HepG2 cells, either alone or together with HNF4α. In contrast, HNF4α alone robustly induced UGT1A9 expression in both Caco-2 cells and in HepG2 cells (Figure 7A).

These data are broadly consistent with our promoter-reporter data and indicate that the endogenous *UGT1A8* gene requires CDX2 for induction by HNF4α. In contrast, HNF4α can increase *UGT1A9* mRNA expression in a CDX2-independent manner. In further support of these findings, in HepG2 cells, HNF4α siRNA had no effect on UGT1A8 mRNA, but dramatically reduced UGT1A9 mRNA levels (Figure 7B).

To augment these findings with data from an *in vivo* context, we examined whether the expression levels of UGT1A8 and UGT1A10 were correlated with levels of CDX2 and HNF4α in normal colon and in colon cancer using the TCGA database. In normal colon samples (n=41), UGT1A8 and UGT1A10 mRNA levels were extremely tightly correlated. Moreover, both genes showed a very robust correlation with levels of both CDX2 and HNF4α (Figure 8). When we examined colon cancer samples, there was still a strong correlation between UGT1A8 and UGT1A10 levels; however the correlation of both genes with levels of both CDX2 and HNF4α was weaker, albeit still statistically significant, for all comparisons except for UGT1A10 and HNF4α (Supplemental Figure 5)

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

DISCUSSION

Previous work has attempted to define the DNA elements and transcription factors responsible for the extrahepatic expression of the *UGT1A7-1A10* gene cluster. CDX2 and HNF1α were shown to play important roles in intestinal cell expression of *UGT1A8* (Gregory et al., 2004). HNF1α also regulates oesophageal cell expression of *UGT1A7* in cooperation with HNF4α (Ehmer et al., 2010). The sole member of this cluster that is expressed in liver, UGT1A9, is regulated in liver cells by HNF4α and this also involves cooperation with HNF1α (Barbier et al., 2005; Gardner-Stephen and Mackenzie, 2007). Recent genome-wide binding studies have revealed that CDX2 and HNF4α bind at adjacent sites in the developing intestine, and placed these two factors at the centre of an intestine-specific gene regulatory network (San Roman et al., 2015). Our new findings suggest that the tissue-specific patterning of *UGT1A8-1A10* expression is also determined by this fundamental developmental CDX2/HNF4α regulatory nexus.

A major finding of our study was the identification of a new composite 12nt element that binds to both CDX2 and HNF4α in the *UGT1A8-1A10* proximal promoters. CDX2 and HNF4α have been reported to interact (Verzi et al., 2010), however using mutagenesis and EMSA we were able to dismiss the hypothesis that CDX2 was recruited indirectly to this element via interaction with HNF4α, and confirm that the cryptic TATT motif within the element recruits CDX2 directly. The relative positions of CDX2 and HNF4α binding motifs have not been defined at high resolution by previous ChIP studies (Verzi et al., 2010), thus to our knowledge this is the first report of CDX2 and HNF4α binding events being integrated within a such a short (12nt) sequence. One curious aspect of our EMSA data is that interaction of the composite element probe (-44bp HNF4αRE) with extracts containing both CDX2 and HNF4α produced two distinct complexes that migrated equivalently to the complexes formed with separate CDX2 and HNF4α extracts. This suggests that the two

proteins bind different populations of probe molecules, rather than binding simultaneously to the same molecules (which would be expected to produce a slower migrating complex). However, this might be an artefact of the technique; in particular, steric hindrance may prevent co-binding to the short probe. In contrast, the native element within genomic DNA could undergo conformational changes that prevent such steric hindrance (Ismail et al., 2010). Regardless, simultaneous binding of both factors is the best explanation for the observation that their synergy is lost upon mutation of either motif; in future work this might be further supported by analyses such as re-ChIP. Our observation that HNF4α recruitment to the *UGT1A8* proximal promoter requires the presence of CDX2 is also consistent with the previous report that CDX2 promotes binding of HNF4α through chromatin remodelling (Verzi et al., 2013).

Previous work identified a conserved CDX2 binding site in the *UGT1A8* and *-1A10* promoters (at -70bp in *UGT1A8*) that is important for their activity in intestinal cells (Gregory et al., 2004). It also showed that CDX2 and HNF1α cooperate to transactivate the *UGT1A8* promoter (see Figure 9A). However, this work did not resolve how *UGT1A9* expression is activated in intestinal cells, given that the *UGT1A9* promoter lacks the equivalent functional CDX2 motif (Gregory et al., 2004). This quandary has been resolved in part by our identification of the novel composite HNF4α/CDX2 element that is fully conserved in *UGT1A8, -1A9,* and *-1A10* and that can mediate synergistic induction of all three proximal promoters by CDX2 and HNF4α. We also showed that the -70bp CDX2 motif in the *UGT1A8* promoter that mediated HNF1α/CDX2 synergy (Gregory et al., 2004), is not involved in HNF4α/CDX2 synergy. Hence at least two different elements that nucleate different complexes mediate regulation of *UGT1A8* and *-1A10* by CDX2 in intestinal cells (Figure 9A).

With regard to the complex regulation of *UGT1A9* in hepatic and intestinal cells, we propose a model in which two regulatory modules within the 1kb *UGT1A9* promoter are used in different cellular contexts. When examining short *UGT1A9* promoter constructs that omit the upstream HNF4αREs but include the proximal (approximately -57bp) HNF4α/CDX2

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

composite element, we observed the same HNF4α/CDX2 synergy that is seen with the *UGT1A8* and *-1A10* promoters. Thus we propose that this is the core 'intestinal module' for all three *UGT* genes. The function of this module may be augmented by the -70bp CDX2 element specifically in the *UGT1A8* and *-1A10* genes (Figure 9A). Studies of the long *UGT1A9* promoter construct indicate a separate 'hepatic module' involving the upstream HNF4αREs. In this model, HNF4α/CDX2 heterodimers activate *UGT1A9* through the proximal composite element whilst HNF4α homodimers activate through the upstream HNF4αREs (Figure 9B, C). The observation that the -1kb *UGT1A9* promoter construct was activated more by HNF4α alone than by co-expression of HNF4α and CDX2 (Figure 5D), suggests that the upstream HNF4αREs can mediate greater activation than the proximal element.

One observation that is not consistent with the model described above is that co-expression of HNF4α and CDX2 did not increase levels of endogenous UGT1A9 mRNA in Caco2 cells (Figure 7). The result implies that overexpressed HNF4α/CDX2 heterodimers could not access/activate the proximal composite element within the native promoter in this context. It is conceivable that this is due to an unfavourable chromatin configuration in Caco2 cells. Although cancer cell lines represent simple and tractable models for gene regulation studies, they have limitations as a developmental model. In particular, the chromatin structures that underlie developmentally-appropriate gene regulation by master regulators such as CDX2 may not be fully recapitulated in cancer cells. The developmental patterning of extrahepatic UGTs should be further studied normal intestinal models; this could involve mice carrying the human *UGT1* locus, and/or human intestinal organoids. The HNF4α inhibitor BI6015 might be a useful tool in the *in vivo* context as it robustly inhibited UGT1A8-1A10 and CDX2 expression. Interestingly, BI6015 did not alter the level of HNF4α protein (a proposed mechanism of action) and we postulate that it may inhibit the ability of HNF4α to recruit coactivators. In addition, it is now possible to study regulatory elements in a native chromatin context by genomic deletion/mutation using CRISPR. These are directions

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

that we are currently pursuing in order to better understand the roles of the distal and proximal HNF4α REs in UGT1A9 regulation in liver and intestinal cell contexts.

Discrepancies between normal intestinal tissue and cancer models were also apparent in our analyses of the TCGA database. While UGT1A8 and -1A10 levels were very closely correlated with CDX2 and HNF4α levels in normal colon, there were less robust (although still generally significant) correlations in colon cancer samples. This may reflect the deregulation of core developmental programs in cancer. It was previously reported that CDX2 can promote both differentiation and proliferation in combination with different partners (San Roman et al., 2015). Hence tumours with very different degrees of differentiation may have similar levels of CDX2, but express differing downstream programs including drug/xenobiotic metabolism.

Overall, these studies give greater insight into the control of intestinal *UGT* genes by core developmental regulators. Future studies should focus on the interplay of these developmental programs with exogenous signals (e.g. dietary chemicals and microbial metabolites) in order to understand the wide inter-individual variation in UGT levels seen in adult intestine, which in turn leads to variation in drug metabolism and detoxification capacity.

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

AUTHORSHIP CONTRIBUTIONS

Participated in research design: Mubarokah, Hulin, Hu, McKinnon, Mackenzie, Meech.

Conducted experiments: Mubarokah, Hulin, Hu, Meech.

Performed data analysis: Mubarokah, Hulin, Hu, Meech.

Wrote or contributed to the writing of the manuscript: Mubarokah, Hu, McKinnon, Mackenzie, Haines, Meech.

REFERENCES

Aueviriyavit S, Furihata T, Morimoto K, Kobayashi K and Chiba K (2007) Hepatocyte nuclear factor 1 alpha and 4 alpha are factors involved in interindividual variability in the expression of UGT1A6 and UGT1A9 but not UGT1A1, UGT1A3 and UGT1A4 mRNA in human livers. *Drug Metabolism & Pharmacokinetics* **22**(5): 391-398.

Barbier O, Girard H, Inoue Y, Duez H, Villeneuve L, Kamiya A, Fruchart J-C, Guillemette C, Gonzalez FJ and Staels B (2005) Hepatic expression of the UGT1A9 gene is governed by hepatocyte nuclear factor 4alpha. *Mol Pharmacol* **67**(1): 241-249.

Barbier O, Villeneuve L, Bocher V, Fontaine C, Torra IP, Duhem C, Kosykh V, Fruchart J-C, Guillemette C and Staels B (2003) The UDP-glucuronosyltransferase 1A9 enzyme is a peroxisome proliferator-activated receptor alpha and gamma target gene. *J Biol Chem* **278**(16): 13975-13983.

Boudreau F, Rings EH, van Wering HM, Kim RK, Swain GP, Krasinski SD, Moffett J, Grand RJ, Suh ER and Traber PG (2002) Hepatocyte nuclear factor-1 alpha, GATA-4, and caudal related homeodomain protein Cdx2 interact functionally to modulate intestinal gene transcription. Implication for the developmental regulation of the sucrase-isomaltase gene. *J Biol Chem* **277**(35): 31909-31917.

Boyd M, Hansen M, Jensen TG, Perearnau A, Olsen AK, Bram LL, Bak M, Tommerup N, Olsen J and Troelsen JT (2010) Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). *J Biol Chem* **285**(33): 25115-25125.

Cheng Z, Radominska-Pandya A and Tephly TR (1998) Cloning and expression of human UDP-glucuronosyltransferase (UGT) 1A8. *Archives of Biochemistry & Biophysics* **356**(2): 301-305.

Chouinard S, Pelletier G, Belanger A and Barbier O (2004) Cellular specific expression of the androgen-conjugating enzymes UGT2B15 and UGT2B17 in the human prostate epithelium. *Endocr Res* **30**(4): 717-725.

Ehmer U, Kalthoff S, Lankisch TO, Freiberg N, Manns MP and Strassburg CP (2010) Shared regulation of UGT1A7 by hepatocyte nuclear factor (HNF) 1alpha and HNF4alpha. *Drug Metab Dispos* **38**(7): 1246-1257.

Fujiwara R, Nakajima M, Yamamoto T, Nagao H and Yokoi T (2009) In silico and in vitro approaches to elucidate the thermal stability of human UDP-glucuronosyltransferase (UGT) 1A9. *Drug Metabolism & Pharmacokinetics* **24**(3): 235-244.

Gardner-Stephen DA and Mackenzie PI (2007) Hepatocyte nuclear factor1 transcription factors are essential for the UDP-glucuronosyltransferase 1A9 promoter response to hepatocyte nuclear factor 4alpha. *Pharmacogenet Genomics* **17**(1): 25-36.

Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, Kubota S, Carvalho S, Pennington MW, Owens IS and Popescu NC (2001) Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* **11**(4): 357-368.

Gregory PA, Gardner-Stephen DA, Lewinsky RH, Duncliffe KN and Mackenzie PI (2003) Cloning and characterization of the human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 gene promoters: differential regulation through an interior-like region. *J Biol Chem* **278**(38): 36107-36114.

Gregory PA, Lewinsky RH, Gardner-Stephen DA and Mackenzie PI (2004) Coordinate regulation of the human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 genes by hepatocyte nuclear factor 1alpha and the caudal-related homeodomain protein 2. *Mol Pharmacol* **65**(4): 953-963.

Hryniuk A, Grainger S, Savory JG and Lohnes D (2012) Cdx function is required for maintenance of intestinal identity in the adult. *Dev Biol* **363**(2): 426-437.

Hu DG and Mackenzie PI (2009) Estrogen receptor alpha, fos-related antigen-2, and c-Jun coordinately regulate human UDP glucuronosyltransferase 2B15 and 2B17 expression in response to 17beta-estradiol in MCF-7 cells. *Mol Pharmacol* **76**(2): 425-439.

Hwang-Verslues WW and Sladek FM (2010) HNF4alpha--role in drug metabolism and

potential drug target? *Current opinion in pharmacology* **10**(6): 698-705.

Ismail S, Hanapi NA, Ab Halim MR, Uchaipichat V and Mackenzie PI (2010) Effects of Andrographis paniculata and Orthosiphon stamineus extracts on the glucuronidation of 4-methylumbelliferone in human UGT isoforms. *Molecules* **15**(5): 3578-3592.

Jullien N and Herman JP (2011) LUEGO: a cost and time saving gel shift procedure. *Biotechniques* **51**(4): 267-269.

Kiselyuk A, Lee S-H, Farber-Katz S, Zhang M, Athavankar S, Cohen T, Pinkerton AB, Ye M, Bushway P, Richardson AD, Hostetler HA, Rodriguez-Lee M, Huang L, Spangler B, Smith L, Higginbotham J, Cashman J, Freeze H, Itkin-Ansari P, Dawson MI, Schroeder F, Cang Y, Mercola M and Levine F (2012) HNF4α Antagonists Discovered by a High-Throughput Screen for Modulators of the Human Insulin Promoter. *Chemistry & biology* **19**(7): 806-818.

Li J, Inoue J, Choi JM, Nakamura S, Yan Z, Fushinobu S, Kamada H, Kato H, Hashidume T, Shimizu M and Sato R (2015) Identification of the Flavonoid Luteolin as a Repressor of the Transcription Factor Hepatocyte Nuclear Factor 4alpha. *J Biol Chem* **290**(39): 24021-24035.

Makarenkova HP, Gonzalez KN, Kiosses WB and Meech R (2009) Barx2 controls myoblast fusion and promotes MyoD-mediated activation of the smooth muscle alpha-actin gene. *J Biol Chem* **284**(22): 14866-14874.

Mariadason JM, Arango D, Corner GA, Aranes MJ, Hotchkiss KA, Yang W and Augenlicht LH (2002) A gene expression profile that defines colon cell maturation in vitro. *Cancer Res* **62**(16): 4791-4804.

Meech R and Mackenzie PI (2010) UGT3A: novel UDP-glycosyltransferases of the UGT superfamily. *Drug Metab Rev* **42**(1): 43-52.

Mojarrabi B and Mackenzie PI (1998) Characterization of Two Udp Glucuronosyltransferases That Are Predominantly Expressed in Human Colon. *Biochem Biophys Res Commun* **247**(3): 704-709.

Ritter JK (2007) Intestinal UGTs as potential modifiers of pharmacokinetics and biological

responses to drugs and xenobiotics. *Expert Opin Drug Metab Toxicol* **3**(1): 93-107.

Ritter JK, Chen F, Sheen YY, Tran HM, Kimura S, Yeatman MT and Owens IS (1992) A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *J Biol Chem* **267**(5): 3257-3261.

San Roman AK, Aronson BE, Krasinski SD, Shivdasani RA and Verzi MP (2015) Transcription factors GATA4 and HNF4A control distinct aspects of intestinal homeostasis in conjunction with transcription factor CDX2. *J Biol Chem* **290**(3): 1850-1860.

Silberg DG, Sullivan J, Kang E, Swain GP, Moffett J, Sund NJ, Sackett SD and Kaestner KH (2002) Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. *Gastroenterology* **122**(3): 689-696.

Strassburg CP, Kneip S, Topp J, Obermayer-Straub P, Barut A, Tukey RH and Manns MP (2000) Polymorphic gene regulation and interindividual variation of UDP-glucuronosyltransferase activity in human small intestine. *J Biol Chem* **275**(46): 36164-36171.

Strassburg CP, Manns MP and Tukey RH (1998) Expression of the UDP-glucuronosyltransferase 1A locus in human colon. Identification and characterization of the novel extrahepatic UGT1A8. *J Biol Chem* **273**(15): 8719-8726.

Suh E and Traber PG (1996) An intestine-specific homeobox gene regulates proliferation and differentiation. *Molecular & Cellular Biology* **16**(2): 619-625.

Ting LSL, Benoit-Biancamano M-O, Bernard O, Riggs KW, Guillemette C and Ensom MHH (2010) Pharmacogenetic impact of UDP-glucuronosyltransferase metabolic pathway and multidrug resistance-associated protein 2 transport pathway on mycophenolic acid in thoracic transplant recipients: an exploratory study. *Pharmacotherapy* **30**(11): 1097-1108.

Verzi MP, Shin H, He HH, Sulahian R, Meyer CA, Montgomery RK, Fleet JC, Brown M, Liu XS and Shivdasani RA (2010) Differentiation-specific histone modifications reveal

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Dev Cell* **19**(5): 713-726.

Verzi MP, Shin H, San Roman AK, Liu XS and Shivdasani RA (2013) Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. *Mol Cell Biol* **33**(2): 281-292.

Yeung TM, Gandhi SC, Wilding JL, Muschel R and Bodmer WF (2010) Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci U S A* **107**(8): 3722-3727.

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

FOOTNOTES

This work was supported by the National Health and Medical Research Council of Australia [APP1002123]. PIM was a National Health and Medical Research Council of Australia Senior Principal Research Fellow. RM was an Australian Research Council Future Fellow. RAM is a Cancer Council SA Beat Cancer Professorial Chair.

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

LEGENDS FOR FIGURES

Figure 1. Synergistic regulation of the *UGT1A8* promoter by CDX2 and HNF4α. **A.** Schematic of the 1kb *UGT1A8* promoter region showing the positions of three predicted CDX2 binding sites, two potential HNF4α binding sites, and the Sp1/Inr element, +1 indicates the transcription start site (see text). **B.** CDX2 and HNF4α synergistically regulate *UGT1A8* promoter-reporter constructs containing either the -190bp, -250bp, or -1000bp region of the promoter. **C**. Mutation of the proximal (-44bp) but not the distal (-811bp) HNF4α motif in the *UGT1A8* 1kb promoter blocks synergistic induction. **D.** Mutation of the CDX2 binding site at -70bp site blocks induction by CDX2 alone but not the synergistic induction. **E** Mutation of the Sp1/Inr like element blocks induction by CDX2 alone but not the synergistic induction. For each panel except 1B, the data is the mean of 2 or 3 independent experiments; for panel 1B, a representative experiment performed in triplicate is shown. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ using ANOVA and post hoc Tukey's test.

Figure 2. EMSA analysis of HNF4α binding to the *UGT1A8* -44bp HNF4αRE **A.** Lanes 1, 2, 5, 6: an HNF4α consensus probe incubated with extracts containing either HNF4α or HNF4α+CDX2; addition of anti-HNF4α antibody (lane 5) inhibits complex formation. Lanes 3, 4: wildtype or mutated *UGT1A8* -44bp HNF4α probes incubated with extracts containing HNF4α. Lanes 7, 8: wildtype or mutated *UGT1A8* -44bp HNF4α probes incubated with extracts containing HNF4α+CDX2. **B.** Lanes 1, 2: the *UGT1A8* -44bp HNF4αRE probe incubated with extracts containing CDX2, without (lane 1) or with (lane 2) addition of CDX2 antibody. Lanes 3, 4: the *UGT1A8* -70bp CDX2RE probe incubated with extracts containing CDX2, without (lane 3) or with (lane 4) addition of CDX2 antibody. **C.** Lanes 1-4; the -70bp CDX2RE probe incubated with extracts that contain CDX2 alone (lanes 1, 2) or CDX2+HNF4α, without (lanes 1, 3) or with (lanes 2, 4) addition of CDX2 antibody. Lane 5: consensus HNF4α probe incubated with extracts containing CDX2+HNF4α.

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 3. EMSA mutational analysis of HNF4α and CDX2 binding to the *UGT1A8* -44bp HNF4αRE. **A**. Sequence of the *UGT1A8* -44bp HNF4αRE showing the predicted HNF4α and CDX2 binding motifs. **B.** Lane 1: consensus HNF4α probe incubated with extracts containing HNF4α+CDX2. Lane 2: the 44bp HNF4αRE wildtype probe incubated with extract containing HNF4α+CDX2. Lanes 3-5: -44bp HNF4αRE probes with mutation of either the HNF4α motif (lane 3), or the CDX2 motif (lanes 4, 5) incubated with extracts containing HNF4α+CDX2 . **C.** The -44bp HNF4αRE wildtype probe incubated with extract containing HNF4α+CDX2 without (lane 1), or with (lanes 2-7) various unlabelled competitor oligonucleotides. Lane 2: consensus HNF4αRE competitor. Lane 3: -70bp CDX2RE competitor. Lane 4: wildtype -44bp HNF4αRE competitor. Lane 5: -44bp HNF4αRE competitor with mutation of the HNF4α motif. Lane 6,7: -44bp HNF4αRE competitor with mutation of the CDX2 motif.

Figure 4. Mutation of either the CDX2 or HNF4α motif within the *UGT1A8* -44bp HNF4αRE prevents synergistic promoter activation. **A.** Schematic showing mutations generated in the CDX2 and HNF4α motifs within the -44bp HNF4αRE in the *UGT1A8* -190 promoter construct. **B**. Regulation of the *UGT1A8* promoter constructs by CDX2, HNF4α, or CDX2+HNF4α. For each dataset, n = 3 independent experiments; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ using ANOVA and post hoc Tukey's test.

Figure 5. A HNF4α/CDX2 composite binding element is conserved in the *UGT1A8-1A10* proximal promoters. **A.** Alignment of the *UGT1A8, -1A9,* and *-1A10* proximal promoters shows complete conservation of the HNF4α/CDX2 composite binding element. **B.** ChIP-qPCR analysis testing binding of HNF4α to regions spanning the HNF4α/CDX2 composite binding element in the proximal promoter regions of the *UGT1A8*, *-1A9*, and *-1A10* genes. **C**. Regulation of the -0.19kb and -1kb *UGT1A10* promoter constructs by CDX2, HNF4α, or

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

CDX2+HNF4α. **D.** Regulation of the -0.19kb and -1kb *UGT1A9* promoter constructs by CDX2, HNF4α, or CDX2+HNF4α. For each dataset, n = 2 or 3 independent experiments; * P < 0.05; ** P < 0.01; *** P < 0.001 using ANOVA and post hoc Tukey's test.

Figure 6. Inhibition of *UGT1A8-1A10* gene expression by siRNAs or inhibitors targeting HNF4α and/or CDX2. **A.** Transfection of Caco2 cells with either HNF4α or CDX2 siRNA decreases the level of UGT1A8, -1A9, and -1A10 mRNAs. **B.** Transfection of HT29 cells with either HNF4α siRNA or CDX2 siRNA and measurement of UGT1A8, -1A9, and -1A10 mRNA levels. **C.** Treatment of Caco2 cells with HNF4α inhibitor BI6015 and measurement of UGT1A8, -1A9, and -1A10 mRNA levels. For each dataset, n = 2 or 3 independent experiments; * P < 0.05; ** P < 0.01; *** P < 0.001 using ANOVA and post hoc Tukey's test.

Figure 7. *UGT1A8* and *UGT1A9* show differential regulation by CDX2 and HNF4α in hepatic and intestinal cell lines. **A.** Transfection of HepG2 and Caco2 cells with HNF4α and CDX2 expression plasmids and measurement of UGT1A8 and -1A9 mRNA levels. **B.** Transfection of HepG2 cells with HNF4α siRNA and measurement of HNF4α, UGT1A8 and -1A9 mRNA levels. For each dataset, n = 2 or 3 independent experiments; * P < 0.05; ** P < 0.01; *** P < 0.001 using ANOVA and post hoc Tukey's test.

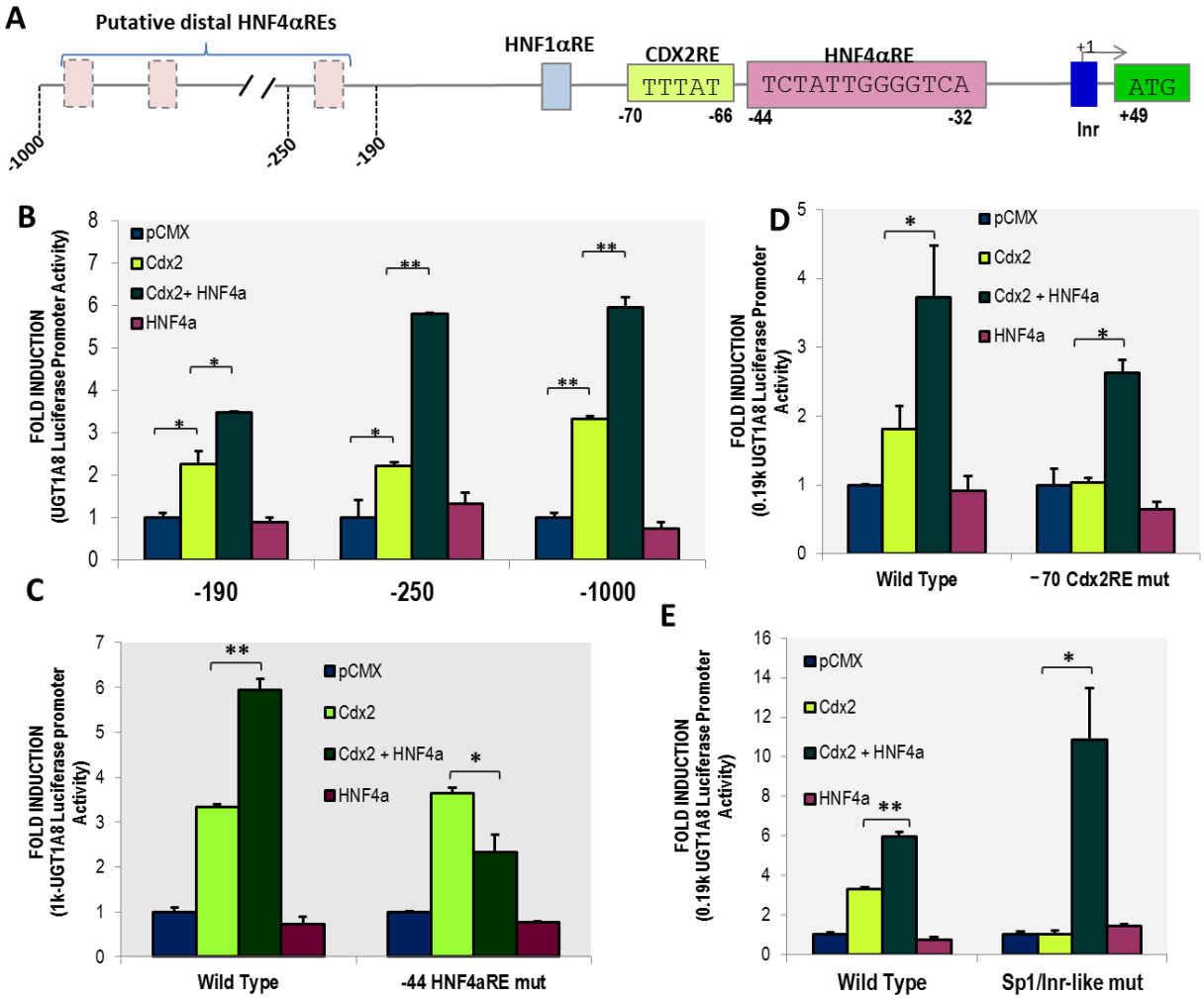
Figure 8. Analysis of UGT1A8-1A10, CDX2 and HNF4α levels in normal colon samples (n=41) using the Colon Adenocarcinoma (COAD) dataset generated by the Cancer Genome Atlas (TCGA) Research Network: http://cancergenome.nih.gov/. **A.** Correlation of UGT1A8 and UGT1A10 levels in normal colon samples. **B, C.** Correlation of UGT1A8 with levels of CDX2 (B) and HNF4α (C) in normal colon samples **D, E.** Correlation of UGT1A10 with levels of CDX2 (D) and HNF4α (E) in normal colon samples. All data analysis used the Spearman rank method with GraphPad Prism 7.03 software (GraphPad Inc., La Jolla, CA); a P value of 0.05 was considered statistically significant; r = correlation coefficient.

Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 9. A model for the regulation of *UGT1A8/1A10* and *UGT1A9* by HNF4α and CDX2. **A.** In the *UGT1A8/1A10* proximal promoters, a two part intestinal module that includes (I) the new HNF4α /CDX2 composite element (-44bp in *UGT1A8*), and (II) the previously defined CDX2 (-70bp) and HNF1α sites (-100bp). When CDX2 is high (e.g. intestine), the proximal HNF4α /CDX2 composite element (I) recruits HNF4α/CDX2 heterodimers, the upstream elements (II) may augment this response (green bracket). **B.** The *UGT1A9* promoter contains the intestinal module (I) centred on the HNF4/CDX2 composite element that is shared with *UGT1A8* and *-1A10*, as well as a hepatic regulatory module involving upstream HNF4αREs. When CDX2 is high (e.g. intestine), HNF4α forms heterodimers with CDX2 that bind the intestinal module; these may also cooperate with HNF1α (green bracket). **C.** When CDX2 is low/absent (e.g. liver), HNF4α forms homodimers that bind the hepatic module, these may also cooperate with HNF1α (green bracket). Chromatin architecture may help determine the relative accessibility of these modules.

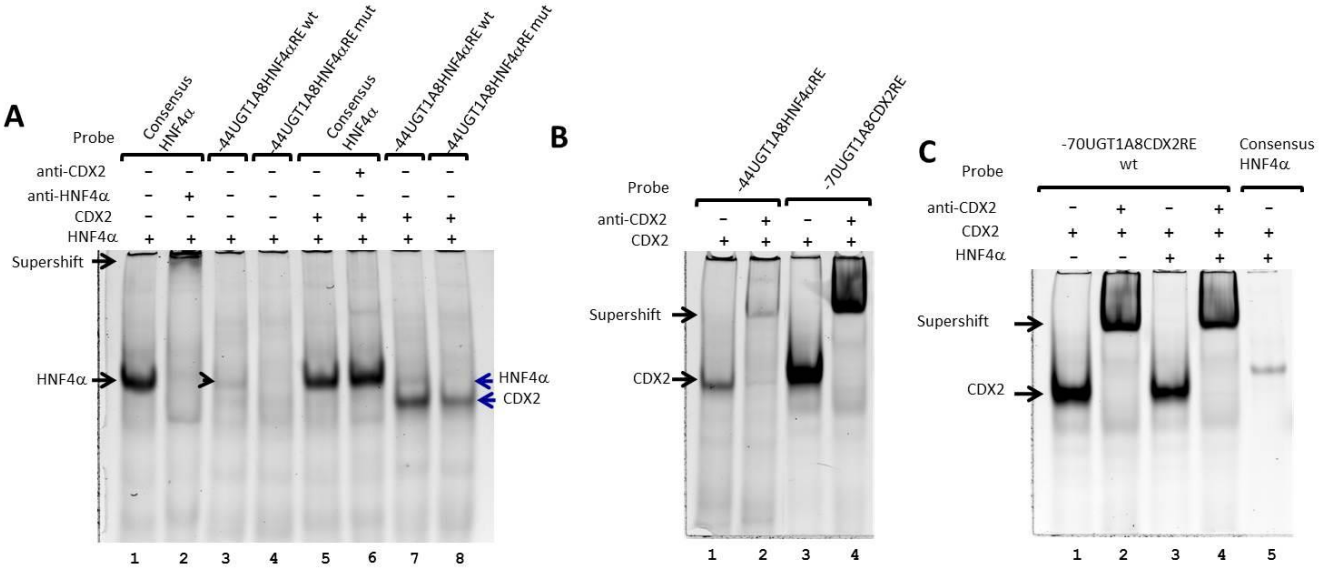
Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 1



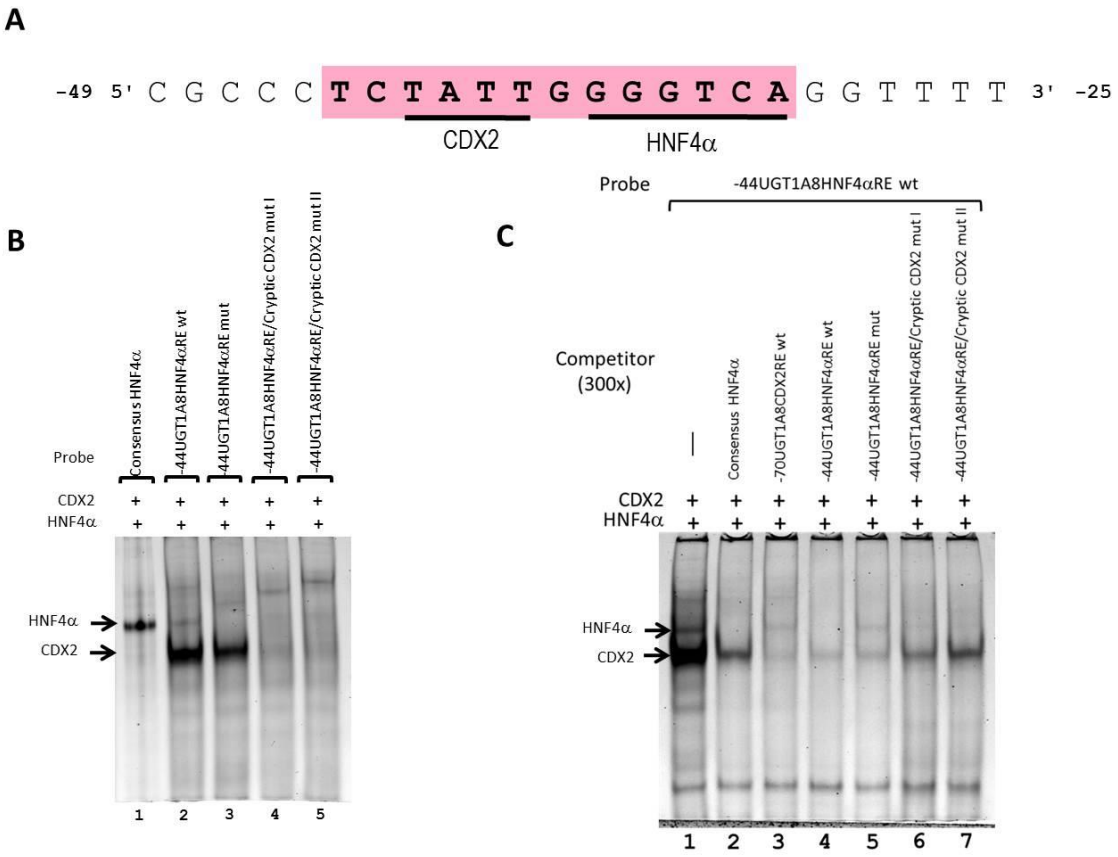
Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 2



Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

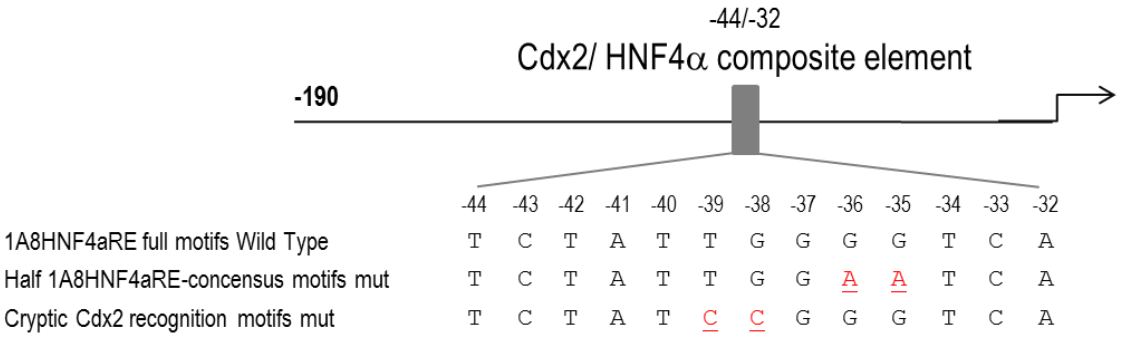
Fig 3



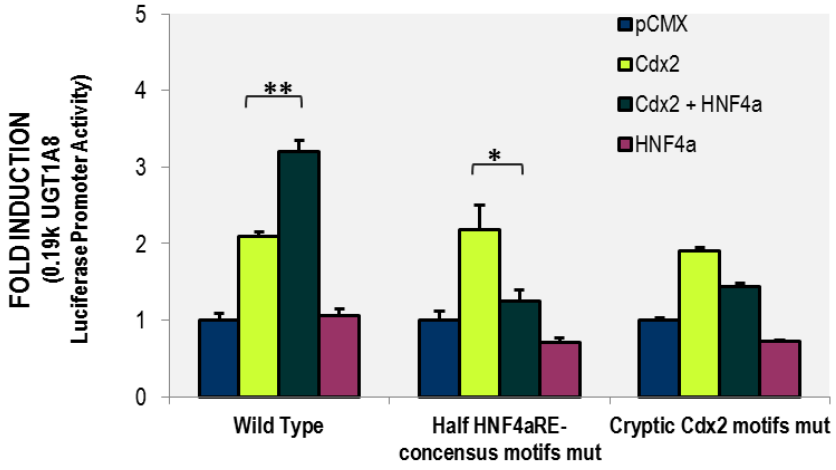
Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 4

A

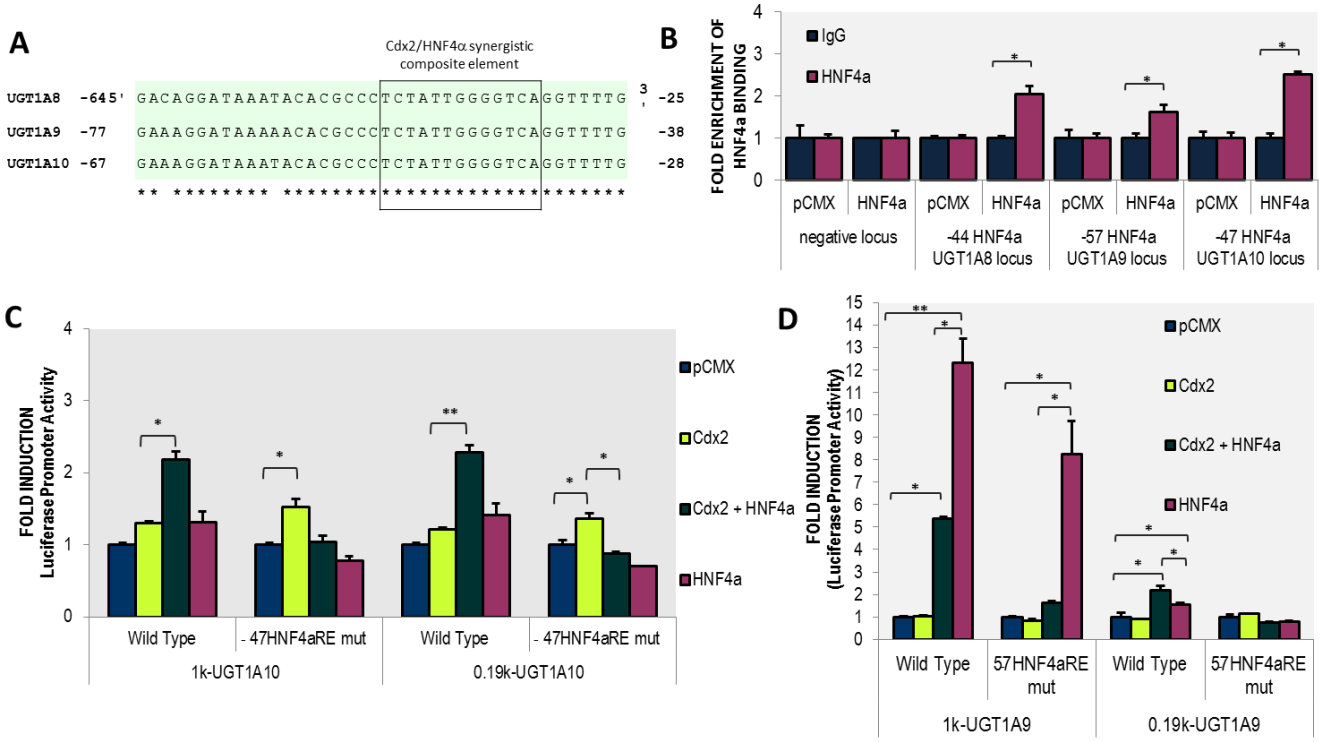


B



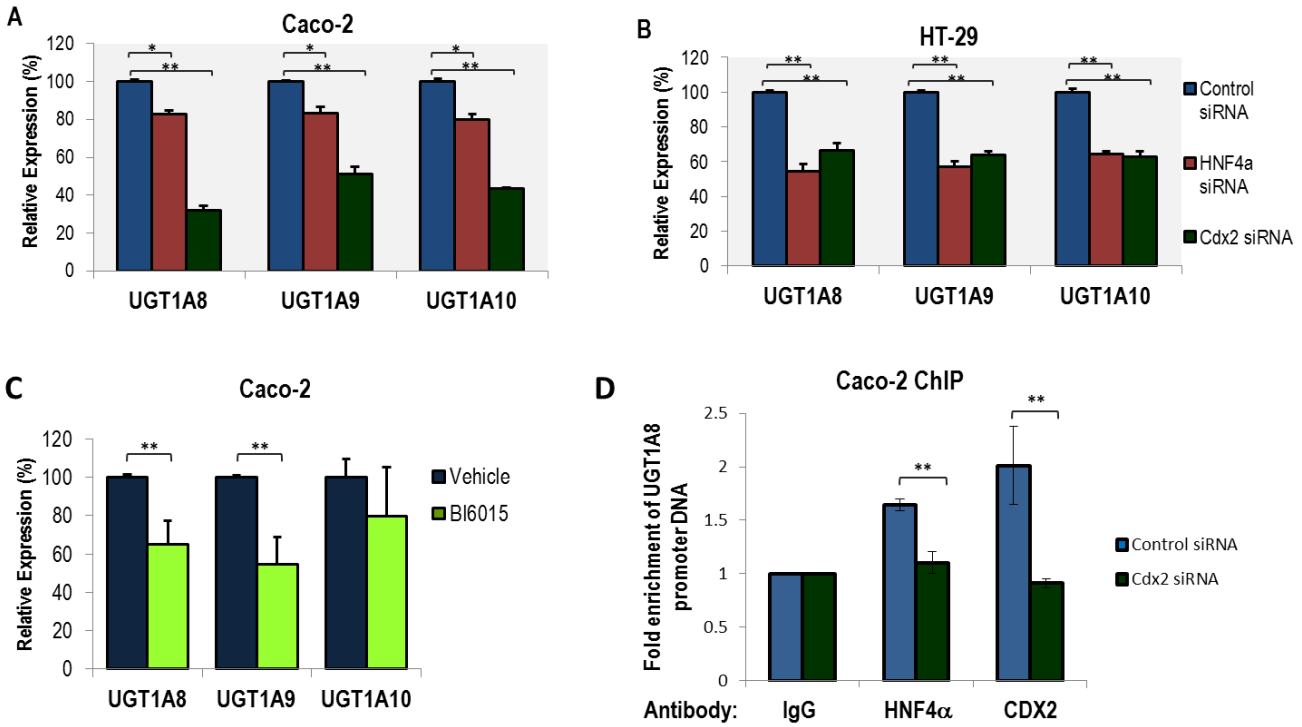
Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 5



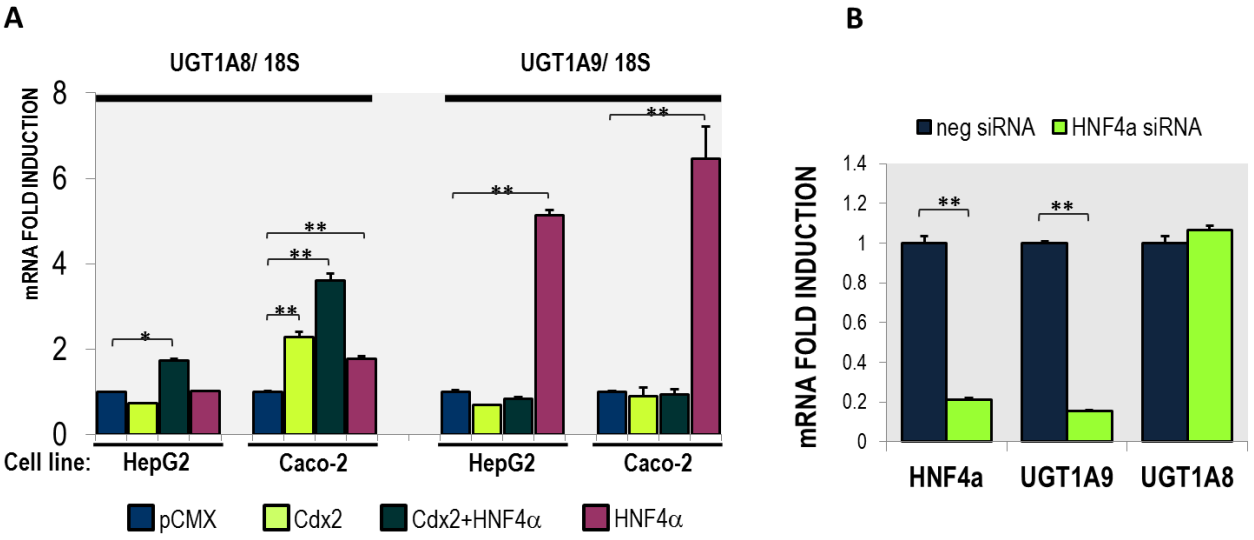
Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 6



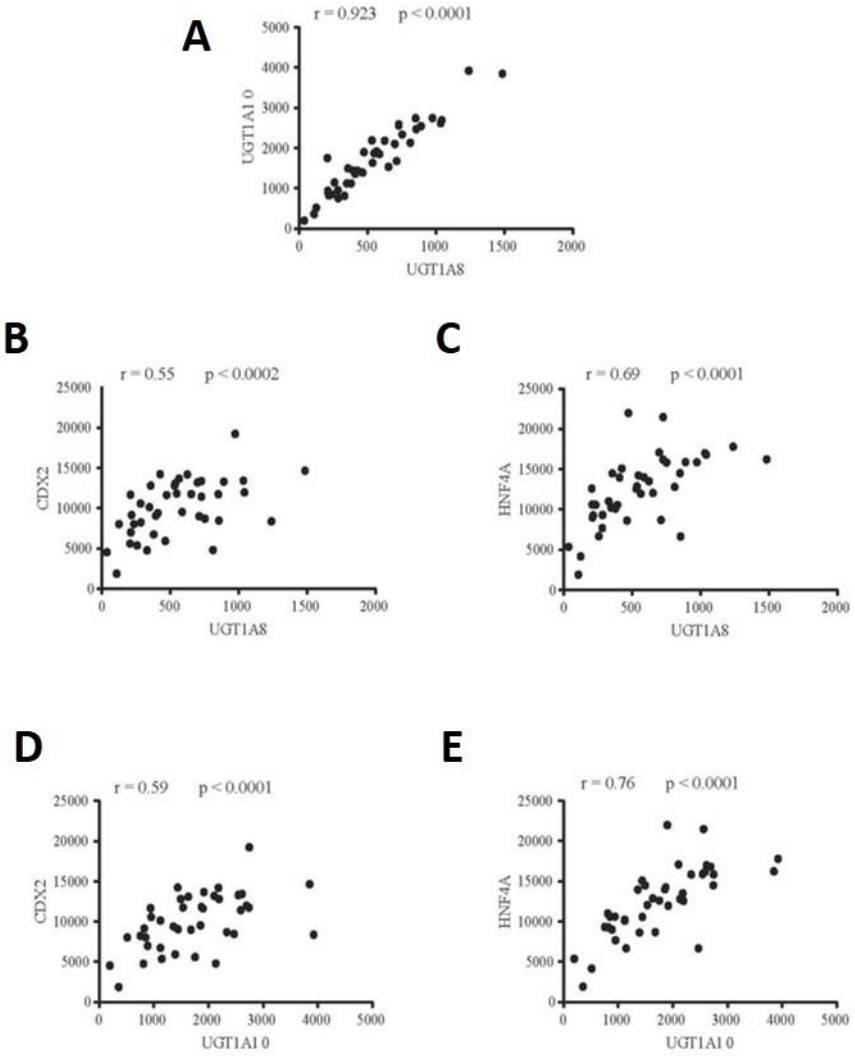
Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 7



Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 8



Molecular Pharmacology Fast Forward. Published on March 8, 2018 as DOI: 10.1124/mol.117.110619
This article has not been copyedited and formatted. The final version may differ from this version.

Fig 9